AN INVESTIGATION OF AUDIO SIGNAL-DRIVEN SOUND SYNTHESIS WITH A FOCUS ON ITS USE FOR BOWED STRINGED SYNTHESISERS

by

CORNELIUS POEPEL

A thesis submitted to the
University of Birmingham
for the degree of
DOCTOR OF PHILOSOPHY

Department of Music
The University of Birmingham
June 2011

Contents

List of Figures

List of Tables

List of Abbreviations

ADAFX	Adaptive Digital Audio Effects
ASDSS	Audio Signal-Driven Sound Synthesis
CBVA	Computer-Based Viola
CWM	Cluster Weighted Modelling
FM	Frequency Modulation
MSSM	Modified Single Sideband Modulation
PD	Parameter-Driven
PDSS	Parameter-Driven Sound Synthesis
PM	Phase Modulation
Rel	Reliability
SD	Signal-Driven
SFM	Simple FM
SS	Subtractive Synthesis
Val	Validity

Acknowledgements

This work would not have been possible without the support and help of many people.

I especially want to thank my supervisors Scott D. Wilson, Erik Oña and Jan Borchers for many hours of reflection, discussion and constructive support, and for their hospitality.

Many thanks go to my wife Damaris, who has encouraged me so wonderfully, put up with long periods of intensive work and kept everything running smoothly. Thank you Oskar and Amrei for your lively smiles, and for the fresh air and inspiration you continually brought to your daddy's mind.

I want to thank Shivaun Heath for doing a marvellous job copy-editing the script. I want to acknowledge Carter Williams and Günter Marx for their interest and for the time invested in working with my developments. Special thanks also go to my father Ewert Pöpel, who helped with language support and philosophical discussions. Further thanks go to the Academy of Media Arts Cologne, Georg Trogemann, Martin Nawrath, Jochen Viehoff, Lasse Scherffig, Christiane König, the Music Academy Basel, the University of Birmingham, the Benedictine monks of Königsmünster Abbey in Meschede who housed me for quite a while to work on the thesis in silence, Charlotte

Stalter-Alexandridis, Garth Paine for various helpful discussions, Dan Overholt for many helpful conversations, Dieter Braun, Eric Lee, Jin Hyun Kim, Carolin Demuth and Carsten Röcker for their scientific support, Ernst Lichtenhahn, Werner Krick, Holger Recker and Heidi Schelhowe for their encouraging reflections on my thesis, Rainer Bühl and Thomas Gwosdz for their supportive assistance, Martin Jann and the company Native Instruments for donating an FM8 and all the members of *hot_strings SIG* for various discussions and for the support they gave by generating the circumstances necessary to pursue the work presented in this thesis.

Abstract

This thesis proposes an alternative approach to sound synthesis. It seeks to offer traditional string players a synthesiser which will allow them to make use of their existing skills in performance. A theoretical apparatus reflecting on the constraints of formalisation is developed and used to shed light on construction-related shortcomings in the instrumental developments of related research. Historical aspects and methods of sound synthesis, and the act of musical performance, are addressed with the aim of drawing conclusions for the construction of algorithms and interfaces. The alternative approach creates an openness and responsiveness in the synthesis instrument by using implicit playing parameters without the necessity to define, specify or measure all of them. In order to investigate this approach, several synthesis algorithms are developed, sounds are designed and a selection of them empirically compared to conventionally synthesised sounds. The algorithms are used in collaborative projects with other musicians in order to examine their practical musical value. The results provide evidence that implementations using the approach presented can offer musically significant differences as compared to similarly complex conventional implementations, and that—depending on the disposition of the musician—they can form a valuable contribution to the sound repertoire of performers and composers.

Chapter 1

Introduction

Bowed stringed instruments have a long tradition in the history of music. Looking at their development, one can see that the knowledge and materials of the time have often had an important impact on the ongoing development of their constructional principles (Schleske, 2004, p. 1). At present, stringed instruments, along with other classical instruments, have an image of being rather old fashioned and representing a musical culture from the past. Music that is considered "up to date" tends to be produced by DJs, laptop performers or pop musicians using instruments like electric guitars, drums or keyboards.

However, music academies and universities with a governmental mandate to educate today's musicians do still train them them to play classical instruments. And despite the decreasing audience, traditional string music is still seen by many to represent an important cultural heritage. These observations raise the question of whether—and if so, how—the connection between traditional string playing and contemporary materials such as computers, electronics and sound synthesis can be made.

In my experience, many traditional musicians are open to increasing their repertoire of musical styles. However, the possibilities for a successful musical expansion are relatively limited. Furthermore, few highly ranked performers exist who are well known for an artistic profile incorporating performances with computer-based stringed instruments. The reasons for this situation do not lie only in a lack of repertoire, a lack of places to perform new music, or the need to be open to improvisation. In my view, this is a problem which is also connected with the available instruments and their sound possibilities.

While studying the viola, I knew fellow students who wanted to play in rock bands and other popular musical forms. Using traditional acoustic instruments or electronic variants did not lead to results which were convincing enough to continue with such attempts. One argument was that the sounds achieved were not suitable for a band. Another argument was that the instruments did not allow the desired subtle methods of playing. This missing subtlety or sound quality in electronic instruments is reported by other musicians too (Zender, 1991, p. 36). In order to bridge the gap between stringed instruments and synthesisers, the company Zeta Music (Zeta, 2010) introduced a MIDI violin in 1984, followed by other instruments from the string family (Graesser, 1996, p.126). Though it offers the possibility of triggering sounds from MIDI synthesisers, the response of the Zeta instrument to the playing technique of a string player is, for many musicians, inadequate, because the system tracks frequency and amplitude only. Thus, its playability is felt by many to be limited. The playability of an instrument is understood here as the ability of the instrument to react adequately, in terms of traditional bowed stringed instrument playing methods, to the input applied using existing instrumental skills.

The computer music research scene has shown a number of developments in the area of stringed instruments in the past. Many of these have tried to overcome the limitations of the Zeta instruments. However, there is still no wider use or tradition in the field of stringed instruments which incorporates the new developments. Since both new instruments and string players willing to work with new sounds exist, the question presents itself as to why these new developments are not used more widely. Do the instruments still lack essential qualities? And if yes, what are the reasons for this lack? What opportunities can be found to overcome the current limitations?

These questions are central to the present research. In conducting the research I have pursued the following objectives: description of related computer-based bowed stringed instruments; discussion of the construction principles of synthesisers, and particularly string synthesisers; identification of possible problems resulting from conventional construction principles; identification of reasons for the wide use of conventional construction principles; proposition, implementation and investigation of an alternative which I have developed and which I believe may be fruitful in the future: the method of *Audio Signal-Driven Sound Synthesis (ASDSS)*.[1]

[1]The term "Audio Signal-Driven Sound Synthesis" was coined by me and Roger Dannenberg while working on the paper entitled "Audio Signal Driven Sound Synthesis" (Poepel & Dannenberg, 2005).

After a more detailed explanation of my motivation for this work in chapter 2, I will present a theoretical apparatus in chapter 3 outlining steps in the design of interactive digital audio systems. These basic steps include the process of abstraction, the building of models, and formalisation, which are fundamental in designing computer-based instruments.

I use these basic steps as a point of departure from which to look at related research, and in chapter 4 describe the most relevant string synthesiser developments that research in computer music has generated in the past 40 years. Since the use of the audio signal to drive a synthesis engine plays a central role in this dissertation, I also present current research which incorporates the audio signal in the sound synthesis process. After this description of related research work, a discussion follows which raises several questions about related research approaches with regard to the use of limited models, problems of formalisation and the tendency to use a physically orientated view of instruments and musical performance.

In order to set the scene for more detailed investigation into the reasons for the shortcomings in conventional approaches to Parameter-Driven Sound Synthesis (PDSS),[2] a description of PDSS is given in chapter 5. By Parameter-Driven Sound Synthesis I understand a method of sound synthesis in which the control applied by the performer to the sound result is based entirely on discrete parameters. These parameters form the pillars on which the sound control is based. Since the objectives of this thesis include the

[2]The term "Parameter-Driven Sound Synthesis" was coined by the author.

identification of shortcomings in the conventional approach, and the search for reasons for the wide use of this approach, a more detailed investigation into historical aspects of sound synthesis and of systems available with regard to the formalisation and measurement of performers' actions is provided.

I am a trained violist,[3] and the developments are designed for use by colleagues and by me. Since it is intended that the newly developed instruments should meet the needs of musicians, it seems to me that it is important to incorporate into this thesis a strong emphasis on my view as a musician. I am aware that physical and functional constraints must necessarily be considered when computer-based instruments are being developed. However, as the decisive assessment of the function or non-function of an instrument is judged here to come from the target users, the musicians' point of view should also be heard. In addition, because my aim is to build instruments which allow performers to make use of their existing instrumental skills, the following section then describes different approaches to violin playing from the point of view of instrumental teaching. The description is used as a basis from which to draw conclusions for the requirements of bowed stringed synthesisers.

One problem that I identify in the design of computer-based instruments is in the use of limited models of performance. The question is therefore considered as to how alternatives to building the interfaces on the basis of predefined and limited models might look. In order to offer an alternative to

[3] I studied the viola with the well-known string teachers Jürgen Kußmaul and Hatto Beyerle.

conventional approaches, the method of Audio Signal-Driven Sound Synthesis (ASDSS), the main subject of investigation here, is proposed and explained in section 5.4. In contrast to PDSS, which needs to define and formalise an input system capable of measuring all essential actions and playing parameters of the performer, ASDSS aims to keep the essentials of performing non-formalised, but to be transparent for the phenomenological parameters that a musician puts into the audio signal when playing. By a phenomenological parameter I understand a parameter that is perceived and used by the musician, independent of the question of whether or not physical descriptors for this parameter are available. If physical descriptors exists the perceived phenomenological parameter may differ from the physically measured parameter value.[4] Additionally in chapter 5 the basic constructional principles of ASDSS are described.

Since the method of ASDSS is almost unknown and unused in sound synthesis, and the current body of research provides only a very small number of publications addressing this approach, a central question of this research is that of whether the ASDSS approach can show musically significant differences as compared to the parameter-driven approach. The main hypothesis of the thesis as presented in section 5.5 addresses this question.

Beyond the investigation of the signal-driven approach, it is my goal to explore how many different variants I can achieve by implementing different

[4]An example can be found in measured temperature (physical descriptor) and individually felt temperature (phenomenological parameter). Examples of musically phenomenological parameters would be tempo, warmth and liveliness of a tone, and intensity of a tone.

synthesis methods based on signal-driven construction principles. Two algorithms I developed for a string synthesiser in my Masters-level work were already available (Poepel, 1999, pp. 21-22). The development of further algorithms, their structure and basic proportions in the sound result are described in chapter 6. These algorithms utilise the raw and unanalysed audio signal of the electric instrument to drive the synthesis engine. Since the raw audio signal is used, conventional methods of sound synthesis must be modified to suit the needs of the new method. Physical parameters such as frequency and amplitude are extracted from the audio signal, but are used to modify the synthesised sound indirectly. After a description of third-party developments of signal-driven synthesis algorithms, the hardware and methods used to capture the instrument's audio signal and to extract explicit parameters are presented. Since the implementation, personal tests and analysis of the method's functioning exhibited remarkable properties, these findings are described in section 6.5.

Chapter 7 makes use of signal-driven synthesis algorithms to investigate the main hypothesis. In order to compare the two approaches to synthesis, a player-based study and a listeners' study were conducted. The listeners' study compares four different instrumental sounds in relation to several musical parameters. The four sounds include parameter-driven and signal-driven synthesis methods, and were generated using exactly the same input signal. The musical parameters selected for evaluation were those which have been found in music psychology literature (Juslin, 2001, p. 316) to be relevant for determining musical expression. They relate to timing, pitch, dynamics, articulation and timbre.[5] A two-sided comparison was used to establish which of the two sounds compared was assessed to offer a better result in relation to a specific musical parameter. The study measured whether the listeners perceived statistically significant differences in this comparison.

The results of the study demonstrate that implementations using the signal-driven approach can indeed show musically significant differences in sound result as compared to similarly complex implementations using the parameter-driven approach. Beyond the empirical approach to investigating the method of ASDSS, personal experiences and statements by performers and composers using the signal-driven algorithms in practical music use are provided and discussed in chapter 7.

Since this method uses an instrument's audio signal to drive the synthesis engine, I assume it can also be used with other instruments. Ideas for future work, including expanded versions of the presented applications, are thus described in chapter 8.

From a musical point of view it is often considered helpful to accompany the written research with audio and video examples. Therefore, in addition to details of the studies' results, the appendices include two audio CDs presenting sounds generated using the ASDSS methods, sounds used in the listener study, and recordings of two compositions which make use of the presented synthesis methods. In addition, a DVD is included which presents

[5]The validity of these parameters for musical expression is, of course, dependent on musical and social context. In order to conduct a study with scientific evidence, however, I had to use a foundation offered by related disciplines.

a video documentation of a composition incorporating ASDSS, and documentation of performers testing a signal-driven viola synthesiser. Finally, a CD-Rom is included which provides applications allowing the reader to test the signal-driven synthesis methods presented in chapter 6.[6]

Bibliographical references are written in APA style (American Psychological Association, 2009) because this style is well known and precisely documented, and is used in literature on interactive music systems.[7] For the sake of gender-neutrality I will refer to performers, user, participants etc. in the male and the female form. Italics are used in the text to highlight important terms. Names of technical devices are written in italics the first time they appear.

---

[6]These applications are implemented in the audio programming environment MaxMSP, which is available on the website of the company Cycling74 (2010b). MaxMSP patches and standalone versions of the patches are provided.

[7]An example may be found in Ng and Nesi (2008).

Chapter 2

Motivation

In order to explain the motivation driving me to carry out the present doctoral research, chapter 2 will describe firstly personal experiences I had with synthesisers and traditional instruments, secondly areas where I see potential for the development and creation of synthesised music and thirdly aspects of music related human-computer interaction that aroused my interest for investigation.

2.1 Personal Experiences

While testing keyboard synthesisers for use in rock bands in the 80s, I was overwhelmed by the sound possibilities offered by the synthesisers of the day. On the basis of my experience as a violist, I expected to be able to acquire an increasing facility to shape the sounds expressively. Although the synthesiser sounds were initially so impressive, however, I reached a blind alley after some months: the possibilities for sound manipulation were exhausted, and the connection between me as the player and the sound quickly reached a limit. Although it was familiar to me from my experience of the viola and the piano, the anticipated ability to get the sound “into my hands” or to shape the sound according to a musical idea or the demands of a live musical context could not be achieved. It felt like a dead end in the process of increasing expressivity by modulating the sound more and more thoroughly, in order to achieve musical goals.

Variations in timbre or articulation had to be pre-programmed. Nearly every change in sound, other than pitch and volume, had to be defined in advance, with no possibility to react to the ongoing flow of music which I was used to dealing with when playing the viola, the piano or the electric guitar. The use of tone wheels, sliders, knobs, aftertouch, foot pedals and the Buchla lightning controller[1] was initially found to be a help in designing the sound in real-time and creating more flexibility. However, while these extensions enabled more parameters to be controlled in real time, the fundamental goal of being able to modify the sound during performance in a subtle and meaningful way by systematically exploring and mastering the instrument could not be achieved, and the perceived dead end did not disappear.

In comparison to keyboard synthesisers, this problem became much more obvious when using the Zeta MIDI string instruments. The ability, gained from long years of training, to generate specific articulations, colours in tim-

[1]The Buchla lightning MIDI controller is a device that measures the position and movement of handheld wands and transforms generated measurement data into MIDI signals (Buchla, 2010).

bre, energy in the tone and the flowing co-ordination of bow, fingers, breath and muscle tension and relaxation in order to achieve a fluid working body of playing techniques, was relatively ineffective with regard to the sound output. Moreover, it was necessary to serve the pitch tracker with specific sounds and articulations in order to enable the tracking algorithm to generate the proper MIDI note-ons necessary to trigger the sounds I desired. The wide pool of sound presets that could be used with MIDI technology was, of course, available. However, in testing the instrument, it was almost immediately clear that the spectrum of variety inside a sound (i.e. timbral variation) was radically reduced in comparison to that of an acoustic instrument.

An aspect relating to the field of an instrument's playability may be mentioned here. In terms of flow theory as proposed by Csíkszentmihályi (1990), an optimal experience is achieved when a dynamic balance between abilities and challenge is present. The flow can be hindered if a person's abilities are not sufficient to master the challenge of a task and the person feels overextended. When performing with a synthesiser, the challenge is to produce the musical results the performer has in mind. If the ability of the player to create this specific music is hindered by restrictions caused by the instruments, the flow is hindered.

In contrast to my experience with the synthesiser, I felt much closer to a flow experience when working with the viola because the instrument allowed for a continuous increase in the potential for musical expression, without running into a dead end—assuming the musical idea could be achieved within the sound a viola can generate. In learning the instrument, more and more features appeared which it was necessary to master. While initially any kind of bowing was sufficient, after a while three bowing parameters—bow speed, arm weight applied to the bow, and bow position relative to the bridge—became important. Flexibility of finger joints, amount of bow hair on the string and left-hand finger pressure were aspects—to name just a few—that had to be learnt. And there was (and is) always the possibility of incorporating personal additions to the methods of playing in order to give the sound a specific personal timbre or style of performance, for example.

Since I am referring here to musical aims or ideas, one might ask what kind of ideas I am talking about. However, the territory of specific musical goals towards which any musician may aim is wide and cannot be defined in advance because it is the performer who determines what she seeks to achieve.[2]

As the experiences described here are subjective and personal, it is interesting to consider whether there are other performers who have had similar experiences. The experience might simply be an illusion and have more to do with the player than with the instrument. Assuming, however, that more of these kinds of experiences can be found, a second question might be how this feeling of limitation comes about.

[2] I myself am interested in incorporating a kind of breathing, liveliness and depth into the sound. Therefore I want to get the sound as closely "into my hands" as possible, to "go into the sound", explore its structure and learn how I can shape and modify it so that I can fit the sound to the musical needs that arise while I am performing. This might be a slowly fading sound, for example, or an aggressive sound, a very thick and inflexible sound, etc.

2.2 Using Potentials

It is well known that one of the factors responsible for the widespread use of computers today is the improved usability which has been achieved by adapting operating systems, applications and computer hardware to the user. This adaptation focused on the use of the already available user skills. As an example the WIMP paradigm (windows, icons, menus, pointers) may be mentioned.[3] Offering a desktop with files that could be opened, manipulated, saved in a folder or put into a dustbin, meant that much of the users' pre-existing knowledge of how to deal with everyday things could be used. This usage created a win-win situation by using the potential offered by computer technology as well as the potential offered by the users.

Building new musical instruments involving computer hard- and software can be seen as a similar thing. As explained in the introduction, traditionally trained musicians embody a huge musical potential. The question is how this potential might come together with the potential of computer technology. First one must discover whether there are indeed musicians open and willing to investigate approaches to computer-based instruments. In the case of stringed instruments, one can say that there are indeed such instrumentalists. Examples are Yo-Yo Ma and Joshua Bell (section 4.4.2). As explained later in section 7.6, Hatto Beyerle, a founding member of the Alban Berg String Quartet, and I have formed a group called *hot_strings SIG*, which provides an

[3] An early usage of the WIMP paradigm is found in the "Star User Interface" developed at the Xerox Corporation (D. C. Smith, Irby, Kimball, & Verplank, 1982).

organisation for exactly such musicians who are interested in the expansion of stringed instruments using contemporary materials.

The question of *how* computer technology might be structured to make use of the potential of traditional musicians is an important question for many researchers. Examples of such researchers are Poupyrev, Lyons, Fels, and Blaine (2001, p. 2), and Marrin (2000, pp. 16-17). This question is a main driving force behind the present research too.

This research focuses on string players and their specific potential, and the question of how their pre-existing skills might be used to play synthesisers. Since the instrument's audio signal plays a central role in the transmission of the actions of the player, I assume that this technique will also be fruitful for other instruments and musicians. It is my goal to build instruments that make use of both the potential of traditional musicians and the potential of computer technology.

2.3 Human-Computer Interaction in Music

The aforementioned WIMP paradigm is an example of a successful outcome of research in human-computer interaction (HCI). This huge research field is represented by conferences such as the ACM Conference on Human Factors in Computing Systems (CHI, 2010) and the ACM Symposium on User Interface Software and Technology (UIST, 2010). In 2001 the Conference on New Interfaces for Musical Expression (NIME, 2010) evolved out of the CHI conference in order to focus on music-related human-computer interaction.

Building interfaces for musical purposes is a very particular type of interface development. In comparison to digital computers, music and musical instruments for performance have a long tradition. This history includes different cultures as well as an ongoing redefinition of the musical languages over human generations. Besides this, music is known as one of the most basic forms of expression used by human beings to deal with all aspects of life itself. Music is often seen as an area where the contradicting elements of mind and matter meet. It has always been a highly interesting field for research, principles of organisation, and for training, control and mastery. Furthermore, it is seen as a field which includes the unknown, involving phenomena which cannot adequtely be described in words (at least in some aspects). Considering these issues relating to music, it is clear that developing musical interfaces can be seen as a task of a different nature from, for example, developing the interfaces of mobile phones.

The NIME conference focuses on the design of novel controllers and computer-based applications for the very particular type of human involvement which is music. It is one of the goals of the conference to "identify current and promising directions of research and unsolved problems," and "to focus on the major practical concerns involved in the design of interfaces for musical expression" (Poupyrev et al., 2001, p. 1). Besides this, it is a goal of the conference to "identify key interface technology developments that offer the most exciting new opportunities for musical expression."

As explained in section 5.4, the approach to the interplay of interface and synthesis investigated in this research differs in key aspects from conventional approaches. This different approach means on the one hand that it is difficult to find a place for my research in this research field. On the other hand, this provides a particular motivation to take on the challenge of investigating a more or less unexplored approach to sound synthesis. However, mastering a challenge would not in itself be an adequate reason to go in this research direction. A survey of the available literature on the subject suggested that the majority of researchers in this field are focusing on the conventional approach to sound synthesis as described in section 5.1. In studying this approach I was confronted with a set of remarkable problems that motivated me to consider them more thoroughly and to take a different route to improving computer-based instruments. These problems were as follows:

- In the conventional approach to sound synthesis, traditional instruments are often described as "controllers" driving a synthesis engine. In contrast to a controller such as a MIDI device, however, acoustic instruments do not have a *formal input*. The formal input of a digital instrument is understood here as a set of sensitive devices that are built to extract values for a predefined set of input parameters based on the actions of a performer. These devices send discrete streams of parameter data on the basis of the measurement results. It is crucial to understand that this pre-definition is based on a *model* of the instrument, which is necessarily distinct from the instrument itself. Traditional acoustic instruments can be described as formal systems (as in physics), but the formal structure of an instrument's model is not necessarily the complete instrument that a performer relates to. When measuring the actions of a performer, the question of which kind of input should be measured, and how an algorithm should react, may differ depending on the playing techniques applied. Which playing techniques are considered relevant or meaningful is generally dependent on the performer and the musical context. Scratch tone was not a common playing technique in the classical era, and a computer-based string instrument designed for classical music would not need to identify scratch tone, or track the ways in which it might be varied. In contemporary music, however, a performer may use scratching as a musically meaningful action, perhaps even one which has a significance distinct from other similar sounds, such as accented note onsets. The success and validity of a model are similarly context dependent: a model derived from classical string technique might not include a specific input or inputs for measuring or distinguishing scratching, while a model derived from contemporary music possibly would.[4]

- Though not usually completely unrelated, many descriptions in player-instrument research do not map well to those found in instrumental teaching. For example, bow tracking technologies are intended to measure bow pressure (here applied or contact force at a certain point of a physical system), while players are often taught to play not with bow pressure but with the weight of the arm. While it is tempting to assume that the two are equivalent, or at least roughly so, by doing so

[4] The building of models will be described more thoroughly in section 3.6.

we risk losing musically meaningful input that, while not immediately tangible, is nevertheless implicit in playing with the weight of the arm. Pedagogical descriptions are not necessarily geared towards building models that digital systems can deal with. This is not surprising, given that pedagogical descriptions are related to perceived phenomena and not to measurement results of physical parameters, but this does not mean that they are meaningless or imprecise.

- In the field of computer music research, meaningful playing techniques and parameters understood by musicians—such as “playing more intensely”, “digging into the string”, “grabbing the string with the bow”, “pulling the note”, “being in touch with the tone”, “warmth of tone” and “liveliness of tone”—are often described as “abstract” playing actions. In terms of what “really” happens, research into concrete physical correlates can be found. If such actions cannot be described physically they are often ignored in the process of developing instruments in many conventional approaches, despite the fact that musicians really use these playing techniques.

- In my experience, when testing a Zeta MIDI violin, many musicians become frustrated with the instrument relatively quickly, and some find that it is not usable at all. This happens despite the fact that it tracks the two most important parameters—pitch and amplitude—and offers the huge potential of all the different sounds of MIDI synthesisers. In such a case, where the most relevant parameters are tracked and many new sounds are available, one would expect a positive attitude to synthesisers. I conclude that an essential quality may be missing in the sound, despite the fact that the apparently most important playing parameters are covered.

- Although there are many developments in the field of computer-based bowed stringed instruments (see chapter 4.2) that have been developed by cutting-edge institutions such as the Massachusetts Institute of Technology (MIT, 2010) and used by top performers such as Yo-Yo Ma (Paradiso & Gershenfeld, 1997, p. 76), and Joshua Bell (Jehan, 2001, p. 5), there is no wider community of string players using such instruments.

These observations motivate me to analyse the design process of the conventional approach to sound synthesis, and the conditions and assumptions found in related research, and to search for possible improvements and alternatives. I believe that it would be valuable for researchers to look at sound synthesis from a range of different perspectives, and to investigate available possibilities which may provide fruitful results in the ongoing renewal of our musical culture.

Chapter 3

Design of Interactive Digital Audio Systems

In parallel to the problems I listed in chapter 2, other researchers point to issues in the development of computer-based instruments which likewise suggest a need for a more thorough analysis of the design process of computer-based musical instruments. The researcher Garth Paine, for example, has made the criticism that, although new interfaces and computer-based instruments have been designed, the number of developments actually used by the target users is relatively small (Paine, personal communication, January 11, 2008). This raises the question of why so many developments have not succeeded. Concerning the wider and standardised use of new controllers, Tod Machover mentions that he is “somewhat uneasy and dissatisfied with the current state of the arts.” In order to improve the quality of computer-based instruments and create what he calls “inevitable” instruments, he proposes to "evaluate our progress, to discuss openly why it is so hard to get to the next step" (Machover, 2002, p. 1).

In order to investigate common construction principles in sound synthesis, I suggest the use of tools which enable the identification of problematic issues in the development process of new computer-based musical instruments. The question presents itself as to how this investigation might be conducted and which tools are likely to be of use in the process of investigation.

The musical instruments that are the focus here are a subdivision of interactive digital audio systems. General constraints which apply to interactive digital systems must be considered when instruments are being developed and analysed. This process falls within the wider area of software engineering and the field of human-computer interface design. Therefore, methods used in this area can also be examined and used here. In touching now on topics from the fields of system design and software engineering, I do not wish to give the reader the impression that the present work is a piece of research in system design. Rather, this is simply a necessary excursus which is intended to provide the reader with tools and knowledge which will enable problems in related research work to be illuminated and particular aspects of the ASDSS method examined in this thesis to be described more clearly.

Computer systems are formal systems. Thus, the process of designing interactive digital audio systems necessarily involves formalisation. In order to formalise objects, models of these objects are developed and abstraction is used to build the models. In order to give the reader an insight into these topics, I present a design model with stages and transformations between stages as are often used in software engineering. The stages of the design model are debated and the issues of abstraction, models and formalisation are addressed in the following sections. This is intended to enable the reader to view the related research described in chapter 4 with a sensibility for the problems that may occur in formalisation. Section 4.7 then discusses problems revealed by this specific perspective on related research.

Acoustic instruments differ from computer-based musical instruments because acoustic instruments do not involve formalisation or offer a formal input (see definition of formal input on p. 17). By studying the question of how formalisation is achieved, I aim to reveal assumptions made in existing approaches, use the outcome to improve existing systems and argue for alternative strategies in constructing computer-based synthesisers.

Although it is not certain that this particular design model was used by the developers of the related research which will be described in chapter 4, since all related research work will have dealt with formalisation (no matter whether it was explicitly addressed by the developers or not), and since the model includes steps that are necessary in formalisation, the model is considered here to be of use in analysing core steps in the development of related research.

3.1 From Idea to Application

In the realisation of any application, there will be goals, intentions and requirements to be met. These requirements can be expressed in words in some form or another. The document containing this information is often called the "requirements document". Once the development has been completed, an interactive application includes, at the very least, a computer program, the code and a user interface including both a soft- and hardware surface.

What happens in between initial idea and application? How can the intermediate stages be defined and organised? These questions and the different strategies for organising this process are addressed in software engineering. The focus in software engineering is primarily on producing applications that will work properly and be commercially successful. It would be off-topic to describe those here, but the reader may find descriptions of strategies in Sommerville (2004), for example.

The "waterfall model" is one of the "general models or paradigms of software development" (Sommerville, 2004, p. 9) and is often considered the classic approach to the design of computer applications. It divides the development process into several stages organised in a sequential and linear order. This method has distinct goals for each level in the development process and provides a document for each level once the work on that level is completed.

I am aware that the waterfall model is seen by many researchers and developers as old fashioned and unsuited for the fast production of highly saleable applications. However, since this model is not only still used widely, but is also seen as the classic approach, I will refer to it here. I do not intend to present an in-depth and detailed description of the waterfall model, nor do I necessarily recommend using the waterfall model for the design of interactive applications. However, in order to enable a clear identification of steps in the design process, I divide the process of development into the following stages: requirements, specification, formal specification, implementation. Figure 3.1 illustrates these stages. The important further step of verification is discussed in section 3.8.

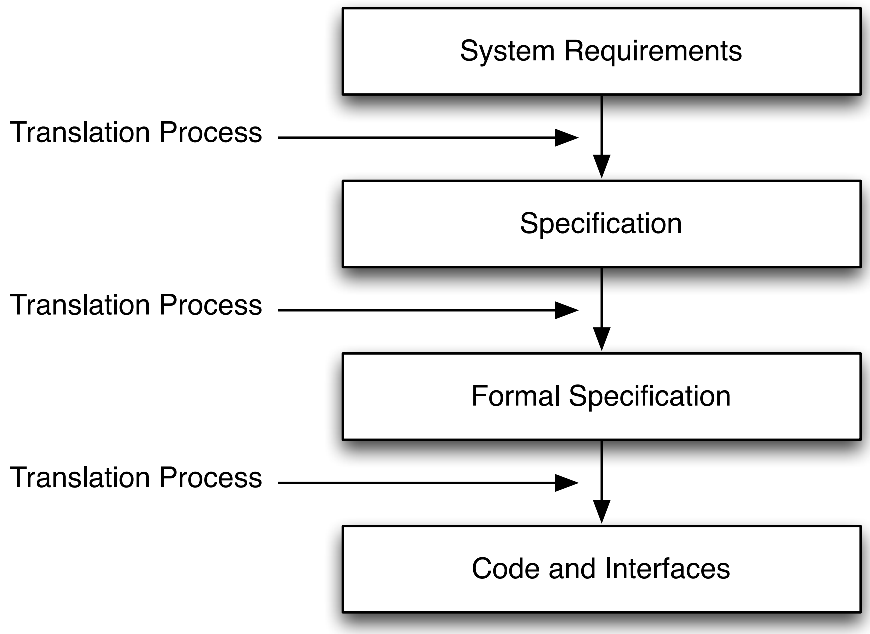

Figure 3.1: Levels and transformation in the development process.

The specification is a transformed version of the requirements document that can provide elements such as an outline of the application or abstractions which can be used to implement objects or methods mentioned in the requirements document. In the formal specification the application concept and the abstractions are translated into a formalised version that in turn can be translated completely and unambiguously (losslessly, i.e. without losing any element, property or function) into the code and interface.

Looking at the illustration in figure 3.1 and at a given application, two questions can be asked:

1. What is in the boxes or written in the document at the end of a specific phase?
2. How is the process of transformation from one level to the next level done?

The following sections will offer a more detailed look at the requirements, the specification, the formal specification, and the transformation between these in order to use these stages later on in this thesis.

3.2 System Requirements

The system requirements document includes essential information relevant to the creation of the application. It often contains elements explained in an everyday and non-formal language.[1] I refer to this everyday language here as "natural language". A template of a system requirements document for the purpose of designing a computer game can be found on the website of the company "Runaway Studios" (Taylor, 1999). Examples of questions one can find in this system requirements document for a computer game are: what is the aim and philosophy of the application? What is the game? What does the user control? What does the world of the game look like? Which game engines are used to generate the world? What is the main focus of the game?

Since musical instruments and, in particular, bowed stringed instruments are the focus here, I provide examples of elements below that could be included in a system requirements document for a synthesiser to be played by a string player. These examples are not intended as an exhaustive list, but provide qualitatively varied and context-dependent items to illustrate the variety of elements that can be expected in such a requirements document.

- The instrument should allow double stops.
- It should be able to transmit bow pressure and speed.
- It should allow the user to generate a brilliant tone similar to that which can be produced on a high-quality violin.

[1] A formal language is understood here as a finite and "organized set of symbols which can be precisely defined in terms of just the shapes and locations of those symbols, without any reference to any meanings or interpretations" (Mastin, 2010). In contrast, a non-formal language uses a set of symbols with meanings and including the possibility of interpretation. It is not finite, in the sense that new words can be created, for example, and that the meaning of sentences can include contents which may be read "in between the lines". For a more thorough description of the distinction between a formal and a non-formal language, see Mateescu and Salomaa (1997, p. 1-39).

- It should be open to the personal style of expression of a player.
- It should be open to and convey the intentions of the performer.
- It should handle pitch, volume, tone colour and articulation.
- It should be able to preserve the character of the timbre a musician is producing.
- It should allow the performer to make use of pre-existing skills.
- It should be connectable to MIDI devices.
- It should offer the methods of FM and additive synthesis as well as physical modelling.

While some of these requirements are related to the technical structure of the instrument (for example the MIDI compatibility or the need to offer FM synthesis), other requirements are related to the actions of a performer using the instrument. By the *performer's actions* I mean all the actions a human performer makes while playing an instrument that are relevant to the realisation of the performance.

I propose to distinguish between phenomenological, physical and formal requirements. A formal requirement would be, for example, the need for a MIDI output. Since the MIDI specifications[2] define the formal structure of MIDI data and the electronics required to generate and transport MIDI data, the question of what formal structure a MIDI output should have has

[2] MIDI Manufacturers Association (MIDI, 2010)

already been answered. This is independent from the question whether the full set of MIDI specifications are implemented in a device or not, since a limited MIDI implementation makes use of the formal structure of the MIDI specification already.

A physical requirement would be, for example, the tracking of frequency. Frequency is defined physically and the requirement could be passed directly to the specification level, but the formalism (the tracking of frequency using the method of an FFT-based monophonic maximum-likehood frequency detector (Puckette, Apel, & Zicarelli, 1998, p. 109), for example) would need to be chosen.

A phenomenological requirement would be a requirement coupled to a phenomenon perceived or imagined by a human. Examples are the needs to: capture the subtle actions of a performer; be open to the expression of a performer; transport the character of tone; allow the player to generate a brilliant tone. Terms such as the actions or expression of a performer, character of tone or brilliant tone are seen as phenomenological because they are actual phenomena for musicians to work with when they are talking about or performing music. They exist, as it were, in the musicians' quality criteria, and are among the requirements of a musical instrument, although they do not necessarily have a one-to-one relation to physical descriptors.

Once the requirements document for the system has been completed, the next level has to be reached. In order to break these requirements down to allow transformation to subsequent levels, an interpretation of the requirements must be made. According to Winograd and Flores (1986, p. 96) this task is critical. Therefore this problem will be looked at in more depth in the following sections. It is known that computer systems function according to formal principles, so I will describe issues of formalisation in the next section, before looking at the process of transformation from the requirements to the specification in section 3.4.

3.3 Formalisation

Following the definition of Hans Herbert Schulze (1989, p. 1281) I understand formalisation in this context to be:

> The transformation of an operation or object into a form in which they each can be described completely and unambiguously by a finite algorithm. [...] Formalisation is related to formatting, i.e. to the unambiguously describable control structure of the relevant magnitudes and values.[3] [Translation by Cornelius Poepel]

With reference to the requirements document for a computer game (mentioned in section 3.2), one can assume that the description of the world the game takes place in (which will be seen on the computer screen) may include things such as: road, gun, house, cloud, mountain etc. In order to formalise the object “cloud”, for instance, one would have to transform this object into a form in which it can be completely and unambiguously described by

[3]Die Überführung eines Ablaufs oder eines Gegenstandes in eine Form, in der sie jeweils restlos und eindeutig durch einen endlichen Algorithmus beschrieben werden können. [...] F. [*sic*] geht einher mit der Formatierung (Format) d.h. mit eindeutig beschreibbarer Regelung der betreffenden Größen und Werte.

an algorithm. This raises the question of whether a cloud can in fact be described completely and unambiguously in this way. In other words: is a cloud formalisable?

Krämer (1988, p. 1) presents three criteria any operation or object has to meet in order to be formalisable:

- Written: it must be able to be described completely in a written form.
- Schematic: it must be able to be organised in a schematic form of procedural operations.
- Exempt from interpretation: it must be able to be described in written symbols of a formal language that have only one proper reading in that formal language.

It is clear that in the case of a marketable computer game one will not be able to formalise a cloud in a way which results in the user being confronted with "any visible mass of water droplets, ice crystals, or a mixture of both that is suspended in the air, usually at a considerable height",[4] as a cloud might be defined. But it is possible to transform the understanding of a cloud into a form suitable to the framework of the application. One might have the goal of presenting something on a computer screen and making the user amazed by the brilliant clouds in the computer game.

[4]Definition of "cloud" found in the New Encyclopædia Britannica, 15th Edition, 2007

In this case one could analyse an image of a cloud, schematise the appearance and parts of a cloud, and write a text that has only one proper reading for the machine (exempt from interpretation) and completely and unambiguously describes the cloud. This formalism can then be implemented in a computer program, and can generate an image on a computer screen which the user will perceive as a brilliant cloud.

The process of formalisation described in the two paragraphs above already includes the levels of specification (here: selection of a model of a cloud), formal specification (here: transforming the model into a form that can be implemented losslessly in a computer language), and implementation (here: writing the code for the program to control the pixels on the computer screen). Thus, in this case, formalisation includes the transformation from the requirements through the specification and formal specification to the point at which the requirements fall into the categories of objects or operations that can be implemented directly. Should elements in the requirements document not fall into these categories (not be formalisable), one would have to replace these elements by elements that are formalisable or one would have to exclude those elements from the requirements document.

3.4 Specification

The specification is the first step in formalisation. Models of the elements to be formalised are defined, the context in which the models are to be valid is set, and requirements assumed not to be formalisable are rejected. An outline of the application, its general structure and its function is drawn up.

I would like to focus here firstly on the definition and use of models. If a model for a specific element of the requirements document is not available, a model has to be generated based on research studying that element. If the intention is to build a musical instrument capturing the expressive actions of a performer, one must first research the question of what the expressive actions of a performer are, so as not to be faced by the problem of capturing something other than the expressive actions. So the first necessary question is: “What is the element to be specified?”

In answering this question, one is confronted with the need for abstraction. Coming back to the aforementioned example of a cloud, the question of what a cloud is must be answered by making an abstraction of an object known as a cloud. As the process of abstraction is crucial for building and properly using models, the following section will deal with this process.

3.5 Abstraction

Abstraction is often considered to be an important subject in computer science. A high capacity for abstract thought is arguably one of the most desirable qualities for a computer scientist (Hazzan & Kramer, 2007, p. 6). If models are considered to be abstract, then there must be a corresponding descriptor that is the opposite. Following Trogemann and Viehoff (2005, p. 148) I refer to this here as “the concrete”. Thus, an abstraction is made from a concrete thing. Accordingly, the objects and operations in the system requirements document are the concrete elements, and models of these are the abstract elements.

It is possible to describe the process of abstraction as a "distillation" of the essentials. According to Winograd and Flores (1986, p. 97) this "distillation" comes at a cost. The cost is that an abstraction always includes a degree of blindness. This view is based on the concept of Martin Heidegger, who introduced the attributes "ready-to-hand" (zuhanden) and "present-at-hand" (vorhanden) (Heidegger, 1967, pp. 42 and 69). When humans come into the world at birth, there are many objects around them that they can see. As infants and small children they do not have names for these objects. They do not know their properties, possibilities, limitations, or the meaning the objects have. At this stage these objects are "ready-to-hand" for them.

When they have learned to name the objects and their properties, when they know their possibilities and limitations—in other words, when they can think of these objects in an abstract and modelled way—then these objects have become "present-at-hand" for them. Abstraction includes this fundamental transition. Through abstraction, objects become usable in thought processes, but then often in these thought processes only the model of the object is considered, rather than the object itself. So questions relating to its properties, possibilities or use are no longer answered in relation to the object, but to its model.

While putting together a jigsaw puzzle, small children often try out places where the puzzle piece, seen from the perspective of an adult, obviously cannot fit. It would not make sense to an adult to try to fit a piece into a clearly incorrect place in the puzzle, because the abstraction of the puzzle piece and the places it may be put allows adults to pre-think the result of this experiment. As the child cannot yet think about the puzzle piece in the abstract way, as a model with specific properties, it cannot pre-think the result of the experiment. On the other hand, children often do strange things with objects, making adults say things like: “That’s funny, I would never have thought of doing that with that object.” This may mean that the adult’s model did not allow for the use that the child made of the object.

This example illustrates that one is used to thinking about a given object and the things one might want to do with it primarily in the way that one’s abstraction allows. The abstraction is, as it were, blind to things that one could, in principle, do with the object, but cannot do because the abstraction does not include the properties necessary for it. So the abstraction refines one’s concept of objects and can thus offer the possibility of acting with objects in an expanded way. At the same time, however, it closes one’s concept for possibilities of the object one’s abstraction does not allow. One might conclude that abstraction is therefore not particularly helpful. Far from it! This is, of course, not what I am trying to say here. Abstraction is absolutely necessary, and humans use it all the time with significant results. What I would like to point out is that abstraction has a double-edged effect. Along with offering new possibilities for acting, it also limits possibilities for acting.

The description in the requirements document can be seen as a first level of abstraction because words are used to describe phenomena, objects, relations, operations and other elements. On the other hand, this is a description which, in relation to the code of the desired program, does not make clear exactly what is meant, and thus cannot be the abstraction that is required for formalisation. In addition, the description in the requirements document is often fragmentary, in the sense of a proper working application. Such a description may look precise and complete to a non-expert, because it fulfils the degree of precision necessary in relation to the personal framework of the non-expert. Nonetheless, for the system developer this description may be very fragmentary, because the precision of the description must be seen in relation to the algorithm that is to be implemented, and not in relation to any other framework.

3.6 Models

As abstraction is used to create models, and models play an important role in specification, I will now take a closer look at the idea of models. According to Stachowiak (1973, p. 131), the term “model” can be understood to include three features:

1. the feature of mapping,
2. the feature of reduction,
3. the pragmatic feature.

Trogemann and Viehoff (2005, p. 33) explain Stachowiak's statement as follows. Fundamentally, models are always models of something. They can be mapped to originals. This something can again be a model, which in turn also refers to further models. "Every entity that is perceptible or 'producible' by a natural or mechanical cognitive subject can be understood in this comprehensive sense as an original of one or more models"[5] (Stachowiak, 1973, p. 131) [Translation by Cornelius Poepel].

The feature of reduction is a result of the fact that models do not usually include all the attributes of the original. The model is reduced to those attributes that seem relevant to the people developing it.

The pragmatic feature describes the fact that models are not in a one-to-one relation to their originals. In other words, the proper relation between the model and the original exists solely for a specific group of people using the model, and only for a specific time and for specific operations. Stachowiak concludes: "Besides the question of *what* a model is representing, a pragmatically complete definition of the term 'model' must also ask with respect to each specific function *for whom*, *when* and *for what purpose* it is a model"[6] (Stachowiak, 1973, p. 133) [Translation by Cornelius Poepel].

[5] "Jede von einem natürlichen oder maschinellen kognitiven Subjekt erfahrbare, allgemeiner: erstellbare Entität kann in diesem umfassenden Sinn als Original eines oder mehrerer Modelle aufgefasst werden."

[6] "Eine pragmatisch vollständige Bestimmung des Modellbegriffs hat nicht nur die Frage zu berücksichtigen, *wovon* etwas Modell ist, sondern auch, *für wen, wann* und *wozu* bezüglich seiner je spezifischen Funktionen es Modell ist."

The example of the formalisation of a cloud for use in a computer game showed that the physical abstraction of a cloud—"any visible mass of water droplets, ice crystals, or a mixture of both that is suspended in the air, usually at a considerable height" (see section 3.3, p. 31)—would not make sense because the pragmatic feature of this model does not suit the functions of the context in which the cloud is used here. Assuming that the developers wish to formalise brilliant-looking clouds, the question arises of the conditions under which clouds on a computer screen are perceived as brilliant. According to the pragmatic feature it is clear that a model of a brilliant cloud is dependent on the model and its feature of reduction (that is to say, how complete the model is in relation to perception) as well as the pragmatic feature (some people may perceive a specific cloud as brilliant while others do not). This means that, in addition to validity according to its contextual use, the model is dependent on the user and her background.

Coming back to formalisation and the question of whether something can be formalised or has been correctly formalised, it must be concluded that it is not enough to look solely at the requirements document. Besides the requirements one has to look at the user and the context in which the application will be used in order to achieve a proper formalisation. It is important to bear in mind that models based on abstractions will always be blind in parts.

3.7 Formal Specification

In order to progress from the specification to the formal specification, the constraints, relations and attributes of the models must be transformed completely and unambiguously into a formal language that can be translated one-to-one (losslessly) into the code and interface. If models of physical abstractions are used and the models refer to processes inside the computer, the translation of the model into a specification of mathematical functions, described in a formal language, is a task that can often be satisfactorily achieved. With issues addressing the human-computer interface, however, the situation is a different one.

In the context of a musical instrument, the input of the player is considered as important. If an instrument is designed to be open to the personal expression and actions of the player (see example of requirements in section 3.2), the developers will model the player and build a device that is capable of tracking the parameters considered to be relevant to the player's expression and actions in this model. It is important to mention here that the question of which parameters are relevant depends on the musical context and the people who will be using the instrument. This relates to the pragmatic feature of the model described in the previous section.

Once the model has been defined, a method of measurement must be developed to measure the player's behaviour and to provide streams of discrete data representing each output parameter of the model of the player. A problem occurs if the measurement either fails to measure exactly what it is intended to measure (validity) or provides varying results when offered the same input (reliability). Besides this, the process of measurement itself can have an influence on the object being measured. This problem is familiar from measurement in physics research, artificial life and robotics. Regarding the question of how the measurement process can be formalised, computer scientist Prem (1997, p. 6) expresses the following conviction:

> It was already John von Neumann who pointed out that results of measurements, choices of observables, the construction of measurement devices and the measurement process itself cannot in principle be formalized (Neumann, 1955), (Neumann, 1966). The reason lies in the fact that the process of measurement is not of a purely formal nature.

In the measurement process two dynamical systems interact. Prem thus states that the measurement process involves an "inherently dynamic nature" (p. 6) which cannot be formalised because the process of measurement will always influence the object being measured. With regard to feedback systems that include the tracking of a player, one has to say that the process of measurement may indeed disturb the player—for example in cases where limited reliability is found. If the tracking method does not work completely reliably and the player wishes to generate sounds that depend on proper tracking data, the player is forced to change her playing style in order to generate the necessary tracking data. In this case the process of measurement influences the player being measured.

3.8 Verification

Once the code and the interface have been implemented, a verification will help to establish whether the goals, intentions and requirements of the application have been met. I have expanded figure 3.1 in figure 3.2, where arrow B illustrates this verification.

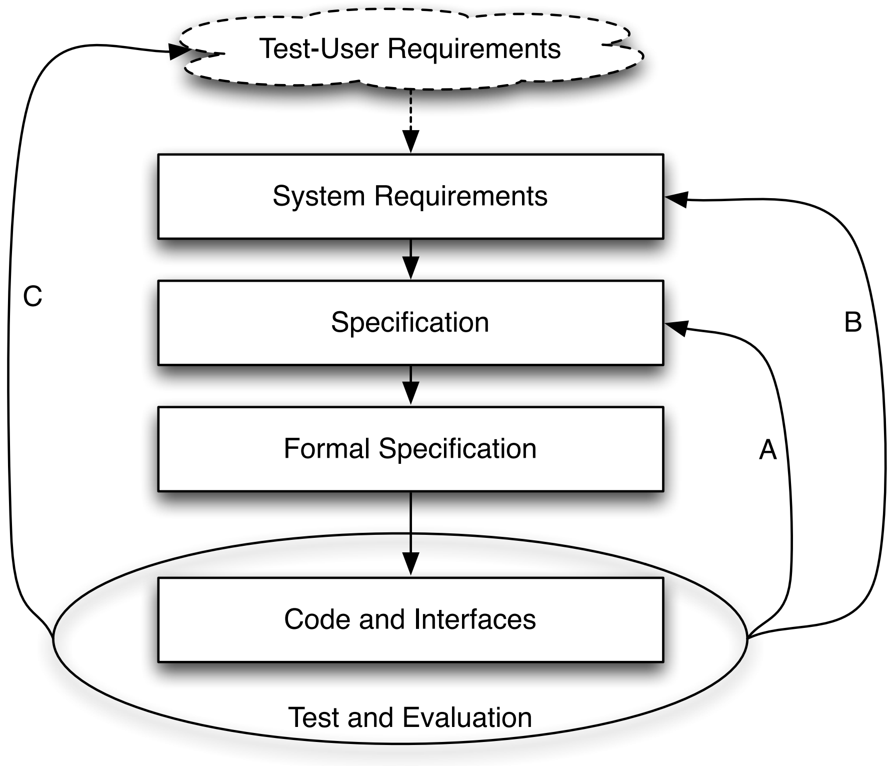

Figure 3.2: Verification relating to different levels.

Looking at commercial products, it should be mentioned that the contract between a customer and a company is not usually based on the requirements document. Instead it is based on the specification (arrow A in figure 3.2).[7] In this case the verification is not conducted in order to confirm whether the requirements have been met, but whether the specification has been implemented correctly. Since descriptions in the requirements are written (at least in parts) in natural language, they can be interpreted in different ways. Thus, the requirements document is not considered to be an appropriate basis for a contract. A specification, however, has to be unambiguous (IEEE, 1998, p. 4).

A company dealing with substantial sums of money has to ensure that it is legally safe with regard to the promises that have been made to its customers. For the customer, on the other hand, it is important to know the limitations of the models which were used for the specification, and to understand in which context and under which circumstances these models will continue to work, or indeed no longer do so. In cases where these limitations are not understood a customer risks being disappointed by the function of the product in relation to the requirements.

In this research the creation of an instrument for a group of musicians is being considered. When an evaluation is conducted it is important to mention that the users' critique and suggestions for improvements to the system are mostly given in natural language describing perceived phenomena.

[7] For a contract between an organisation and a customer the IEEE Standard 830-1998 recommends using the specification as “a baseline against which compliance can be measured” (IEEE, 1998, p. iii). The software company MicroTools Inc., for example, writes the following: “We use [the specification] as the basis of our contract with our clients all the time” (Japenga, 2010, para. 4).

They are related to the users' ideas of how the system should be. Therefore I consider the evaluation feedback in relation to specific requirements of the test-users (arrow C in figure 3.2). The feedback from the users can be used to improve the system requirements.

Chapter 4

Related Research

This thesis focuses on sound synthesis systems which are used by string players and which seek to take advantage of the skills of string players. Accordingly, this chapter focuses on string-specific playable synthesisers. First I give a brief overview of the research field and pointers as to where in the thesis I take a more detailed look at specific research questions. On the basis of an early analogue synthesiser, I then present two main directions in the design of musical interfaces. In addition, three sections describing implementations of related research are presented. Analysing these implementations, conclusions are drawn in the final section 4.7.

4.1 Research Field

Many newly developed instruments have been presented at conferences such as NIME. Chadabe (1997) presents an overview of the history of electronic instruments including digital synthesisers and musical interfaces. A collection of many new musical interfaces such as the *sensor chair* and the *rhythm tree percussion controller* are described in an article by Paradiso (1997). A current overview of controllers for digital musical instruments can be found in Miranda and Wanderley (2006).

Traditional instruments are known for their great potential for musical expression (Settel & Lippe, 2003, p. 197). In order to create the same expressive potential using computer-based musical instruments, a considerable amount of research has been done to understand how this expressive potential is achieved. Thus researchers have investigated the actions of the performer and their relation to sound (O'Modhrain & Chafe, 2000; Wanderley, 1999). Some of the questions asked are: what happens when an instrument is played? What is the performer doing (Paine, 2007, p. 70)? How are the actions of the performer connected to what we hear, and how can we measure these actions (Cadoz & Wanderley, 2000, pp. 44-46)? Which meaningful parameters can be identified in generating expressive performance (Canazza, Roda, & Orio, 1999, p. 381)?[1]

Gestures play an important role in the performance of music. In this research, a gesture is understood to be a particular body movement which is clearly visually identifiable and is made in order to realise a specific action. It is a common conviction that the physical gestures of the musician are a key element responsible for the expression of a performer. As a result of this conviction one can see the European research project *cost 287 ConGAS* en-

[1]Section 5.2.1 will address these issues more thoroughly.

titled “Gestural Controlled Audio Systems” (ConGAS, 2007), the publication “Trends in Gestural Control of Music” (Wanderley & Battier, 2000), and the Norwegian research project “The Musical Gestures Project” (Aksnes, Godøy, Kvifte, & Ruud, 2007).[2]

In cases where it is desired to control a digital synthesis engine with parameter data derived from captured gestures of the performer, the question of how these gesture parameters can be mapped to the input of the synthesis engine will be of importance. Rovan, Wanderley, Dubnov, and Depalle (1997), Hunt and Kirk (2000), and Arfib, Couturier, Kessous, and Verfaille (2002) give examples of different strategies regarding how to map gesture data to synthesis inputs. Part of the aim of this research is to understand how the input to a traditional instrument is mapped to the sound result. A further question is that of which mapping strategy best enables a computer-based instrument to be open and transparent to the intentions of the musician.[3]

4.1.1 Two Directions in the Design of Musical Interfaces

Looking at development strategies for synthesisers, I would like to point out two strategies often used. Both start from different positions but have the same goal: to build a better synthesiser. These strategies can be described as follows:

[2]A more detailed look at gestures will follow in section 5.2.2.

[3]This topic will be considered more thoroughly in section 5.2.4.

1. There is a synthesis engine and we want somehow to control it.

2. There is a musician and we want to build an interface for her.

4.1.2 Controlling a Synthesis Engine

Early synthesisers available on the market were analogue synthesisers. An example is the EMS AKS Synthi A (see figure 4.1).[4]

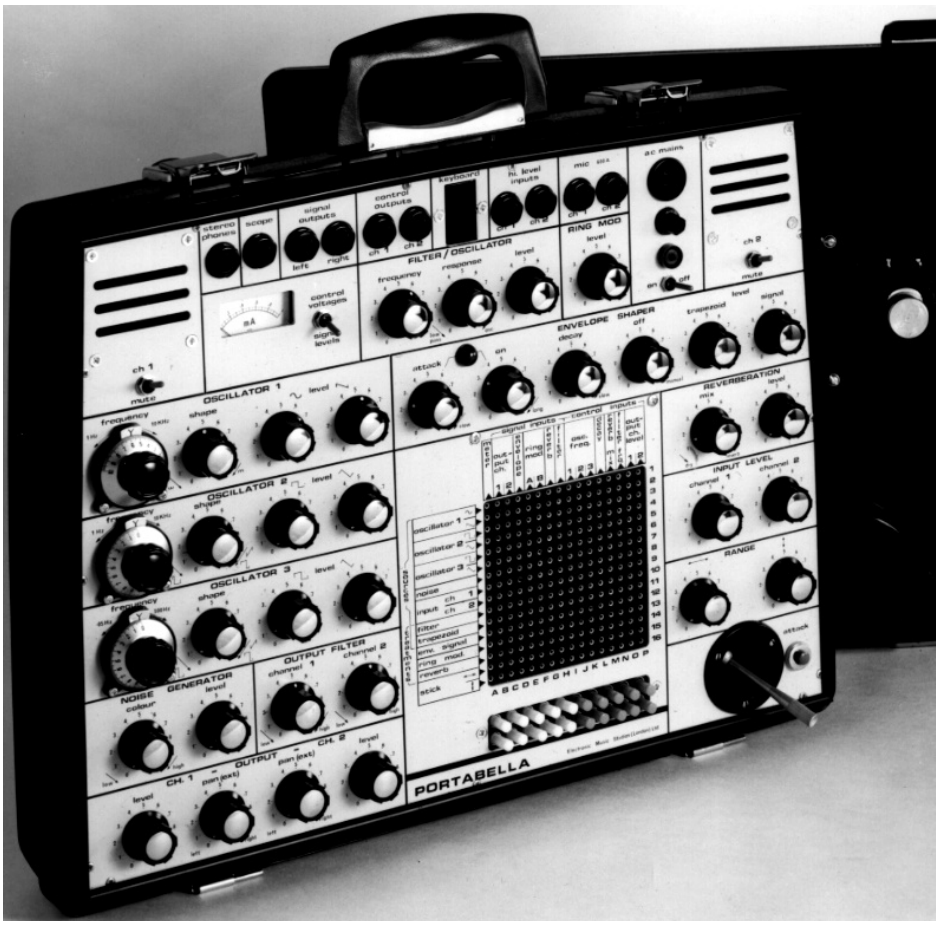

Figure 4.1: EMS AKS Synthi A.

[4]Image from Electronic Music Studios (Hinton, 2010).

This synthesiser was equipped with keys, adjusting knobs, a joystick and a patchbay. These electronic parts suited the mode of synthesis very well. They were common parts, easily available and easy to integrate into an electronic circuit. The operation possibilities available were conceived and planned from the perspective of the needs of the synthesiser: the inputs were given and had to be equipped with parts enabling a human to manipulate the voltages at these inputs. When playing the instrument, the performer had to adapt to this control surface.

A more complex approach to working with synthesis engines is found in the *reacTable* (Jordà, Kaltenbrunner, Geiger, & Bencina, 2005). The control of the synthesis engine and the synthesis engine itself can be defined by a variety of knob-like objects placed on a table. The table and the objects together offer a "tabletop tangible user interface" (p. 579). Objects of modular synthesis such as oscillators, filters, mixers, or control filters can be placed and connected on the surface while performing. By moving the objects on the table connections are made and severed. Turning the objects generates control parameters that are used to drive the input parameters of the modular synthesis objects. This concept aims to maximise the user's ability to create and control the structure of modular synthesis.

4.1.3 Building an Interface for a Trained Performer

There were, of course, musicians who did not want to adapt completely to instruments such as the EMS AKS Synthi A and felt restricted in their musical expression by the options offered by a synthesiser. This raised the question of whether it might be possible to adapt the instrument to the player, and, more precisely, to the skills a musician already has. The keyboard was a first answer to this question. This approach was conceived and planned from the perspective of the user. Since many musicians were trained piano players and it was possible to develop keyboards able to generate the necessary electronic signals, the solution of a keyboard interface worked very well.

If a synthesiser is defined as a machine, the knobs and the keyboard can be seen as human-machine interfaces. In research the question is still open as to how one can build ideal interfaces for synthesisers or interactive music systems. There are researchers focusing on the challenge of building interfaces that enable not only keyboard players but any kind of trained instrumentalist to make use of their available skills. The requirements are conceived primarily from the point of view of the performer and her possibilities. Many examples of resulting interfaces can be found in Miranda and Wanderley (2006). As my approach is based on the requirement to make use of the skills of trained string players, I have selected related research which uses (at least in some form) the skills of traditional string players.

4.2 Computer-Based Stringed Instruments

Building interfaces for strings is different from building interfaces for keyboards or guitars, because the playing parameters of strings are, similarly to those of woodwinds, modified in a continuous way by the players. With a piano, most of the playing parameters of a tone are established once a pianist has pressed down a key and the tone has started. At this point it can no longer be changed (except by using foot pedals). A cellist, however, continues to perform the tone for as long as it is audible. This may be a reason why the MIDI keyboard interface was far more successful than the MIDI clarinet or violin.

In order to understand the interaction between the player and the instrument, research has been undertaken to investigate the process of violin playing from a physical point of view (Askenfelt, 1986; McIntyre, Schumacher, & Woodhouse, 1983, pp. 1331-1335; Schelleng, 1973).

The mechanical interaction between the bow and the string is described by Pitteroff and Woodhouse (1998). Schoonderwaldt, Rasamimanana, and Bevilacqua (2006) use accelerometers and video tracking to investigate bow velocities in violin playing. Rasamimanana, Fléty, and Bevilacqua (2006) use an augmented violin[5] to analyse performers while using common bowing techniques.

Dan Overholt and I have presented an overview of publications in the period between 1995 and 2005 concerning research on computer-based bowed stringed instruments (Poepel & Overholt, 2006, pp. 390-391). Both the proceedings of the ICMC (International Computer Music Conference) (ICMC, 2009) and NIME were analysed. All string instrument related papers were collected and categorised into five string instrument related groups: controller, sound synthesis, sound processing, complete instruments and others.

[5]The bow of this violin is described in section 4.3.7.

Results can be found in figure 4.2. ICMC papers are presented in blue and NIME papers in red. Due to the fact that the NIME conference was first held in 2001, more papers are available from the ICMC.

Despite the fact that there are already many synthesis methods available that could be used to build string synthesisers, research has been focused on new or expanded methods of sound synthesis. Analysing the topics in this field, it was found that it was primarily the physical modelling of bowed strings which was addressed.

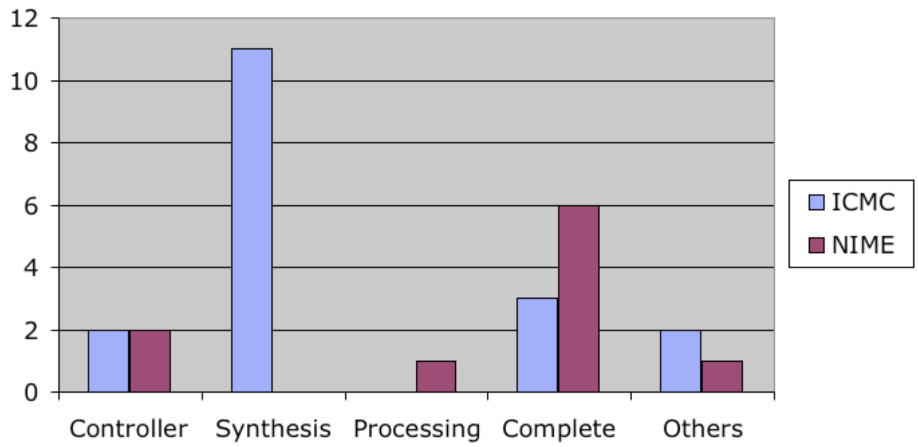

Figure 4.2: Number of string-related research papers at ICMC and NIME between 1995 and 2005.

The publications of Young and Serafin (2003, p. 107) and Serafin (2004, pp. 142-146) present research results on the playability of physical modelling synthesis algorithms. Like Woodhouse (1993, pp. 125-126), they see an important quality criterion of stringed instruments in their playability. This conviction is supported by my experiences.[6]

[6]String players testing my prototypes (see section 7.2) often mentioned the instruments' playability as an important factor.

In the following sections, the projects I found to be related to this research will be described. As a basis for this, literature and personal information I have collected about these instruments will be used. Where the information collected allows, the focus will be on the goals and quality criteria of the developments, the ways in which these goals were reached, and the evaluations conducted with the instruments.

4.3 Using Sensors

These sections present related developments that mainly use sensor technology providing explicit parameter data to control the sounds of computer-based instruments.

4.3.1 Sensor-Bow

In the late 1980s Jon Rose collaborated with researchers at STEIM.[7] In this project their aim was "to bring together the physicality and dynamics of improvised music (as played on a violin) with the quick change and virtual possibilities of computer music" (Rose, 2005, para. 3).

The *Sensor-Bow* which they developed measures bow pressure with a pressure sensor under the index finger and bow position using a sonar sensor (Trueman, 1999, chapter 3, p. 8). In an extended version it also measures acceleration using an accelerometer fixed on the bow arm. Sensor data is

[7] Centre for research and development of instruments and tools for performers in the electronic performance arts, Amsterdam, Netherlands (Steim, 2010).

translated into MIDI data, which is used to control several MIDI devices, including a sampler.

Although he does not say this explicitly, one might hypothesise that Rose did not attempt to keep the connection between movement and sound as familiar from traditional violin playing. He did not intend to sound or play like a traditional violin player, but preferred to create new musical experiences using the sensor-bow and its data output. Therefore the question of playability as compared to a traditional instrument was not important to him. I assume that Rose did not seek explicitly to gather all data describing the rough and subtle actions made by the performer in order to provide musical expression.

The Sensor-Bow was used in several concerts on international stages such as the Ars Electronica Festival in Linz in 1986 and at international festivals of electronic music in Stockholm, Bourges, or Berlin (Rose, 2005, para. 10).

4.3.2 SuperPolm

In 1995 Suguru Goto began developing the virtual violin *SuperPolm* in collaboration with the IRCAM[8] engineers Patrice Pierrot and Alain Terrier. The goal of the SuperPolm project was to build a virtual violin in order to perform the composition *VirtualAERI* for virtual violin and live video (Goto, 1999, p. 115). The virtual violin was designed to look good, not like a simple sensor, and be easily recognisable as an instrument. It needed to

[8]Institute for music/acoustic research and coordination, Paris, France (IRCAM, 2010).

work properly and be easy to handle. The virtual violin was not intended to replace the violin, but to use the model of the gestures of a violinist. The first performance of *VirtualAERI* was given in 1996 at Espace de Projection, the main concert hall at IRCAM.

Inspired by the composers Dieter Schnebel and Earle Brown, *VirtualAERI* is a composition playing with gestures on stage. Gestures are categorised into small gestures (finger bubbling), larger gestures (for example arm movement of the violin player) and large gestures (walking on stage). The composition plays with gestures in the sense that the sound sometimes corresponds to the performer's gesture on the virtual violin (as one would expect), but at other times is completely different from one's expectations. The score does not provide notes to be played. Instead, gestural instructions are written to be executed with the SuperPolm by the performer (Goto, personal communication, December 13, 2007).

In order to build the desired instrument, a wooden body was created. Instead of four strings there are four sensors on the fingerboard, each of them measuring pressure and position. The traditional bow (with hair) is replaced by a rod of a similar length to that of a bow. The rod is connected to a mechanism allowing a longitudinal movement of the rod and measuring its absolute position. Bow speed can be calculated from the difference in absolute position over time. In addition, the instrument is equipped with a one-dimensional accelerometer to measure how much the instrument is inclined. A pressure sensor under the chin rest allows the generation of data, and a field of keys is mounted on the body to select presets. The output of the sensors is transformed into MIDI data and sent to a MaxMSPJitter patch which performs sound synthesis and video processing.

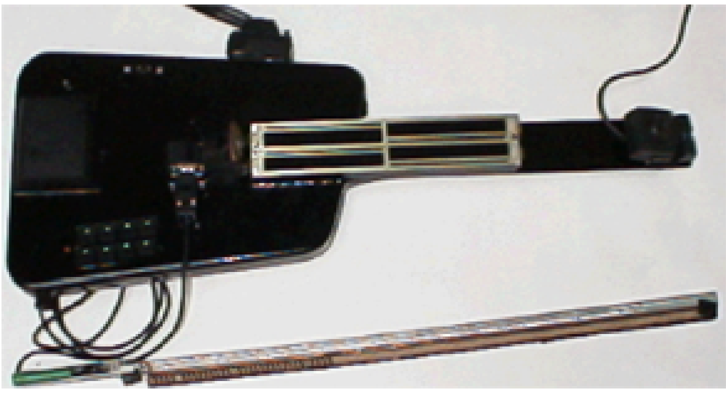

Figure 4.3: Bow and body of the SuperPolm.

The composition *VirtualAERI* is intended to both satisfy and confuse people, by using sounds that are related to the visually perceivable gestures of the performer and sounds that are unrelated to the performer's gestures. In this composition the instrument is used to play with the public's expectations. In the experience of the composer the public does not like to be confused. He explains (S. Goto, email communication, March 4, 2008):[9]

> People want to understand everything. Final thing is that something quite clear happens on stage and people can listen more on musical aspects, without concerning much about technical points. As soon as public sees the complex things and unexpected things they start to confuse and sometimes people even start to doubt about technical points wondering if the system is really working or if I am cheating on stage. The composition is almost like a game of this psychological aspect. To explain more, the poles of this perceptional aspect is: A: expected and B: unexpected as well as: A: clear and B: complex. I play with this two poles as a kind of psychological game. This is the concept behind SuperPolm.

[9] The quotation retains the exact wording of the email.

As a consequence of this audio-visual interplay shown in the context of a stage performance, the composition would not work if played from an audio-CD.[10]

In the composer's estimation the initial system requirements of the instrument have been met. In order to be played, the SuperPolm has to be learnt as a new instrument. It requires the skills of extended violin gestures, MaxMSP programming skills and the ability to read the score with gestural notation. By December 2007 the instrument had been used in approximately 70 concerts on international stages across the world. When used in a performance the instrument was always played by the composer (Goto, personal communication, December 13, 2007).

4.3.3 Funny Fiddle

Neal Farwell's *funny fiddle* is a violin-like controller that was built in 1996 and described in his doctoral thesis (Farwell, 2001). The controller was modelled on the violin because of the "potential gestural subtlety a classical training yields" and because of the violin's "rich cultural resonance of larger-than-life virtuosity" (p. 104). The instrument is not played via a string but similarly to the SuperPolm via sensors. The read head of a tape recorder serves as a string. Describing the abilities of the instrument, Farwell points out that it not only controls when and what sounds are played, but also "where sounds occur: it foregrounds the [sound] diffusion act" (p. 104).

[10]This issue raises the general question of the validity of auditive documentation of media-based composition, because this is a work of both sight and sound. While the question of documentation is worth considering, it will not be treated in any more detail here because the focus of this chapter is on the design and function of the instrument.

To generate control information the performer can press down one string using his fingers. The control information generated is dependent on where the string is pressed (p. 132). The string is used to offer the player a "passive haptic feedback; that is to say, the string *feels* as it should to the player" (p. 129). This is considered to be of importance "for the 'injection' of regular violin technique" (p. 129). The bow of the funny fiddle is tracked with respect to its position and velocity using a "40KHz pulsed ultrasonic transmitter-receiver pair" (p. 130) and a "bow-on-string" (p. 133) signal is provided.

Gipsy fugue, a composition by Neal Farwell for the funny fiddle, loosely follows the Gipsy tradition of Eastern European music and includes samples of J.S. Bach's work (p. 124). The composer was seeking to "find satisfying relations between violin-like gestural activity and the sounds as experienced in performance" (p. 124). *Gipsy fugue* was performed by the composer in two concerts at the University of East Anglia. Furthermore the funny fiddle also featured in an Open University network television program (Farwell, 2001, p. 104).

4.3.4 BoSSA

The *BoSSA* (Bowed-Sensor-Speaker-Array) was built by Dan Trueman and Perry Cook. The instrument was built to use both physical and visual aspects of violin playing (Trueman, 1999, chapter 3, p. 13). Control data to drive a synthesis engine is generated using a set of sensors. The BoSSA consists of a speaker array equipped with 12 audio channels. On top of the speaker array a *Bonge* (chapter 3, p. 14) is bowed with an *R-Bow* (chapter 3, p. 10). In addition, a so-called *Fangerbored* (chapter 3, p. 13) is mounted with a hinge onto the speaker array and the R-Bow. The instrument is under permanent reconstruction. By 2005 it had been rebuilt twice. According to Trueman (email communication, November 30, 2005), the sensors and information transmitted are as follows:

R-Bow:

- sonar sensor (transmitter on frog, receiver on speaker) to provide information about bow position;
- accelerometer (on frog) to provide information about tilt and acceleration of bow;
- force-sensing-resistor (FSR)[11] (under the index finger) to provide information about bow pressure;
- piezo pickup (at the tip of the bow) to send an audio signal to the computer;

Bonge:

- four FSRs (arranged like "strings" on speaker) to sense bow pressure;
- sonar receiver (on speaker) for bow to measure absolute longitudinal bow position;

Fangerbored:

- one linear position sensor to provide information about position of finger on the Fangerbored;

[11] An FSR is an electronic component part whose resistance changes when a force is applied.

- FSR (on back of the Fangerbored) to provide information about finger pressure on the Fangerbored;
- accelerometer to sense tilt of the Fangerbored.

The sensor output is transformed into MIDI data and sent to a computer holding the synthesis engines. The most time-sensitive streams of data (R-Bow sonar, for example) are updated as quickly as possible (for example every 8ms for the sonar). This, combined with the 12-bit resolution (the original BoSSA sensors from 1999 were all 7-bit), has made the instrument far more responsive and playable (D. Trueman, email communication, November 30, 2005).

Trueman writes about his playing technique on the BoSSA (Trueman, 1999, chapter 3, p. 13):

> I have learned a variety of gestures that are meaningless without the sensor data. For one, I can simply press down with varying degrees of weight and play a virtual instrument without pulling the bow across a string—the grit of the bowed-string relationship is non-existent. Secondly, I can move the bow, or shake the bow, and depending on the kind of gesture, play a virtual shaker or adjust a signal processing parameter. Again, the bowed-string [*sic*] is irrelevant. In fact, the violin itself is irrelevant. By combining these new gestures with the familiar ones available, I have an entirely new way of playing the violin—it is an instrument with a unique, and only partially explored, expressive potential.

I conclude that from a macroscopic point of view, the gestures made when performing with the BoSSA still look like those involved in playing a violin or, more closely, a cello. However, on the microscopic level, a new playing technique must be learnt. Since Trueman points out that “the violin itself is irrelevant”, I conclude that the connection between gesture, feel and sound has left the paradigm of a traditional violin.

The BoSSA has been used in concerts on international stages many times in recent years, including its very first performance at Princeton University in 1999, and performances at the “Music at the Anthology Festival” in New York City in 2005 and the ICMC in Barcelona in 2005 (D. Trueman, email communication, September 17, 2009). The instrument was played by Dan Trueman.

4.3.5 Hyperbow

The *Hyperbow* was presented by Young (2001) as a successor to the bow of the *Hypercello* (see section 4.4.2). It was among the goals of the Hyperbow project “to create a violin bow capable of measuring the most intricate aspects of violin technique, the subtle elements of physical gesture that immediately and directly impact the sound of the instrument while playing” (Young, 2001, p. ii). In addition, it was desired to “enable a player to use the same postures of the right hand wrist and fingers on the bow” as the player was accustomed to, and thus to build a bow “as similar to a traditional bow in size, weight, and weight distribution as possible” (Young, 2002, p. 4). Similarly to the bow of the Hypercello, the Hyperbow is built to measure bow position. In addition, it measures acceleration and strains (downward and lateral) inside the bow stick.

To achieve these goals the technology of electric field sensing was used (Paradiso & Gershenfeld, 1997). A carbon fibre violin bow was equipped with a resistive strip working as an antenna and attached to two oscillators (50 and 100 kHz) on each end of the strip. The signal from the resistive strip is received by an antenna mounted onto the bridge of the violin. Depending on the amplitude of each of the received oscillations, longitudinal bow position can be calculated. Accelerometers mounted on a circuit board, which is mounted onto the frog, provide information about the 3D acceleration of the bow. Strain sensors placed around the middle of the stick are used to determine the downward and lateral strains of the bow (Young, 2002, p. 5). The output of the strain and acceleration sensors is transmitted via an RF system implemented into the small circuit board mounted onto the frog. Thus the Hyperbow works wirelessly. The finished Hyperbow has a weight of 89.751g whereas the raw carbon fibre bow weighs 60.930g. One may expect this difference in weight to make a noticeable difference to the player in terms of the feel of the bow.

In terms of an evaluation, Young (2002, p. 5) mentions that "the playability of the *Hyperbow* was evaluated by several accomplished players and two professional violinists and found to be adequate by all." The Hyperbow was also used for the evaluation of the physical modelling synthesis algorithms that were mentioned in section 4.2. While the bow was built for a traditional western virtuoso violinist, it was also intended as an alternative "for any player of a bowed string instrument" (p. 5). The bow has been used in several concerts, including the performance of the composition "Toy Symphony" by Tod Machover in 2002.

It has been used by highly regarded string players such as Joshua Bell and other international composers and performers (Young, Nunn, & Vassiliev, 2006, p. 396). A refined version of the bow was used as a measuring tool for an investigation into bowed string performance (Young, 2007, p. 2).

4.3.6 vBow

Charles Nichols began work on the *vBow* in the context of his doctorate at Stanford University (Nichols, 2003). The vBow is a bow built for playing virtual instruments, and incorporating the haptic feedback of a traditional violin bow. It seeks to be an "expressive human computer interface, based on the paradigm of the violin" and "a controller that solely [maps] the violinist's gesture to synthesis parameters" (Nichols, 2002, p. 215). Furthermore, it is built to allow the computer musician to use most of the gestural freedom of a bow on a violin string. "The vBow [...] provides the performer [*sic*] with an amplification of their musical intent, through the translation of their physical gesture into the expressive manipulation of timbre" (Nichols, 2003, p. 12).

Addressing the question of what a player does when playing the violin and, in particular, when bowing, Nichols provides insight into the literature of violin pedagogy and into physical research on violins and player-instrument interaction. Nichols uses such descriptions as a basis for the model of a player the interface will be built for. "To begin with, it should be understood what intricate and precise mechanics are involved when bowing, a gesture that requires a most flexible instrument and technique" (Nichols, 2003, p. 36). He goes on by citing the violin teacher Hodgson (1958): "The arm as a whole forms a wonderful bowing machine, and without the supple adjustments of the smaller members the work could not be done" (Nichols, 2003, p. 37). Following on from this, and in order to achieve the goals of the vBow project, a device had to be built which would measure the relevant parameters of the arm (bowing machine), and be capable of tracking the fine movements of the arm and fingers. Cadoz, Luciani, and Florens (1984, p. 61) and Chafe (1993, p. 79) assume that haptic feedback can play an important role in instrumental performance. On the basis of findings by Chafe and O'Modhrain (1996, pp. 68-69) and O'Modhrain (2000, pp. 76-78), Nichols integrated the concept of haptics into the architecture of the vBow.

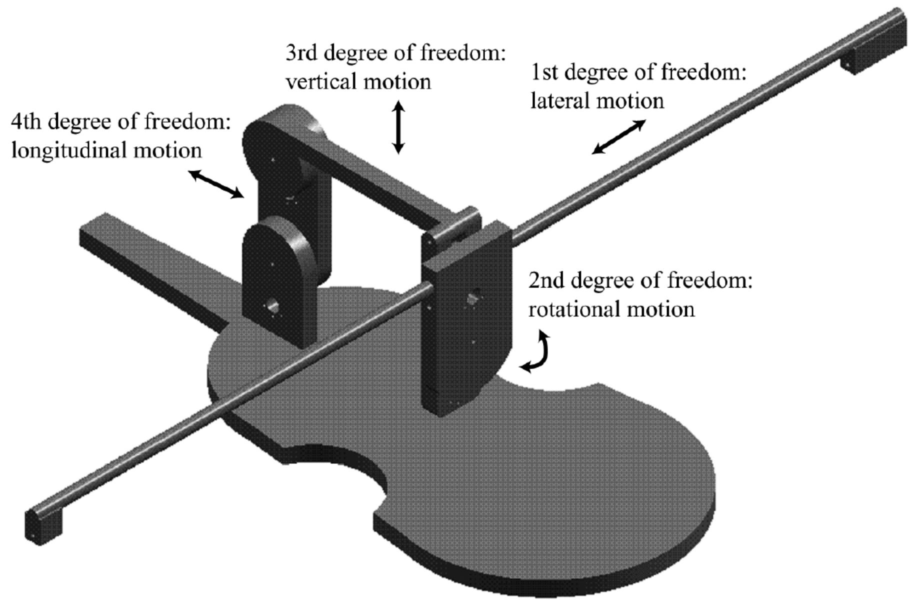

Figure 4.4: vBow architecture.

In order to realise the plans, a ground plate was built holding a set of arms, pivots, sensors and actuators, in turn holding a movable bow stick. Figure 4.4 illustrates the basic configuration (Nichols, 2003, p. 76). Four servo motor systems track four degrees of freedom of the bow stick: lateral, rotational, longitudinal and vertical motion. In order not only to track the player's bow movements, but also to give the player a more realistic feeling of bowing, the system is able to simulate haptic feedback. This is achieved by using servo motors and thus generating physical forces against the bow stick. The tactile feedback is able to simulate the vibration, friction and elasticity of the strings, as well as the feedback perceived when moving the bow across different strings.

In order to implement this feature, Nichols developed control software that "produces the haptic feedback of friction and vibration for the lateral motion, detents for the rotational motion, elasticity for the vertical motion, and friction for the longitudinal motion" (Nichols, 2002, p. 218). Encoders are used to measure the physical forces applied to the bow, and servo motors create the physical forces felt by the user. The physical forces applied to the encoders are mapped to the input of a bowed-string physical model. Streams of data for sound as well as for the tactile feedback are calculated and sent to the speakers and servo motors.

An explicitly stated evaluation is not available. However, a report on the proper working functions of the system is given, for example, in a description of the vibrational feedback applied to the bow stick and felt at the frog of the bow (Nichols, 2003, p. 116). The controller was used in live demonstrations at CCRMA,[12] IRCAM and the Banff centre.[13] It was played by the developer Charles Nichols.

4.3.7 IRCAM Bow

In early 2004 a team of researchers at IRCAM including Emmanuel Fléty, Nicolas H. Rasamimanana and Frédéric Bevilaqua began building a bow tracking system for the IRCAM augmented violin project (Bevilacqua, Rasamimanana, Fléty, Lemouton, & Baschet, 2006). The researchers wanted to develop their own bow tracking system which could be used for research into bowstroke recognition. The construction of the *IRCAM bow* was intended to measure the bowing of players, both for scientific reasons and as measurement data for composition.

The bow implements similar technologies to those used for the Hyperbow (see section 4.3.5). It incorporates acceleration sensors and a bow position sensing system using the principle of electric field sensing (Paradiso & Gershenfeld, 1997). An RF transmitting system enables the bow to operate wirelessly. In contrast to the Hyperbow, the IRCAM bow was not built with the aim of tracking all the subtleties of performance when bowing the instrument. It is rather intended to provide a set of measurement data which is then investigated for its usability. Figure 4.5 presents an image of the bow and the receiving antenna mounted on the augmented violin.[14]

[12]Center for Computer Research in Music and Acoustics, Stanford, California, USA (CCRMA, 2010).

[13]Department of Music & Sound, The Banff Centre, Banff, Alberta, Canada (Banff Centre, 2009).

[14]Photo source: IRCAM, Emmanuel Fléty.

Figure 4.5: IRCAM bow and augmented violin.

The measurement output makes it possible to distinguish between different bow strokes as well as differences in dynamics (Rasamimanana et al., 2006, p. 145). Thus, bow strokes can be used to enhance a score following system and to influence or control live electronics in performance. Schoonderwaldt et al. (2006) have presented an additional video tool for tracking bow movement. By synchronising the accelerometer and the video stream (point tracking) they were able to achieve a drift-free reconstruction of the bow velocity (p. 200).[15]

The IRCAM bow was used in the composition and performance of *Bogenlied* by the IRCAM-associated composer Florence Baschet. There are plans to use the bow in future compositions.

[15]Another, simpler method for measuring velocity was developed by Neal Farwell. He uses an ultrasonic transmitter mounted on the body of the instrument and a receiver mounted on the frog of the bow. In order to increase the velocity resolution Farwell developed a "*Doppler* velocity measurement" (Farwell, 2001, p. 132) which is inherently drift-free and thus offers a valuable alternative to the system Schoonderwaldt et al. (2006) have presented.

4.4 Using Sensors and the Audio Signal

Several newly designed instruments use combinations of sensors and audio signals coming from vibrating strings. The following sections describe such instruments.

4.4.1 Celletto

The celletto is an electronic cello developed by the researcher, composer and performer Chris Chafe. He uses the celletto as a research tool as well as for performances.

The instrument senses bow force by means of a strain gauge (bend sensor) mounted at the middle of the bow stick. In addition, an accelerometer is placed at the end of the bow (at the tension screw) (Nichols, 2003, p. 56). In order to research the tactile feedback between string and fingering hand, Chafe placed an accelerometer on the nail of the left index finger. He recorded the audio signal of the instrument and compared those recordings with the data from the accelerometer. The results allowed him to show the tactile feedback cues of the fingering hand (Nichols, 2003, p. 31).

Sensor data is translated into MIDI data and used to control parameters in interactive compositions. Chafe uses the bow with both traditional gestures and new gestures such as shaking the bow in the air to generate accelerometer and bending sensor data (Trueman, 1999, chapter 3, p. 8).

The celletto has been used in several concerts on international stages, for example at the Center for New Music and Audio Technologies (CNMAT, 2010) in Berkeley in 2003, and in Seoul in 2007.[16] An image of the celletto can be found on the front page of the Computer Music Journal, Volume 14, Number 4, Winter 1990.

4.4.2 Bowed Stringed Hyperinstruments

The *Hyperinstruments* project was started in 1987 by Tod Machover at the MIT Media Lab (Machover, 1992, p. 1). Its aim was to expand stage performance through the development and use of intelligent and interactive musical instruments as well as computers (p. 4).

One aspect of the development was focused on bowed stringed instruments. The Hypercello was designed "to control an extensive array of sounds through performance nuance" (Hypercello, 2010, para. 1) and uses sensors to track the actions of the performer. "The goal of the sensing was to unobtrusively and responsively detect the player's actions" (Paradiso & Gershenfeld, 1997, p. 73). In addition, the audio signal coming from the instrument is analysed in pitch and loudness in order to control the devices and algorithms generating music.

After tests with a Zeta MIDI cello, an RAAD electric cello was selected as the instrument's body. In order to detect the player's actions and performance nuance, the wrist angle of the bowing hand is measured using a hand skeleton. In addition, measurements of bow pressure (pressure sensor under

[16]Online concert involving two locations, duo with Juan-Pablo Caceres (2007) who was at CCRMA in Stanford.

the right index finger), bow position relative to bridge and strings, and bow speed (calculated from bow position over time) are taken. With regard to actions of the left hand, the position of the finger pressing a string down onto the fingerboard is measured by a resistive thermoplastic strip on the fingerboard (Paradiso & Gershenfeld, 1997, pp. 72-76).

The data collected from the measurement apparatus is used to control samples, synthesised sounds, sound processors and music computer programs. Yo-Yo Ma, the well known cellist, commissioned Tod Machover to compose music making use of the Hypercello. The composition *Begin Again, Again...* (Machover, 1991) for Hypercello solo was a result of this fruitful cooperation. The piece was premiered by Yo-Yo Ma on August 14, 1991, at the Tanglewood Festival.

A formal evaluation with multiple participants in a controlled environment has not been not published, and, to my knowledge, has not been conducted. However the performer's description of its practical use can be seen here as an aid in gaining a more detailed insight. In 1994 *Begin Again, Again...* was performed at the Concertgebouw concert hall in Amsterdam. As described by Levenson (1994), the team around Machover, together with Yo-Yo Ma, realised that the sensors and tracking systems used did not capture all the information the player imparted to the string. While the Hypercello is capable of detecting dynamics and bow movements, it "cannot measure all the motion a skilled player can impart to a string" (Levenson, 1994, p. 16). Levenson concludes that the Hypercello is not yet able to "capture the full range of expression that a musician uses to create his own interpretation of a piece". And Yo-Yo Ma comments, "If you could put expression in, it would be really good" (p. 16).

With regard to the measurement apparatus, Ma reports that the performer has to serve the system with appropriate input in order to avoid sending wrong control data to the sound engines: "I have to make sure that certain signals are magnified. You can't be too subtle. I exaggerate what I do. I have to, to minimize the chance of error" (p. 17). The instrument has been used in several concerts on international stages. A CD is available which includes the composition *Begin Again, Again...* performed by the cellist Matt Heimovitz (Machover, 2003).

A *Hyperviola* similar to the Hypercello was also built. A composition by Tod Machover was commissioned by Betty Freeman (Los Angeles Philharmonic New Music Group). The first performance of this composition entitled *Song of Penance* was played by violist Kim Kashkashian and the Los Angeles Philharmonic New Music Group in February 1992.

In 1993 the *Hyperviolin* was introduced. Parallel to the development of the Hypercello, a RAAD electric violin was used. The completed instrument was based on the technology of the Hypercello. However, some improvements were made to make the instrument more sophisticated. Firstly, the bow was built to be wireless. Secondly, the instrument was modified to add sensitivity to the measurement of phrasing. Finally, timbre analysis of the instrument's acoustic signal was integrated, in order to interpret subtle changes in tone colour.

A composition by Tod Machover was commissioned by the Saint Paul Chamber Orchestra. The first performance of the composition entitled *Forever and ever* was given in September 1993. Violinist Ani Kavafian performed as soloist and Hugh Wolff conducted the Saint Paul Chamber Orchestra. To my knowledge, published works which contain a formal evaluation or statements of Ani Kavafian's experiences with the Hyperviolin do not exist.

4.4.3 Sbass

The *Sbass* is a bodyless, electronic double bass equipped with pickups and sensors. It was built by composer and performer Curtis Bahn and luthier Bill Merchant. With regard to the aim of the instrument, Bahn writes (Bahn, 2010, para. 3):

> The interactive computer environment is designed to maximize flexibility in performance to generate, layer and route musical material with the same improvisational freedom as he [a bass player] has developed with his string bass. An aim in creating this interface is to enable me to take my electro-acoustic music out of the studio and into a wide range of performance contexts.

The system consists of an electric bass using an array of pickups, a contact microphone mounted under the hair of the tip of the bow, a small mouse touch-pad mounted under the fingerboard and offering two axes of continuous control, several extra buttons, several slide sensors, force-sensing-resistors, turn "pots", and a biaxial accelerometer. Signals from these sensors are scaled into MIDI continuous control data. The instrument is modified to the needs of the actual performance situation. "The configuration of the sensors and the computer performance interface is constantly changing and developing in a way analogous to the musical development of an improviser from performance to performance" (Bahn, 2010, para. 4).

Analysing Bahn's playing movements and the sounds produced, one might hypothesise that he is not interested in sounds that are similar to those of a traditionally played string instrument. He rather aims to perform new sounds using an extended version of old playing methods. The instrument is played with both the traditional gestures of a bass player and with newly developed ones. It has been used in more than 300 concerts on international stages, for example at the Boston Cyberarts Festival in 2001, the Orchestra Tech Festival in New York City in 2001, and the International Society of Bassists Convention, Contemporary Music Series, in Richmond in 2003 (C. Bahn, email communication, October 10, 2009). The Sbass was played by Curtis Bahn.

4.4.4 Digital Stradivarius Project

In 1998 Bernd Schoner started the project *The Digital Stradivarius* within the context of a doctorate (Schoner, 2000b). As a starting point the violin was viewed as an analogue input/output system with specific constraints that processes input data to generate sound. Subsequently, the idea was born to build this system with a computer. One of the goals of the project was to capture the physical actions of the performer in order to provide explicit input data for the computer, and then to generate the sound of a Stradivarius violin using algorithms representing the specific constraints of a Stradivarius violin. The aim was "to build a model that, from the point of view of a listener or player, appears to obey the same physical laws as the acoustic instrument" (Schoner, 2000a, p. 376).

In collaboration with the researcher Neil Gershenfield, Schoner developed a synthesis method called *Cluster Weighted Modelling* (CWM). CWM falls in between the synthesis categories of physical modelling and sound sampling or wavetable synthesis. A probabilistic network is developed, which is able to map a set of input data, in this case gestural data of the performer, to a set of output data; the perceptible parameters of an audio signal.

The network is trained by applying measurement data from a player's actions and recordings (audio signal data) of a Stradivarius violin. The training of the network generates the constraints between input data and the data from the recordings. In order to train the system properly, sound sequences "that best represent a certain playing situation" (Schoner, 2000a, p. 376) must be chosen. It was hoped that once the most relevant constraints had been established, the system would allow a performer playing the interface to achieve the sound result of a Stradivarius violin. In addition, when trained with other sounds, the system was intended to be capable of playing, for example, a trumpet with the violin interface.

In answering the question of which measurements must be made in order to capture the performer's actions, Schoner was guided by three considerations which are repeated here in a shortened form (Schoner, 2000b, p. 136):

1. Collect all input parameters that a performer applies to a violin and identify the most relevant ones by establishing a hierarchy.

2. Define the optimum physical measurement technology adequately for each parameter to be measured. (Schoner mentions: “Some parameters may not be measurable at all.” (p. 136))

3. Bear in mind that the signal-to-noise ratio may disturb, and in some cases destroy, the measurement.

The instrument’s input system is based on slightly improved implementations of the Hypercello technology (p. 173). Using the method of electric field sensing, bow speed, position and pressure are tracked. With regard to the left hand, the position of the finger pressing a string down onto the fingerboard is determined by measuring a varying voltage. A resistive stainless steel strip is mounted on the finger board. Depending on where the string is pressed onto the steel strip, the overall resistance and thus the voltage to be measured changes. The audio signal from the four strings is captured by small magnets mounted below the strings.

Although Schoner would have liked to be able to measure the left hand finger pressure, because this pressure can affect the pitch, this requirement was not implemented because there was no appropriate method of measurement available (Schoner, 2000b, p. 136). Similarly, the measurement of the rotational bow angle driven by the wrist and the angle determined by equal or non-equal bowings was omitted because there was no appropriate method of measurement available, and because it was assumed that these factors would implicitly be captured when measuring bow position, speed and pressure.

For recording, a violin equipped with sensor technology was used (see figure 4.6, left side), while a cello inspired controller (see *Marching Cello*, figure 4.6, right side) was built for use in performances.[17] The synthesis engine operates with *Cluster Weighted Sampling* (CWS). This method is an improved version of CWM, and is able to process the inputs in real-time and generate the sound according to the pre-trained network.

Figure 4.6: Traditional violin with sensors and Marching Cello.

[17]Images in figure 4.6 by Bernd Schoner (2000b, p. 141 and p. 150).

Cluster Weighted Modelling and Cluster Weighted Sampling both use sinusoidal and wavetable representations as a basis for the prediction and reconstruction of the synthesised sound output. In comparison to pure sampling or additive synthesis the system was found to be more flexible. A pure sinusoidal representation has difficulty generating the transient phase of sounds, while wavetable synthesis is less flexible and less responsive in terms of continuous player input. Since the wavetable representations are used when the start of a tone is detected, and are replaced after about 0.5 seconds by the sinusoidal representation, the system can react with more flexibility to continuous control. Furthermore, it can generate more complete instrumental models (Schoner, 2000a, p. 379). The cello controller was used in demonstrations of this synthesis method.

4.4.5 Eviolin

The *eviolin* project was started by Camille Goudeseune in 1998. The idea of the inventor was to "take advantage of existing performance skill" to play sound synthesis (Goudeseune, 2001, p. 174). In addition, it was intended that the instrument should work well in chamber music. The development of the eviolin is included in Goudeseune's doctoral thesis, which presents a theory for designing musical instruments making use of sound synthesis. As Goudeseune's project addresses "the problem of controlling an n-parameter synthesizer with fewer than n (two or three) controls" (Goudeseune, 2001, p. iii), it is clear that the design of the instrument will be based on a set of parameter outputs.

Similarly to the Digital Stradivarius (section 4.4.4), the instrument is seen as a device connecting input parameters to output parameters. However, the model used here is not merely based on gestural input parameters, but also on parameters derived from the audio signal of an electric violin. In addition, it differs from the approach used with the Digital Stradivarius in the sense that the mapping is explicitly defined, rather than being defined by the training of a probabilistic network.

The hardware of the instrument is based on the body of an electric violin built by luthier Eric Jensen (2010). The audio signal from the electric violin is analysed with respect to fundamental frequency, amplitude and spectral brightness (proportion of higher frequencies in relation to the fundamental frequency). Parameters relating to gestural activities are gathered by using an *Ascension SpacePad motion tracker* (Ascension, 2010). This device measures the position of sensors relative to an antenna emitting a time-varying magnetic field. One sensor is mounted on the back of the body of the Jensen electric violin. A second sensor is placed on a glove worn by the violin player on the bowing hand. The SpacePad measures 12 dimensions (two sensors each measuring x, y, z, yaw, pitch and roll). The approximate position of the violin bow is computed by subtracting bow position from violin position (vector subtraction). The bow speed in turn is computed by subtracting this value from the value measured a moment earlier.

Two computers, connected via OSC (Open Sound Control) (OSC, 2010), are used for sound synthesis and sound processing. A Macintosh computer running MaxMSP provides algorithms for sound processing based on resonance filters,[18] delay lines, low- and high-pass filters. Sound synthesis is done on a Linux PC connected to a Macintosh and using the real-time sound synthesis package *VSS* (Bargar, Choi, Das, & Goudeseune, 1994). FM synthesis, additive synthesis (Goudeseune, 2001, p. 198), double modulator FM, a clarinet of the *Synthesis ToolKit* (Cook & Scavone, 2009), a chant synthesiser and a theremin-like generator are used. Spherical speakers (four channels) are used, and sound is spatialised according to the position of the player (Goudeseune, 2001, pp. 196-198). Furthermore, the spatial position of the electric violin can modify filter parameters of the resonance model.

Goudeseune presents a specific method for mapping parameters coming from the interface to the input parameters of the synthesis algorithms. He believes that the performer's intuitive understanding of the instrument can be increased if a compound mapping cross-couples several of the synthesis unit's controls. The compound mapping is seen here as a continuous function built by "associating several sets of d control values with corresponding sets of e parameter values (i.e. sounds)" (Goudeseune, 2002, p. 85). This method is then extended by a so-called "simplicial interpolation": a geometric technique which allows for continuous mapping. In order to adapt the mapping to a specific context, the mapping can be modified by moving existing points' geometric structure or by adding new ones.

Using different synthesis methods, sound effects and mappings, eight different instruments were developed (Goudeseune, 2001, pp. 199-204). Describing the eviolin Goudeseune mentions:

[18]CNMAT resonators~ MSP object (Freed, 2008).

> The latitude and longitude of the violin is mapped to timbre roughly according to perceptual dimensions: spectral brightness (relative prominence of higher partials) varies with latitude, spectral richness (number of partials) with longitude (Goudeseune, Garnett, & Johnson, 2001, p. 3).

When playing parameters are mapped in ways which differ from the expected ones (not pitch to frequency for example) the playability of the instrument alters. With regard to the instrument named “Rarae Aves”, which still maps loudness to amplitude, but no longer links pitch to frequency, Goudeseune is convinced that “[t]his instrument is more difficult for a violinist to learn than one which maps pitch to pitch” (Goudeseune, 2001, p. 201). Goudeseune’s experiences show that “the result in what performers call the feel of the instrument, its responsiveness and controllability, its consistency, continuity, and coherence” (Goudeseune, 2002, p. 85), is crucially influenced by the mapping process.

Although Goudeseune has described some of the personal experiences he has gained with the eviolin through composition and performance, no formal evaluation based on an evaluation standard has been done (C. Goudeseune, email communication, November 15, 2006). The instrument was used in a composition by Guy Garnett called *E-Violin Study*. It was performed by the violinist Chad Peiper.[19] In 2001 the instrument was presented at Microsoft Research and at the NIME 2001 workshop in Seattle.

[19] A video is available (Goudeseune, 2010).

4.4.6 Overtone Violin

In 2004 Dan Overholt started to build the *Overtone Violin* as a research tool for his doctoral thesis. The Overtone Violin is a “specialized instrument that continues the evolution of the violin” (Overholt, 2005a, p. 34). It consists of an electric violin equipped with several sensors, buttons and sliders as well as with a camera. The goal of the development is “to preserve the expressive elements of the expert violinist, while incorporating the added benefits of gestural controllers via embedded sensors” (p. 34). In addition, “[o]ne of the primary motivations behind the Overtone Violin is to put real-time signal processing under direct expressive control of the performer” (Overholt, 2005b, p. 604).

By using sensors Overholt increases the functionality of the Overtone Violin as compared to a traditional violin. The sensors invite the development and capture of a new and separate set of gestures for performing music. This can be done either by generating control data to be mapped to a synthesis engine or by applying these control parameters to the parameter inputs of signal processing algorithms.

The hardware of the instrument is based on a custom-built wooden body. The audio signal from six strings is transmitted using an optical pickup. This gives a warm sound and permits a separate audio signal without the need to mount a pickup directly to the strings, onto the bridge or below the bridge. The audio signals are pre-amplified directly on the violin. The machinery for tuning the strings is placed at the bottom in order to leave space for keys at the end of the fingerboard which can be operated easily with the left hand. The sensors used are a 2D accelerometer x/y, 2 channels of sonar distance sensors, knobs, faders and a joystick. The system is wirelessly connected to a computer. Figure 4.7 presents an image of the instrument.[20]

While this development offers a lot of new possibilities it does not focus on the question of how to capture the expressivity of the performer's actions when playing the Overtone Violin in a traditional way. "The philosophy behind this approach is to use gesture sensors to add completely different functionality to an instrument rather than capturing playing techniques that already have their own outcome, in this case, the sound of the strings" (Overholt, 2005b, p. 604). The development is, as it were, pushing the limits of violin playing by both providing an expanded pool of raw material for building stringed instruments, and inviting extended playing techniques, gestures, sounds and thus compositions.

The instrument has been used in several concerts on international stages, for example at the Dutch Electronic Art Festival (DEAF), Rotterdam in 2004, the ICMC in New Orleans in 2006, and the Spark Festival in Minnesota in 2007. Up until 2009, it had only been played by the developer.

4.4.7 Augmented Cello

The *augmented cello* was built by the researchers Adrian Freed, David Wessel and Michael Zybszynski (Freed, Uitti, Zbyszynski, & Wessel, 2006). During development they "decided not to try to measure and track traditional cello-

[20] Image from Dan Overholt (2009)

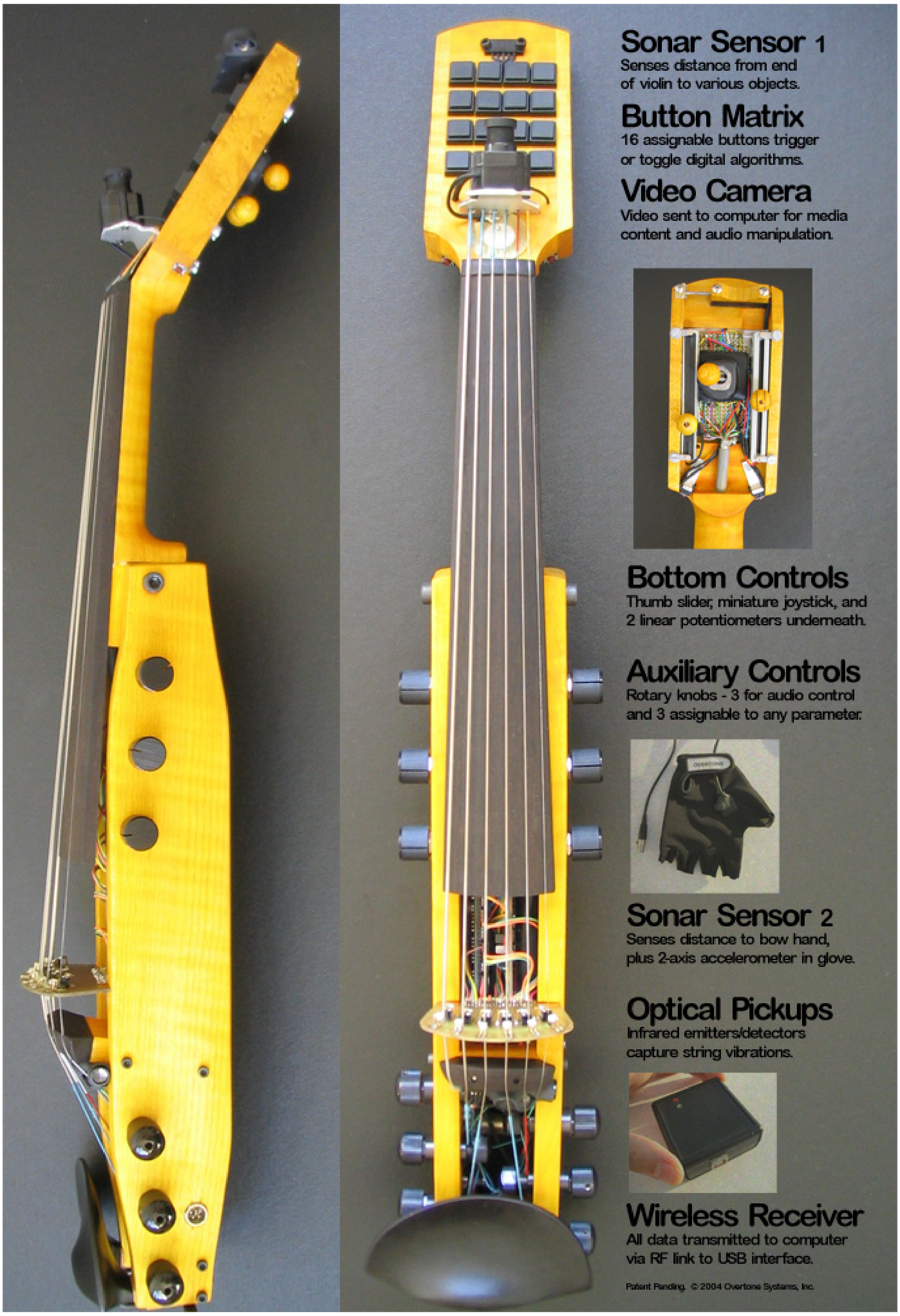

Figure 4.7: Overtone violin.

playing gestures but instead augment the instrument with new possibilities" (p. 410). To an electric six-string cello they added machinery to tune the instrument and a slide sensor at the side of the neck (as Curtis Bahn did with the Sbass, see section 4.4.3), as well as pressure sensors at the upper end of the body. They also added a button matrix and a wheel that can be driven by the bow to generate control data.

4.5 Using the Audio Signal

My doctoral research focuses on a synthesis method that is driven by the audio signal. Therefore, applications using an instrument's audio signal will be described in the following sections.

4.5.1 Max Mathews' Electric Violin

Computer music pioneer Max Mathews built several electric violins (Roads, 1980, p. 20). He developed specific pickups (Nichols, 2003, p. 50) and filters capable of generating sounds similar to high quality acoustic violins (Mathews & Kohut, 1973). One of his inventions is presented here because of its close relation to the Mu-Tron effect box explained in section 4.6.1, and to the approach to sound synthesis investigated in this thesis.

In most of Mathews' violins the source of the vibration is similar to that found on a traditional violin: the vibrating string. In one of his violins each string is picked up separately and the sound of each signal modified electronically. After pre-amplification, the audio signal is passed to a set of filters, each with two sets of coefficients. One set of coefficients can be adjusted to a fixed setting and the other can be controlled by extracted parameters of the audio signal. The output of an amplitude measurement of each string's audio signal is used to drive the dynamically operable coefficients of the filters. Using this method, Mathews was able to make the sounds similar to those of a brass instrument or even a human voice (Pierce, 1989, p. 159).

4.5.2 Meta Viola

In 1998 I started to build the so-called *Meta Viola* (Poepel, 1999). The goal was to create a synthesiser allowing string players to use as much of their available skill set as possible to play synthesised sounds, in other words, to be string-specifically playable. The instrument made use of modified versions of FM (frequency modulation) and subtractive synthesis, and was driven by the raw audio signal of a traditional viola equipped with a Zeta MIDI pickup system UR-205 Retro-Pack.[21] While I was interested in playing this instrument myself, it was planned and built to be used by other performers too.

The present thesis represents ongoing research initiated in the work done with the Meta Viola. In order to keep the description of the algorithms falling within the field of Audio Signal-Driven Sound Synthesis concentrated in one chapter, the methods and algorithms used in the Meta Viola are not explained here but later in sections 6.2.2 and 6.2.3. Figure 6.3 presents a

[21] The Zeta MIDI pickup system UR-205 Retro-Pack is now known under the name Zeta Jazz Series pickup system.

flowchart of the modified FM synthesis algorithm in the Meta Viola, and figure 6.1 presents the modified method of subtractive synthesis used.

To answer the question of whether the goal of string-specific playability was achieved, I personally evaluated the instrument. The sound result was compared with the sound result of a synthesiser driven by the MIDI output of the Zeta system. It was estimated that the string-specific playability was increased in comparison to the playability of the MIDI-driven method. The instrument was presented in demonstrations. Although composers showed interest in composing for the instrument, it was not used in concerts at that time.

4.5.3 Audio-Driven Timbre Generator

Tristan Jehan developed a synthesis engine for a new version of the Hyperviolin (see section 4.4.2). This instrument was intended to be used by the violinist Joshua Bell in the composition *Toy Symphony* by Tod Machover. One of the goals of the instrument was to overcome the weakness "either on [*sic*] the quality of sounds [...] or on [*sic*] the controls [offered] over the synthesis" when controllers such as the Yamaha WX5 wind MIDI controller or Roland GR-1 pitch-to-MIDI converter were used (Jehan, 2001, p. 15). In addition, the new Hyperviolin was designed to react "closely and intuitively to the player's music" (p. 22). To this end, the instrument controls its synthesis engine via the audio signal created by the performer on an electric violin, in contrast to the way the input device of the Digital Stradivarius Project (see section 4.4.4) uses the "input gesture on the physical instrument" (p. 22).

The specification of the system is based on the assumption that the "timbre of a musical signal is characterised by the instantaneous power spectrum of its sound output", and that "any given monophonic sound is fully described by the perceptual parameters pitch, loudness, and brightness and by the timbre of the instrument" (Jehan & Schoner, 2001, p. 382). The architecture of the system is built in such a way that the audio signal is analysed, and the features extracted from the audio signal then drive, without a specific mapping, the inputs of the pre-trained probabilistic network of a Cluster Weighted Modelling (CWM) synthesis. The features extracted are pitch (fundamental frequency), amplitude, brightness (spectral centroid) and noisiness (spectral flatness). The audio signal comes from the pickup of a Jensen electric violin (Jensen, 2010).

Since the system tracks features from the audio signal rather than gestures, Jehan concludes that an interpretation of the extracted features (fundamental frequency, spectral centroid and spectral flatness) is not necessary. The instrument has been trained to model the timbre of sounds such as a male and a female singing voice, a Stradivarius violin and woodwind instruments.

On the basis of the assumption that any given monophonic sound is fully described by the perceptual parameters pitch, loudness and brightness, and by the timbre of the instrument, Jehan believes that the sound output of the synthesis offers "identical perceptual qualities" (Jehan & Schoner, 2001, p. 23) to the sound produced with an electric violin. In addition, Jehan believes that the sound result "is close to the target instrument", preserves playability and "the musical intent of the player" and "behaves intuitively and predictably" (p. 58).

The violinist Joshua Bell took part in the development of an instrument consisting of the Jensen violin connected to the analysis of its audio signal and to the synthesis engine using CWM. Testing the instrument, Bell estimated that it is "certainly groundbreaking and is sure to inspire the minds of many musicians" (Jehan, 2001, p. 5).

4.5.4 Voice-Driven Synthesis

Although I am focusing here on bowed stringed instruments, a controller using the human voice will be described because it addresses the same task as the general focus of my research: that of using an audio signal to drive sound synthesis. In his doctoral thesis Jordi Janer proposes a singing-voice interface to drive sound synthesis (Janer, 2008, p. 7). He sets out to use the "high degree of expression and the nuances of the singing voice in order to exploit it as a musical controller" (Janer, 2005b, p. 132) and to drive synthesis. The aim is to determine and musically transport the performer's vocal intentions.

To achieve these goals he first "examines the characteristics of the captured acoustical signal, which at its turn, [*sic*] drives the parameters of a synthesis engine" (p. 132). For the feature extraction of the audio signal of a voice, first a decision is made on which features to extract. The "chosen features include basic attributes, such as pitch and energy, in addition to other timbre descriptors" (p. 132). For voice analysis a Phase-Vocoder based method is used.

The analysis of the singing voice is based on the assumption that the expression and nuances of singing and vocal intentions can be captured by measuring control aspects in excitation, vocal tract, voice quality and context. The following reason is mentioned: "Excitation descriptors are elemental for the user, since they are related to the instantaneous sung energy and fundamental frequency" (p. 133). In addition, Janer assumes that timbre is associated with a particular vowel in voiced sounds. Vowels are measured by estimating the first two formant frequencies determining the vowel (Janer, 2005a, p. 2). The voice quality is estimated by two algorithms, and an attack unvoicedness descriptor is used to determine the harshness of a synthesised sound (Janer, 2005b, p. 133). In order to perform sound synthesis, Janer uses an adapted Karplus-Strong algorithm as well as a Spectral Morph algorithm. The question of mapping is solved here by using both instrument-dependent and synthesis-dependent mapping.

While the implemented version of 2005 was tested by the developers and found to be usable as a virtual instrument for real-time performance it is mentioned that it was "still limited in terms of musical control" (p. 135). Thus, it had not yet been used on stage by the time of the paper describing it (Janer, 2005b).

4.6 Adaptive Audio Effects

Two developments will be described in the following two sections that again do not fall into the category of stringed instruments, but are of interest because they parallel the basic architecture of the ASDSS method.

4.6.1 Mu-Tron

The *Mu-Tron III*, an interactive effects box, was developed in the early 70s by the engineers Aaron Newman and Mike Beigel of the company "Musitronics" (Gill, 1997). Their goal was to build an automatic wah-wah pedal: a filter that can be controlled live by the input audio signal. As with Max Mathews' violin (section 4.5.1), the cutoff frequency of a low-pass filter was controlled by the measured amplitude of the audio signal to be filtered. The Mu-Tron III "used a state-variable filter because that way we [the developers] could get three different kinds of filter response out of it" and photo mods[22] were used to control the filter (Gill, 1997, para. 7).

The evaluation of the system took place in a pragmatic way. The developers took a prototype to dealers to see what they thought about the device. The Mu-Tron III sold very well. It was used in numerous concerts all over the world, by musicians such as Stevie Wonder, George Duke, Lee Ritenour or Frank Zappa.

[22]Photo mods were an early type of optocoupler that used internal photo resistors instead of photosensitive silicon diodes to transfer electrical signals by utilising light waves to provide coupling with electrical isolation between their input and output.

In the late 70s Musitronics started to develop a guitar synthesiser "that was not strictly a synthesizer because it did not use the guitar to control an oscillator to generate the sounds" (para. 19). The developers "tried to make enough sophisticated modifications to the sounds by basically combining sound effects to give the function of a synthesizer and keep the versatility of a guitar sound" (para. 19). Finally, they "were about 95 percent there" (para. 19). As they were not sufficiently convinced to release a product that was only 95 percent satisfactory, the instrument did not come onto the market. A developer explains: "Although a few people saw the product, which was called the Mu-Tron VII and looked like a Bi-Phase with more knobs, we never went into production because at that point we still hadn't perfected it" (para. 19). With regard to the basic idea of creating an instrument with the function of a synthesiser, but using the raw audio signal to drive it, the Mu-Tron VII can be described as an early version of the approach to sound synthesis investigated in this thesis.

4.6.2 Adaptive Digital Audio Effects

The method of *Adaptive Digital Audio Effects* (ADAFX) was introduced by the researchers Vincent Verfaille and Daniel Arfib in 2001 (Verfaille & Arfib, 2001). The aim was to extend the sound varieties of existing digital audio effects and create sounds that were interesting according to the perceptions of the researchers and of musicians.

To achieve this, an audio signal is fed into an effects unit. The control of the effect unit's parameters is performed on two levels. On the first level, parameters extracted from the audio signal are fed via mapping into the parameter inputs of the effect unit. In addition to driving the parameter inputs of the effects unit, the extracted parameters from the audio signal can be used to control the mapping process itself. So called "warping functions" (non-linear transfer functions) are used to transform extracted features according to specific rules. On the second level, parameters derived from tracked gestures can be used to control the mapping process of the extracted audio signal parameters to the inputs of the effects unit (Verfaille, 2003, p. iv and 213).

The basic ADAFX architecture is similar to that of the Mu-Tron III (see section above) and to my approach to sound synthesis. However, the effect algorithms are not designed and evaluated in order to preserve the musical intent which a player has put into the audio signal. ADAFX aims to create effects which produce interesting sounds with a specific effect character. The main questions in the development of ADAFX were (Verfaille & Arfib, 2001): which parameters can be extracted (p. 10)? How can these parameters ideally be mapped to the input parameters of the effect (p. 11)? Which effects generate good sound results when built adaptively (p. 12)?

Examples of effects gained by this method are: adaptive ring modulation, spectral warping, robotisation,[23] hybridisation and gender change (Verfaille & Depalle, 2004). With regard to evaluation, one may say that the aims of the developers were achieved, since new varieties of sound processing were

[23]To transform a sound (i.e. the sound of a human voice) into a machine-like sound.

created. However, a formal evaluation in a controlled environment was not found in the ADAFX publications.

4.7 Discussion of Related Research

The descriptions of related research presented in the previous sections raise several questions. There is, for example, the question of the goals pursued by the researchers with their instruments, and the methods used to achieve these. Other questions arise once the basic tasks such as formalisation or modelling that have been addressed in chapter 3 are taken into account. The present section will address questions raised by related research work and discuss those.

Looking at the many instruments described in the previous sections, it should be mentioned that research has produced a notable body of work. I am not aware that any of the presented instruments have been built in series or large numbers, but one might hypothesise that instruments which have been used on stage on numerous occasions might possess some notable qualities. While comparing numbers of performances should not be considered a reliable way to measure the relative quality of the instruments, one might reasonably imagine that those instruments which were used in a large number of performances were considered reasonably usable and "stage-worthy" by their performers. The Sbass (section 4.4.3), with more than 300 concerts on international stages, may be mentioned here as an example.

4.7.1 Aims and Requirements

Concerning the goals the developers had when building their instruments, it is observable that these goals vary. There are instruments that use the violin, or violin-like interfaces, but are not intended to offer a connection between performers' actions and sound as would be expected with a traditional string instrument. An example can be found in the sensor-bow of Jon Rose (section 4.3.1). While tracking an essential part of violin playing—the movement of bow—Rose does not aim to generate sounds that correspond to the shape of the sound one would hear from a traditional violin. The success of his development is not characterised by the sound matching the actions of the performer in correlation to the expressive potential known from traditional instruments (Rose, 2005, para. 5).

Besides this, there is the goal of building instruments that on the one hand offer the possibility of making use of traditional playing techniques with corresponding sounds, but on the other hand have their potential in offering new playing techniques (e.g. new arm, hand and finger movements) and corresponding new sounds. Examples of this would be Dan Trueman's BoSSA (section 4.3.4) or Dan Overholt's Overtone violin (section 4.4.6).

Some researchers use bowed stringed instruments as a basis for their developments because they want to use the existing skills of string players and offer instrumental qualities known from traditional string instruments. Since the subject of this thesis, the method of ASDSS, also follows these goals, developments falling into this category will be looked at more closely now.

How do developers describe such goals, and thus parts of the instruments' requirements? With regard to the development of the vBow (section 4.3.6), Charles Nichols writes: "Since I am a violinist, it is important that I use a controller that takes advantage of the years of training I have had as a violinist" (Nichols, 2003, p. 36). Part of his goal is to build the instrument in such a way as to "maximize the expressive potential of the instrument" (p. 33). Since he has learned "how to use a violin bow to subtly vary [his] performance, producing a wide range of dynamics and types of musical expression" (p. 36), he seeks to build "an instrument that would sense [his] bowing gesture with acute resolution, affording [him] the greatest dynamic and timbral range possible" (p. 33). In conclusion, the instrument should allow the performer to make use of existing skills and should offer qualities found in traditional string instruments such as their expressive potential or the ability to offer the performer a subtly varied performance.

According to the goals mentioned in the descriptions of related research in the sections above, instruments designed to use the existing skills of performers and to offer qualities known from bowed stringed instruments include the Hyperbow (section 4.3.5, goals p. 60), the Hypercello (section 4.4.2, goals p. 68), the Audio-Driven Timbre Generator (section 4.5.3, goals p. 85), and the Digital Stradivarius Project (section 4.4.4, goals p. 72). Since these instruments overlap with the present research work in their goal of using the existing skills of performers, I will focus on these instruments in the sections which follow.

4.7.2 Formalisation

Given that formalisation is a necessary step in the development of a computer-based instrument, the question arises as to how formalisation has been achieved. With respect to the stages of the above-mentioned design model, one may ask for the transformation from described aims (requirements level) into, firstly, the specification, secondly, a formal specification and thirdly, the development result.

I view two areas as most interesting with regard to the question of formalisation: firstly, the question of how the developers specify objectives at the requirements stage, because this transition from the non-formal to the formal description of objects involves abstraction and the use of models. If, for example, the requirements include aims such as “react to subtly varied performance” or “give extra power and finesse to a virtuosic violinist” (Jehan, 2001, p. 22), it is interesting to ask how specification has been done, because such a requirement does not in itself offer a formal structure that can be implemented directly.

Secondly, as outlined in section 3.7, the formalisation of measurement is a critical task and it might be worthwhile to ask how developers address this problem. In the case of the Audio-Driven Timbre Generator (section 4.5.3), for example, one might ask how latency issues affect the playability of the system. While Jehan (2001) states that with his new development “the playability of the [traditional] instrument is preserved” (p. 58), it is hard to imagine that the system is unaffected by the latency of the tracking algorithms Jehan uses.

It would be of interest to consider the question of whether the goals of the developments have been achieved, and whether the formalisation used works properly according to the requirements. However, since the published details of the related research addressed here did not offer formal evaluations with multiple participants in a controlled environment, there is no data on which to base an answer to this question.

4.7.3 Specification and Measurement

What insight does a survey of related research provide into specification and measurement? With regard to the documentation of the development process, I notice that some researchers sketch the goals of their instruments briefly and then present the development results. The translation from the level of requirements to the level of code and interfaces is not explained and the question of how the implemented system meets the requirements is not answered. Other researchers, however, mention models they selected in the design process.

In order to use the "most intricate aspects" of violin playing and the "subtle elements of physical gesture that immediately and directly impact the sound of the instrument while playing", Young selects a physical model of violin playing (Young, 2001, p. 5) to specify the requirements of the Hyperbow (section 4.3.5). The model was developed by the physicists Schelleng (1973) and Askenfelt (1986) while they were researching the interaction of the bow with the string of a violin. Young's specification includes the definition of the essential bowing parameters relevant to the sound according to this model. The formal specification includes the definition of the measurement apparatus needed to measure these parameters. The questions of the completeness of the model used (feature of reduction, see section 3.6 p. 36) and its validity in relation to the context in which it is used (pragmatic feature, p. 36) are not found to be addressed nor the problems that can result for a user due to the limited model used in the design process.

A similar approach to translating the requirements is used by Nichols in the development of the vBow (section 4.3.6). In parallel to Young, Nichols makes use of physical models of violin playing and defines according to these models the essential parameters that must be measured and the way these parameters control the synthesised sound (Nichols, 2003, pp. 108-121). The problem of blindness which accompanies abstraction (see section 3.5) and thus also these models, is not addressed.

The eviolin (section 4.4.5) seeks "to take advantage of existing performance skill" (Goudeseune, 2001, p. 174) in the player's use of sound synthesis. The parameters defined for measurement are the position and orientation of the physical instrument and the bowing hand. Based on this measurement the approximate contact point of bow hair on string (bow position) and bow speed are calculated. The audio signal of the instrument is analysed by measuring the parameters pitch, amplitude and spectral centroid (Goudeseune, 2001, p. 179). An explicit explanation of why and how these parameters relate to the requirements of the instrument is not found. This raises the question of the assumptions the developer had when translating the requirements to the specifications. It is remarkable that bow pressure is not measured, despite the fact that Goudeseune is a trained violin player and will thus have been aware of the importance of bow pressure for the sound result. Looking at the measurement devices used one might speculate that he reduced the formal specifications to correspond with what the apparatus used allowed him to measure.

Jehan uses timbre modelling in the Audio-Driven Timbre Generator and assumes that "the timbre of a musical signal is characterized by the instantaneous power spectrum", and that "any given monophonic sound is fully described by the perceptual parameters pitch, loudness, and brightness and by the timbre of the instrument" (Jehan, 2001, p. 29). It is a well-known psychoacoustical phenomenon that the perceived timbre is affected by time-dependent features such as onset, envelope and articulation. Thus, the question is raised as to whether the model of timbre Jehan uses includes a specific blindness concerning performed timbre. To what extent will such a model react adequately if a player's input makes use of time-dependent features such as onset, envelope and articulation in order to create a specific sound? How will the model react during the non-steady state of a violin tone, or a relatively noisy tone such as a fast tremolo in pianissimo? These questions seem important to me as a performer because the steady state of a tone is surely not the entirety, but just a specific part of the tones with which I am working.

The system requirements of the Digital Stradivarius Project (section 4.4.4) differ from other requirements in that they include only elements that are, in themselves, specified. The violin is considered from the outset as an "analog computer" in which the physical input is connected with the physical output via computation (Schoner, 2000b, p. 133). The system requirements do not include verbal descriptions by performers with respect to the instrument. Since the instrument is—according to the developers' definition—a computational input/output system, the requirements may only include objects (models, parameters, algorithms) that a computational system can process. Because such objects already have the structure of a formal specification (the precondition for their computability in a computer), the variety of objects to be put into the requirements document is reduced, and does not include natural language descriptions. In other words one might say that the requirements are already pre-specified by their definition in a physical language using computable objects.

Schoner is a trained cellist. I assume that he will have known that a number of physical playing parameters exist and can impact the sound in performance. According to Schoner, "defining and carefully executing [the physical input] measurements" (p. 136) is crucial to the development of the instrument, because it is not necessarily possible to measure all parameters, and the question must be answered as to which of the parameters are more or less relevant. In order to tackle this problem he suggests that the "search for optimal measurement" (p. 136) be conducted by defining a hierarchy of parameters, and by selecting the measurable parameters according to available and valid methods of measurement. Thus, he solves the problem of the reductive feature of models by defining a hierarchy of parameters and by pragmatically reducing the measurements to those that work properly. However, the pragmatic feature and thus the question of the context in which the defined hierarchy is adequate is not addressed. The problem, for example, that a specific tone or playing technique might require a different parameter hierarchy, and that—depending on the tone or playing technique—different methods of measurement might be appropriate, is not considered. If a performer uses a different set of physical input parameters or a different hierarchy of playing parameters, it might be that the instrument is found to be unsatisfactory because of an inadequate response to the performer's physical input.

In summary one can say that in cases where the researchers explicitly mention the models used in the specification, these models are physical ones. In view of the fact that the pragmatic feature and the feature of reduction were not addressed by the researchers, the question remains open as to how far the models used are in fact suited to meeting the needs of the requirements. This is particularly the case when the requirements include natural language descriptions such as the need to provide performers "with an amplification of their musical intent" (Nichols, 2003, p. 12) or to build a device that is "capable of measuring the most intricate aspects of violin technique" (Young, 2001, p. ii).

4.7.4 Mechanical Models of the Performer

Studying research that states the use of physical models in specification, I noticed that these models are embedded in mechanical models of performers. The mechanical image of a performer drawn by Nichols (2003, pp. 36-38), for example, describes the performer as a unit essentially applying measurable physical forces to the instrument, considering the arm as a “bowing machine” (p. 37). This is an understanding of performance which is not necessarily in accordance with all successful concepts of instrumental teaching.

From my perspective as a performer who was—like many others—trained as a viola player with an organic rather than a mechanical idea of playing, it is astonishing that the mechanical model of the output apparatus of a performer plays such a role in the developments discussed here. The violinist Biesenbender (1992) explains the problems which may arise for a performer if she understands herself as a mechanistic device (pp. 18-20). He looks at the history of human self-concepts and concludes that the description of a mechanical string player is more influenced by the dominant contemporary world view than by the essential elements of playing a bowed stringed instrument. In order to create an instrument which is responsive to the actions of the performer, the question of what precisely a performer does when playing is highly important.[24] Thus, the question arises as to how far the abstraction of a performer seen as a mechanical model may cause a problem in the design of computer-based instruments.

4.7.5 Playing or Controlling

While analysing related research I noticed that the term *controller* is used widely and conveys the idea of controlling a number of discrete playing parameters. Researchers such as Nichols (2002, p. 215) or Jehan and Schoner

[24]The question of what a performer does when playing is worth considering in more depth. As this chapter discusses related research, the pedagogical descriptions of performers actions will be looked in more detail in section 5.3.1.

(2001, p. 386), for example, refer to the musical interface as a controller. The colloquial term for controlling an instrument is playing an instrument. The pianist is said to play the piano, rather than control the piano. One may ask whether there is a difference.

If somebody is controlling something, I consider that the aim is to have everything under control: one has it in one's hands, it cannot slip away, get damaged or do something one does not want it to do. One *has* it, one knows and can control what is going on. In an interface all control data can be viewed exactly and explicitly. In playing a traditional musical instrument, however, it may happen that things are not always under control, and one must expect that unwanted situations may occur and need to be dealt with.

An example to illustrate the action of "playing" may be found in a ball that is bounced on the floor. One knows roughly how the ball will rebound but not 100% exactly. According to the above-mentioned understanding of controlling and playing, controlling would mean controlling the ball completely, while playing would mean having the ball partly, but not entirely, under control. Not having the ball under complete control, but seeing what happens when the ball bounces and reacting to unexpected situations can be fun.[25]

Many musicians understand a performance as an object that has to be brought to life in an improvisatory manner (Biesenbender, 1992, pp. 46-47).

[25] A similar situation can be found in football. If football players were able to control the ball completely, part of the tension and fun of the game would be gone.

The cellist Pablo Casals mentions that he has to learn to play the instrument every day, again and again (Casals, 1983, p. 10). In my experience as a violist, a major part of the performance of music is playing with the resistance of an instrument, and going to the limits of tone creation, timbre and desired sound in order to obtain the highest musical quality. Inherent in this method is the risk of losing control. The conductor Nicolaus Harnoncourt reflects on this issue as follows (Harnoncourt, 1992, Track 14, 3:38):

> It is my opinion that two incompatible opposites determine the quality of a performance in music: striving for security—avoiding a cracked or wrong note—and aspiring to beauty or truth. These two things are incompatible.[26] [Translation by Cornelius Poepel]

In his experience, for the performance of music to achieve a magical quality, it is necessary to take an extremely high risk. He explains that he seeks:

> at a very high level of playing to forget about technique and risk the maximum. For example, if you think about playing high notes in pianissimo with wind instruments—that is a big risk. There is a point at which the instrument does not respond properly anymore. The musician does not—precisely—know where this point is. Therefore he prefers to stay in a safe zone, and so he is actually too loud for the music. Or [the tone] can crack, and then everybody can hear that something unintended has happened. The true magic, the unbelievable in music, happens at the very brink of a catastrophe. But one has to understand that beauty

[26]Ich bin der Meinung dass in der Musik zwei unvereinbare Gegensätze die Qualität der Aufführung bestimmen. Und das ist das Streben nach Sicherheit – dass kein Kiekser entsteht, dass kein falscher Ton entsteht durch [*sic*] – und das Bestreben nach Schönheit oder Wahrheit. Das lässt sich nicht vereinbaren.

> and security are incompatible with each other.[27] (Track 14, 5:06)
> [Translation by Cornelius Poepel]

I conclude that the area of insecurity, of incomplete control, is for some high-level musicians, at least, an area in which they will often play in order to achieve the desired quality of music. It is the task of the player to enter and explore this area. The better the musician is, the further she can go. However, she never exactly knows the breaking point at which an undesired sound result will occur. The instrument is partly under control and partly not. In parts, the instrument is under control, in parts it is not. In parts the musician knows what will happen, in parts the musician does not know what will happen. The musician generates quality by risking losing the fight with the music and the instrument, walking on the border of disaster. Thus, I consider the concept of "playing"—which essentially includes having incomplete control—as important in the performance of music, and one may ask whether the idea of controlling an instrument completely with a controller in fact covers the whole spectrum of the performer's interaction with a musical instrument.

In Harnoncourt's terms, an instrument that does not offer this uncontrolled and "free-running" area will not allow the musician to produce this

[27] auf einem sehr hohen technischen Niveau die Technik zu vergessen und das Äusserste zu riskieren. Es ist zum Beispiel, wenn sie denken, hohe Töne auf Blasinstrumenten im Pianissimo zu spielen, ist ein grosses Risiko. Da gibt es einen Punkt wo das Instrument nicht mehr anspricht. Diesen Punkt – ganz genau – weiss der Musiker nicht. So bleibt er lieber in einem Sicherheitsbereich, dann ist er für die wirkliche Musik ein bisschen zu laut. Oder es kann abreissen, dann hört jeder Zuhörer, dass da etwas passiert ist, was nicht vorgesehen war. Die wirkliche Magie, das Unglaubliche in der Musik, das ereignet sich an der äussersten Kante zur Katastrophe. Aber wir müssen uns klar sein, dass Schönheit und Sicherheit zusammen nicht gehen.

desired quality. Thus, putting this—the idea of building a controller with which the performer can have everything under control, or even play music which is highly expressive with no great effort—on the list of requirements seems problematic because it represents a limited model of instrumental performance and—at least according to Harnoncourt—one at odds with the thought processes essential to the creation of music of the highest quality.

4.7.6 Views on Musical Instruments

Descriptions of musical instruments and performer-instrument interaction offer insight into views on and thus abstractions of musical instruments. This section will present some views on musical instruments in order to illustrate different possible abstractions of instruments and raise the question of whether a purely functional and physical understanding of an instrument might cause problems. In asking what acoustic instruments can do and how they can be played the researcher Jordà (2005), one of the main developers of the reacTable (see section 4.1.2), provides the following view:

> Most acoustic instruments consist of (a) an excitation source that can oscillate in different ways under the control of the performer, and (b) a resonating system that couples the vibrations of the oscillator to the surrounding air. Parts of the performer's body generate the energy for the excitation of the instrument, interacting through the instrument's control interface. (p. 20)

This model is not wrong or false with regard to the physics which it concerns. However, it is reductive in the sense that the whole instrument is described by the words "consist of", while the abstraction covers only specific acoustic properties. In order to explain "*how these digital instruments can easily go beyond limitations governing acoustic instruments*" (p. 19), Jordà tries to answer the questions of "*what* [acoustic instruments] can do (output), and *how* they can be played, i.e. how their sonic possibilities (input) are controlled" (p. 20). I assume that in his understanding of playing an instrument the above mentioned uncontrolled area does not play an important role.

With regard to the question of *what* acoustic instruments can do, Jordà argues that the criteria of pitch variation, dynamics variation and timbre variation (p. 22) are sufficient to describe, "with very few exceptions, their simplified possibilities" (p. 21). He further remarks that the sonic outputs of digital instruments can cover "all the aspects" included in his list of acoustic instruments' possibilities "without any restrictions" (p. 25) and add several new possibilities, for example in "parameter variation", "small periodic changes", "unlimited and precise control" and the potential for "[a]ny imaginable timbre" (p. 26). While digital sound will certainly offer spectra which acoustic instruments cannot produce, one might ask whether Jordà's understanding of acoustic instruments in fact addresses all aspects essential to the quality of traditional acoustic instruments.

Interestingly, the computer music pioneer Max Mathews expresses his view on the limitations of traditional instruments in Buxton et al. (2000, p. 625) as follows:

> I believe that traditional instruments and electronic controllers are equal in their sensing sensitivity. I also believe that electronic sounds overcome basic limitations in instrumental sounds. A violin is beautiful, but it always sounds like a violin.

It might be worth considering the question of whether Mathews understands a traditional instrument as an input device with a limited number of parameters to measure. Furthermore, Mathews' belief is dependent on a proper formalisation of measurement. As explained in section 3.7 (p. 40), however, according to Prem (1997, p. 5) "the construction of measurement devices and the measurement process itself cannot in principle be formalised". Thus, the question arises as to how electronic controllers can be expected to have an equal sensing sensitivity if one assumes that the formalisation of the measurement process will always be to some degree wrong or incomplete.

It is undoubtedly right that a violin always sounds like a violin. However, musicians may perceive the violin to be far more flexible in terms of sound than for example a MIDI violin which offers the performer the whole world of MIDI synthesiser sounds. Once adapted to the framework of a traditional instrument's sound variance, the performer can play with a huge variety of timbres.[28] While Mathews talks about the limitation of a violin always sounding like a violin, the violinist Yehudi Menuhin describes the fingerboard of his instrument as a "galaxy" (Menuhin, 1985, p. 48) to be explored for the creation of sounds. I conclude from this that models describing the differences of perceived sound may vary from person to person, and are thus dependent on context. I conclude further that, at least for some musicians, the question is more about how an instrument behaves within a specific range of sounds (i.e. the sound of a violin) than about how many different synthesiser sound-presets an instrument offers.

[28]The generation of the many different variations inside the spectrum of sounds possible with a traditional instrument was the subject of many of my viola lessons and music rehearsals.

Answering the question of which instrument offers the greatest variety of timbres, the famous singer and conductor Bobby McFerrin remarks in an article on his research into sound as an artist: "Ultimately, the human voice offers the widest timbre palette of all instruments"[29] (Klot, 2007, p. 103) [Translation by Cornelius Poepel]. Given his wide musical experience, one can hypothesise that Bobby McFerrin is familiar with synthesisers and their capabilities, and that he must have experienced sounds which can be produced by a synthesiser or sampler, but not by a human voice. Despite this, he sees the human voice as the instrument with the widest range of timbre.

I conclude, that his argument is not based on a physical understanding of sounds, because seen from the point of view of a physical description of sounds, the variety of sounds that can be generated with sound synthesis is far greater than the variety of sounds one can generate with the human voice. One might further hypothesise that he is arguing here from a phenomenological point of view, where the question of the meaningfulness of a timbre is essential, and where the phenomenologically perceived variety of timbre palette forms the basis for his statement. In this case, the model underlying his assessment may not be the physical model of the identity and timbre of a sound, but rather a phenomenological model which corresponds with his musical experience.

[29] "Die menschliche Stimme verfügt schließlich über die größte Farbpalette aller Instrumente."

Kirk and Hunt (1998) mention the problem of a restricted view of human-instrument interaction, and propose that one might learn from the interaction between a traditional performer and the instrument. They describe, not without irony, a "brave engineer" (p. 150) who understands "the process of playing a violin" as follows:

> A violin is a multiparametric acoustic signal processor, achieving significant throughput of hard real-time data. The operator controls the machine through a multivariate interface by means of trained reflex operations, in response to perceived auditory cues interpreted through psychoacoustic analysis, and in response to a complex graphical notation which encodes temporal, spectral and gestural information. The operator typically requires ten years of training to acquire a reasonable facility in using the system. (p. 150)

The system Kirk and Hunt (1998) propose is an audio-visual performance instrument based on a graphically extended version of Max Mathews' unit generator metaphor MUSIC-N.[30] Despite the fact that their system allows an interface of a similar complexity to Jordà's reacTable, they are convinced that they have not reached or overcome the limitations of traditional instruments, but rather have reached the "[l]imitations of [c]onventional [c]ontrol" (p. 160) of real-time computer music systems.

The different views on how to understand instruments and the contradictions arising from such views, raise the question of why such different views exist; why researchers in the computer music area tend to use the physical and functional abstraction; what this kind of abstraction means for the de-

[30]The MUSIC-N metaphor is explained in section 5.1.1.

sign process of instruments; and to what extent the physical and functional abstraction will meet the needs of musicians working with a different understanding of an instrument. As the instruments I have focused on in this chapter seek to allow the user to make use of pre-existing skills, I think it should be possible for the user to continue with the methods of working that she is used to, that is to say that she should be provided with an instrument which can be used with her already existing understanding of playing an instrument.

4.7.7 Conclusion and Questions

In conclusion I have found that some but not all related works refer to models used as a basis during the process of designing interactive digital stringed instruments. Those models mentioned explicitly are based on physical abstractions of musical performance. It is according to these models that the essential playing parameters to be measured are selected, defined and formalised, and the measurement apparatus is built.

Putting the method used in the development of the instruments discussed in related research into a very short form, one may describe the implementation principle with the following words: the quality of the instrument is generated by defining and formalising essentials. The problems of blindness in abstraction, and of possibly incomplete models, are not found to be addressed, and the question of the extent to which the formalisation of the process of measurement creates a problem is not found to be addressed either.

These chapters discussing related research have presented several questions. Since natural language descriptions of requirements do not in themselves offer a specification or formal structure, the question that remains most interesting to me is that of why the researchers specify the requirements merely by using physical and functional abstractions of the player-instrument interaction. Further questions which I assume to be of importance are: are there alternatives to this method of specification? With regard to the problems that may occur as a result of using incomplete models, and to problems of formalisation in measurement: how exactly do these problems look if one examines a computer-based instrument in detail? If there are shortcomings one may expect that there are researchers confronting these shortcomings: how do they describe these problems and what solutions do they propose?

In order to find answers and alternatives, the following chapter will address a selection of the questions raised by the related research, and will propose an additional approach to conceiving sound synthesis.

Chapter 5

Approaches to Digital Sound Synthesis

This chapter relates to the questions posed in the previous chapter's examination of related research. Firstly, the conventional parameter-driven approach to sound synthesis will be described in section 5.1. Following this, I look at the history of sound synthesis (section 5.1.1) and at formalisation strategies (section 5.1.2) in order to ascertain whether one might discover why sound synthesis is currently constructed primarily in such a way that essentials are formalised and physical and functional models are used to build the instruments.

In order to be able to give a precise illustration of some practical problems that may arise when the actions of the performer are modelled and a system is built to track them, the performer's actions and their measurement are detailed in sections 5.2.1 and 5.2.3. Sections 5.1.3 and 5.2.4 will show where researchers see problems in the conventional parameter-driven approach, and which ideas are pursued in order to solve these problems and make the systems more open to the input of the performer.

With the aim of adding a phenomenological abstraction used by string players in instrumental teaching to the physical abstraction of string instrument playing, section 5.3 focuses on the question of what actions are made seen from the perspective of performers. Additionally, section 5.3.2 draws conclusions on what this understanding of instrument playing might add to a requirements document for a computer-based string instrument.

Section 5.4 then introduces the signal-driven approach to sound synthesis, and describes main differences from the construction principles of the parameter-driven approach. In order to outline the basis on which the signal-driven approach will be examined in the present research, the hypothesis of this thesis and surrounding questions are detailed in section 5.5.

Since digital sound synthesis plays a central role in this thesis, I first want to provide a definition of what is understood here by the term “sound synthesis”. Based on the ancient Greek word “synthesis” (composition, aggregation, catenation), where synthesis is understood as the putting together of pre-existing parts to form a new whole, I understand digital sound synthesis here as a method by which to create electronically generated sound. In this sound, input data, computer hardware and an algorithm are put together in a way which allows a user to create a meaningful sound—that is to say a sound with a certain quality which makes it musically useful to the musician working with it.

Digital sound synthesis is conventionally understood in a more specific way, in which a sound is generated by a synthesis engine driven by a set of discrete input parameters. As mentioned earlier, however, this thesis is written with respect to the perspective of the musician using synthesisers. Thus, an application producing a result which is perceived—according to the definition of sound synthesis—as a new sound put together from pre-existing parts, is seen here as a valid synthesised sound. The conventional understanding of sound synthesis is expanded here by allowing the use not only of the discrete input parameters but also the raw audio signal to form the new whole.

5.1 The Conventional Parameter-Driven Approach

The present research addresses sound synthesis in real-time usage. Figure 5.1 shows three areas of focus which are commonly found in parameter-driven synthesis systems. These areas are the measurement apparatus (for measuring the input of the performer), the mapping unit and the synthesis engine. Explicit and discrete parameters are generated by the measurement unit, and are mapped to the discrete parameter inputs of the synthesis engine.

Once the parameter data (fundamental frequency and amplitude) of a Zeta MIDI violin, for example, is provided by the player via the measurement apparatus, it must be mapped to the input parameters of the selected synthesis method. Hunt and Wanderley (2002) mention three basic meth-

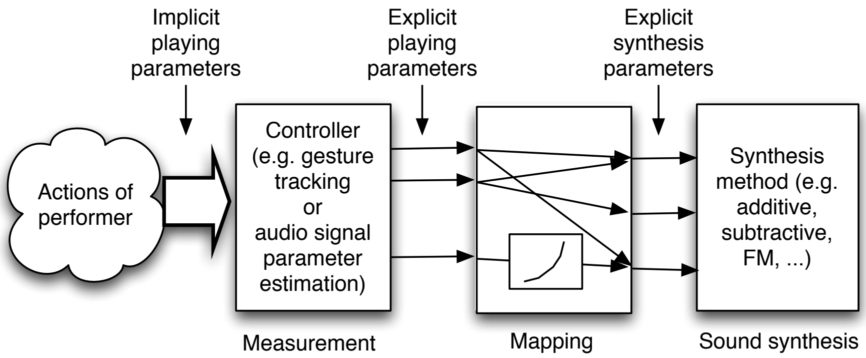

Figure 5.1: Basic principle of Parameter-Driven Sound Synthesis.

ods of mapping the player-related parameters to the synthesis parameters (p. 99). These methods are *one to one*, *one to many*, and *many to one*. A combination of these methods would result in the mapping method *many to many*.

Many synthesis methods have been developed in the past such as FM synthesis (Chowning, 1973), waveshaping synthesis (Brun, 1979) and physical modelling (Poli & Rochesso, 1998). The possibilities for generating different sounds are endless. However, the question of how to generate a sound and control it in real time in order to achieve a specific sound quality is a problem which must be solved by improving the areas described in figure 5.1.

In contrast to the conventional parameter-driven approach, the sound generation of traditional acoustic instruments is not based on the foundation of discrete parameters,[1] and a flute or a violin, for example, have no formal input for formalised parameters such as frequency or amplitude. A

[1]In order to give the reader an image of this understanding I will later refer here to discrete parameters as "pillars" on which the instrument is based.

formalised parameter is understood here as a parameter which is represented by a discrete number or by a stream of continuous discrete numbers. It is possible to understand the position of a left-hand finger on the fingerboard of a violin as a pitch input. The pitch of a violin, however, can be influenced by other factors such as finger pressure when stopping the string, pressure applied by the bow (overbowing can influence pitch significantly), bow speed, and by turning the tuning pegs.

Likewise, with a flute, it is possible to understand the discrete fingering patterns and a vector field measuring the air stream as a parameter set for an input. A performer, however, who wants to generate a specific sound by approximately closing keyholes (leaving the holes slightly open) will not be able to do this with discrete fingerings (closed *or* open keys). Similarly, a performer using a parameter on the mouthpiece which is not covered by the vector field will not be able to use this missing parameter to shape the sound.

As mentioned in section 2.3, a fixed set of explicit input parameters is understood here to be part of the *model* of a traditional acoustic instrument. The instrument is not seen here to be constituted by its abstraction, and thus its model. The model is not a part of the instrument per se, but is rather projected from outside onto the instrument.

In view of the differences between digital instruments and traditional acoustic instruments, one might ask how the parameter-driven instrumental models that are predominantly used in sound synthesis developed. It is possible that early approaches to sound synthesis have had a significant impact on the way sound synthesis is understood and conceived today. Thus, they will be examined in the following section.

5.1.1 Historical Foundations

The engineer Max Mathews played an important role in establishing the basic framework with which we are dealing in digital sound. Looking at the evolution of computer music, it is obvious that he has crucially influenced conceptions of computer-based sound systems. In the history of digital sound synthesis we find the first attempts in the late 50s. At the Bell Telephone Laboratories in Murray Hill, New Jersey, in 1957, Mathews introduced the first program for a digital computer that allowed the synthesis of sound. The program was called *MUSIC I* (Chadabe, 1997, p. 108). Mathews, born in 1926, studied electrical engineering and joined the acoustic research team at AT&T Bell Laboratories in 1955. In order to conduct listening tests, Mathews had developed a converter to digitise sound in the computer and to get it back out again. Mathews concluded (p. 108):

> It was immediately apparent that once we could get sound out of a computer, we could write programs to play music on the computer. That interested me a great deal. The computer was an unlimited instrument, and every sound that could be heard could be made this way. And the other thing was that I liked music. I had played the violin for a long time.

MUSIC I, the so-called *music compiler*, was written in assembler code and ran on the first mass-produced computer with floating point arithmetic hardware, the IBM 704 (Crab, 2005, para. 1). The programme had "one voice, one waveform, a triangular wave, no attack, no decay, and the only expressive parameters you could control were pitch, loudness, and duration" (Max Mathews in Chadabe (1997, p. 108)). Despite its limitations, this programme was a break-through in digital sound synthesis and it had a number of successors, summarised under the name *MUSIC-N*.

In the following years Mathews developed *MUSIC II* through *MUSIC V* (1968). With *MUSIC III* (1959) Mathews introduced the concept of unit generators: modules that can be connected and combined to generate sound. The collection of MUSIC V unit generators included oscillators, filters, adders, multipliers, random number generators and envelope generators (J. O. Smith, 1991, para. 2). These modules had parameter and audio-signal inputs, produced audio signals and were similar to those used in analogue synthesisers at that time. Figure 5.2 presents an example of connected MUSIC V unit generators. It should be mentioned that MUSIC V was not a visual programming language like MaxMSP. Thus figure 5.2 shows a (paper) graphical representation, but not the MUSIC V code itself.

With the concept of unit generators which could be combined to build a so-called *instrument*, came the concept of the *score* (Roads, 1996, p. 787). The *score file* specified features such as notes, duration, pitches, amplitudes and other parameters. Using this information the *instruments* (consisting of unit generators) could be played. Csound (2010), a MUSIC-N successor implemented by Barry Vercoe at MIT using the programming language *C*, is based on the same principle. Csound was introduced in 1986. An orchestra file defines the synthesis engine by a set of instruments, with each based on a

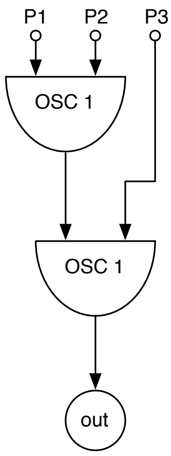

Figure 5.2: Example of unit generators of MUSIC V.

combination of unit generators. A score file consisting of discrete parameter data, including changes to data over time, drives the synthesis engine.

Coming back to the questions (posed in section 5.1, p. 116) of how the instrumental models used in sound synthesis developed, and why sound synthesis in nearly all of its manifestations is built on the method of Parameter-Driven Sound Synthesis (PDSS), I believe that the answer is threefold: firstly, Mathews' PDSS-approach had a great influence on the field of computer music; secondly, his approach was very successful commercially in areas such as keyboard synthesisers, sound modules for computers and software synthesisers; and thirdly, his approach parallels a major method in computer science in which objects are understood as units of functions generating output by computing explicit input parameters.

The assumptions and abstraction used by Max Mathews are viewed here as crucial because his early inventions, with a synthesis engine driven by discrete parameters, are still largely unchallenged. This conviction can be underlined by the fact that, in laying the foundations of digital sound synthesis, Max Mathews has come to be seen often as the "father of computer music" (Maramax, 2010, Biography, para. 4) or even the "great grandfather of techno" (Biography, para. 4). Since Mathews' approaches have been extremely successful and have been used in a huge number of synthesisers and computer audio systems, one might reasonably hypothesise many engineers have internalised basic considerations drawn from the view of Mathews.

One might argue with regard to the origins of this approach that the concept of unit generator based synthesis functions (i.e. the so-called "instruments") driven by a number of parameters was adapted from analogue synthesisers. However, Jean-Claude Risset (1984, p. 444) writes about the unit generators of MUSIC III as follows:

> MUSIC III was already developed in 1959, several years before Moog and Buchla's synthesisers appeared; it is possible to use them [the unit generators of MUSIC III] in different combinations, and to create additive synthesis (generating a sound from the sum of several components) or subtractive synthesis (generating a sound by filtering and modifying a complex sound or a noise).[2] [Translation by Cornelius Poepel]

In asking where the abstraction, the models and thus the formalisation for the conventional parameter-driven approach originate, one might reasonably

[2]Music 3 wurde bereits 1959 entwickelt, mehrere Jahre vor dem Auftauchen des Synthesizers von Moog und Buchla; man kann sie [die Unit Generatoren von MUSIC III] auf verschiedene Arten verwenden und auch additive Synthesen (einen Klang als Summe einfacher einzelner Komponenten bilden) oder subtraktive Klangerzeugung (einen Klang bilden, indem man einen komplexen Klang oder ein Geräusch filtert und modifiziert) herstellen.

argue that MUSIC-N is the model on which all of these are based. As mentioned earlier (see p. 117), Mathews was convinced that any sound could be generated by programming a computer. A meta-model of instruments built from unit generators with discrete parameter inputs seemed for him to be a starting point in developing a synthesis of any sound that could be heard.

On a fundamental level, sound-programming environments such as MaxMSP,[3] Pd (Pure Data, a programming environment similar to MaxMSP), SuperCollider (a text-based, object-oriented music programming environment), JSyn, Common Lisp Music and ChucK can largely be subsumed under the MUSIC-N approach. Although the score and orchestra are not necessarily strictly discrete as they are in the original MUSIC-N paradigm, these programs can be seen as derived from the MUSIC-N model inasmuch as they are all based on unit generators that run at the sampling rate of the audio signal (comparable to the orchestra), and control level functions (comparable to the data streams coming from the score) that operate the inputs of the functions at a lower control rate (with the exception of ChucK which has sampling frequency as the control rate).[4]

One might ask the question of how physical modelling sound synthesis is related to the MUSIC-N approach. Typically, physical models offer a specific set of discrete parameter inputs and an audio output. Given such a case, a

[3] It is called “Max” in honour of Max Mathews. MSP does not stand for anything official but it can be considered to have meanings such as “Max Signal Processing or Miller S. Puckette” (Cycling74, 2010a, Why the names Max and MSP?, para. 1).

[4] While SuperCollider is syntactically closer to the programming language *Smalltalk* (Lount, 2004) than to MUSIC-N, SuperCollider is viewed here as a derivative because of its unit generator approach and the distinction between audio rate and control rate.

physical modelling algorithm is viewed here as a unit generator object that can be driven from a formal input. If the audio signal of an instrument is used, and is fed into a physical modelling algorithm with the goal of using the implicit playing parameters to modify the sound, it is seen to conform more closely to the non-formal input of acoustic instruments, because the audio signal does not formally define which input parameters a performer is able to put into the audio signal. If the audio signal and a specific set of discrete parameter inputs are used, the model can be seen to fall in between the MUSIC-N archetype and the non-formal input of an acoustic instrument.

Up to Csound, music compilers are non-real-time systems. It is worth considering whether the basic abstractions of real-time systems are different. As early computers were slow, early MUSIC-N systems were not able to perform real-time sound synthesis. For this reason Max Mathews invented the GROOVE (Generated Real-time Output Operations on Voltage-controlled Equipment) in 1970, the "first fully developed hybrid system for music synthesis, utilising a Honeywell DDP-224 computer with a simple cathode ray tube display, disk and tape storage devices" (Crab, 2005, The GROOVE System, para. 1). The sounds were generated by a voltage-controlled analogue synthesiser, while the control was provided by a digital computer. The user interface offered a traditional qwerty keyboard, a 24-note "piano" keyboard, four rotary knobs and a three-dimensional rotary joystick.[5]

Mathews' plan was to build a system by which a pre-composed parameter-score and the performer, generating discrete input data with the knobs and

[5]This corresponds approximately, except for the mouse, to the kinds of interface used by many contemporary laptop musicians.

joysticks, would drive the synthesis engine. Describing the performer as a conductor, Mathews explains his idea of performance as follows:

> The computer performer should not attempt to define the entire sound in real time. Instead the computer should retain a score and the performer should influence the way in which the score is played..... [*sic*] the mode of conducting consist [*sic*] of turning knobs and pressing keys rather than waving a stick, but this is a minor detail....... [*sic*] The programme [*sic*] is basically a system for creating storing, retrieving and editing functions of time. It allows the composition of time functions by turning knobs and pressing keys in real time: it stores the functions on the disk file, it retrieves the stored functions (the score), combines them with the input functions (the conductor) in order to generate control functions which drive the analogue synthesiser and it provides for facile editing of functions via control of the programme [*sic*] time... (Crab, 2005, The GROOVE System, para. 3).

Interestingly, the difference between a conductor's action of "waving a stick" and the action of "turning knobs and pressing keys" is seen here as a "minor detail". This does not concur with the current research's understanding of musical conducting gestures. Marrin Nakra (2001) has conducted research into conducting gestures and the development of conducting systems, and writes: "And by continuing to improve upon gestural control systems, we can ensure that performers will not have to change their established techniques – just perform naturally and let the computer systems follow them" (Introduction, para. 2). Thus, I conclude that she considers it a desirable goal to allow the conductor to use trained gestures, which is not the case if knobs are turned. In addition, one might ask to what extent defining the act of performing as an act in which a performer provides additional parameter input to a given stream of parameters (the composition) will satisfy a performer who is used to traditional performance practices.

5.1.2 Formalisation Strategies

With regard to formalisation, one strategy is to understand musical objects and operations as elements that *consist* of formal structures. The known formal structures can be used as a basis for formalisation. If the available knowledge of, for example, the physical structures of an object is not sufficient, research must then be done to achieve a deeper understanding of the formal structure. An example of this research is the investigation of the operation "musical gesture". Following this strategy, once the formal structure is known, the object or operation *itself* can be used in computer applications by making use of its formal structures for modelling or for building interfaces.

Another strategy is to understand musical objects and operations as *describable* using formal structures. The formal structures, however, are seen as an abstraction of the objects and operations, not as the inherent nature of the objects and operations *themselves*. The abstractions are used to build models of the originals. As with the first strategy, the formal structures can be used as a basis for formalisation. In the resulting computer application, however, the output is not understood as the object or operation *itself*, but as a variant of models that *may or may not* be perceived as the objects or operations themselves. In fact they *are not* the original objects and operations because they are built from models.

Although Mathews claimed that the computer was an "unlimited instrument" and could make "every sound that could be heard" (cited on p. 117) by using computer programs to play music, it cannot be stated with certainty that he assumed the MUSIC-N approach, expanded with an input device like the GROOVE system, would itself be successful in accomplishing this. For the purposes of providing an intellectual point of departure for the work presented in this thesis however, it is useful to consider what assumptions might have been made in viewing such an approach as successful and complete, and what shortcomings might be revealed by examining these hypothetical assumptions more closely. These assumptions are as follows:

- Every sound one can hear can be generated using a synthesis engine and a limited number of discrete, independent and separate parameter inputs.

- A synthesis engine capable of generating every possible sound can be constructed using a finite number of unit generators.

- A finite number of discrete parameters form a score. A composition is defined by such a score. An additional number of discrete parameters coming from the musician constitute, together with the score parameters, all the information necessary for the performance of a composition. The performance of a composition is defined by such a score and such a set of performance parameters. In the case of an improvisation, all parameters come from the musician (because there is no score).

- To perform a composition it is necessary to read discrete parameter data from the score and the performer, in order to drive the combination of unit generators.

- With this system's setting, a composition or an improvisation can be performed in which all the sounds that one can, in principle, hear can be generated with the full spectrum of quality we know from the history of music, as well as with new, additional and improved sounds or timbres and qualities of sound.

Following from this, a musical instrument is seen as a device with formal inputs that must be operated according to the discrete parameters of a composition. A performer playing a musical instrument is thus seen merely as a parameter data provider who adds parameters to the already fixed parameters of the composition. In view of these assumptions I conclude that the formalisation strategy found here falls into the first of the two categories described above. It is important to mention here that this strategy is not based on facts about music or musical operations. Mathews did not know, for example, whether every sound one can hear can in fact be generated using a synthesis engine.

Computer applications implement theories (Degele, 2000, p. 139). And the MUSIC-N approach implements a specific abstraction and thus theory of music. It seems probable that the formalisation strategies of Mathews' MUSIC-N implementations had a large influence on the users working with MUSIC-N variants. As a significant part of musical education, especially in electronic or computer music, uses MUSIC-N successors, one can expect that Mathews' understanding of objects and operations in music has influenced contemporary understandings of music to a large extent. It is my contention that sound synthesis is conceived and constructed in a way which is prescribed by the MUSIC-N approach, and that solutions to problems in sound synthesis are thus sought primarily within the framework proposed by the MUSIC-N approach.

Although the MUSIC-N based approach of Parameter-Driven Sound Synthesis was very successful, shortcomings were noticeable with time. In the early stages of the approach, a problem was considered to be the computational power necessary to generate the sounds one was seeking. As Puckette (1991) mentions, the focus has shifted to the problem of how to control a real-time system:

> The sample generation problem has historically been considered 'hard' simply because of its stringent computational requirements. [...] But the control problem, that of making the signal network respond in an instrument-like way to live human control, is not made appreciably easier by the availability of faster and faster hardware. Today, the challenge for a signal processing network editor is to open itself up to a wide range of control possibilities (pp. 68-69).

The sound-control problem was increasingly seen to be of importance. This issue will be considered more closely in the following section.

5.1.3 Problems and Common Improvements

A typical description of the shortcomings of current musical interfaces for parameter-driven synthesis systems is given by Marrin (2000, p. 15):

> First of all, many current interfaces do not sample their input data fast enough or with enough degrees of freedom to match the speed and complexity of human movement. Secondly, the nature of the sensing environment often inappropriately constrains the range and style of movement that the performer can make. Thirdly, the software that maps the inputs to musical outputs is not powerful enough to respond appropriately to the structure, quality, and character of the gestures.

To overcome such problems one would need faster sampling rates on the input side, more degrees of freedom, sensing systems appropriate to the range and style of the performer's movements, and more powerful mapping systems. One would need to establish, define and formalise the essential degrees of freedom, performer's movements and mapping systems. In parallel to what Mathews did when improving MUSIC I, one can find here the basic principle already mentioned in section 4.7.7: improvement is achieved by defining and formalising essential parameters and functions, and adding these to the system.[6]

In the field of interface research three main areas of focus can be identified when it comes to the question of how existing computer-based musical instruments can be improved. These areas are shown in figure 5.3, which develops the PDSS model shown in figure 5.1 (see p. 115).

1. the measurement of the player's actions,

[6] While Teresa Marrin does not explicitly state that she uses the method to define and formalise essentials, one may see in her description of the approach to improving available conducting systems that she indeed defines the essential performance parameters, builds a measurement system and uses functions to map the measured data to a synthesis unit (Marrin, 2000, p. 30).

2. the mapping,

3. the sound synthesis itself.

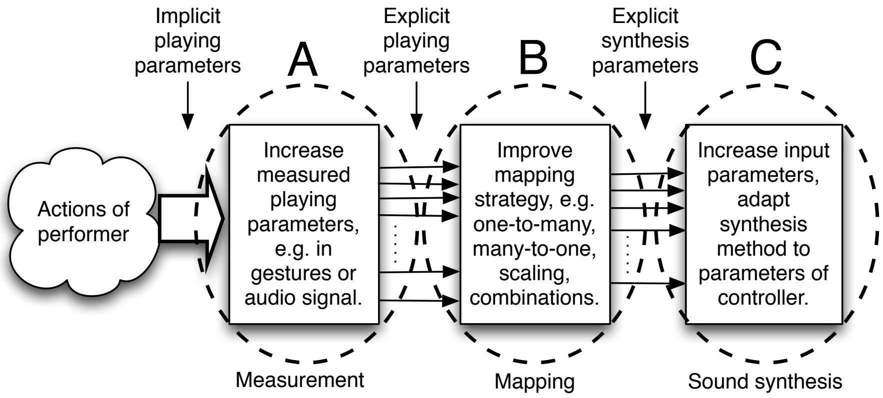

Figure 5.3: Common areas of improvement.

Examples of research related to area A have been mentioned in section 4.1 on p. 45 and to area B on p. 46. With regard to area C two methods can be mentioned here: Physical Modelling and Cluster Weighted Modelling (CWM). Both are built to be driven by formal inputs correlating to the inputs one can describe when looking at musical instruments from a physical perspective. Examples of physical models in relation to stringed instruments can be found in Woodhouse (1992) and Serafin and Young (2003). As described in section 4.4.4, Jehan and Schoner (2001) use the method of CWM in a timbre synthesiser that can be played with a violin interface.

One question that remains is that of whether the view of musical instruments presented in section 5.1.1 is the only possible one, and thus, whether these three areas of work are necessarily the only ones worthy of attention. As mentioned earlier, I assume that the developer of an instrument is led to a great extent by her idea or view of an instrument. Levitin, McAdams, and Adams (2002, p. 171) parallel this assumption with the statement that a new way of conceiving musical events and tones can lead to new construction principles for musical instruments. If the instrument's nature is conceived in abstract and formal terms, it will be built and improved according to those terms.

Another question is that of whether the conventional approach to sound synthesis tends—as a result of its construction principle—to a specific set of problems. In order to establish whether a specific set of problems can be found, and to explain why it might make sense to work in areas other than the three described above, I would like to discuss some problems associated with the conventional parameter-driven approach in more detail in the following sections.

5.2 Performer and Instrument

After looking at the history of sound synthesis and the function of—and improvements to—the conventional approach to sound synthesis, I would now like to focus on the performer and the question of how instruments are built in relation to the actions and expression that the performer puts or might want to put into the instrument.

5.2.1 Performers' Actions

As mentioned earlier, the research field of human-computer interaction in music presently exhibits a growing body of work which is concerned with establishing which of the actions of a performer are crucially related to the sound one hears. In contrast to the early approaches of Mathews,[7] the difference between playing a musical instrument with knobs and faders and the gestures of a traditional performer has for several years now been seen not as a minor issue but rather as a major concern (Cadoz, 1988, p. 1).

This is connected with the belief that part of the quality of the sound (for example its expressivity) cannot be generated in the synthesis unit, but is generated by the performer through his actions. While it is possible for a conductor to tell the musicians to play more expressively, there is no synthesis technique commonly known which allows a single sound to be made more or less expressive.[8] The expressivity is generated by the performer. Thus, one might ask: how is this quality generated? How can one measure the quality and capture it on the computer?

Besides questions relating to sound synthesis, researchers such as Müller and Mazzola (2003) address the question of how musicians generate expressivity in an interpretation of composed classical music. Their research con-

[7] cf. the citation of Mathews talking about the GROOVE system in section 5.1.1.

[8] Interestingly, the General MIDI standard defines the parameter *Expression* on controller no. 11. I speculate that there was an awareness of the need to be expressive and the idea of having a slider with which one could simply adjust the expressivity of sounds. This adjustable parameter, however, did not solve the problem of performing expressive music.

cerns which musical parameters are used to generate expressivity. Hiraga, Bresin, Hirata, and Katayose (2004) conduct empirical studies in which listeners are offered different "interpretations" of written music performed by computer applications and are asked to evaluate which of the performances is the best one. The research of Wanderley, Vines, Middleton, McKay, and Hatch (2005) focuses on clarinetists' gestures and their meaning for the production of sound.[9] O'Modhrain (2000, p. 89) provides evidence of the positive effect which interfaces with haptic feedback have on the actions of the performer in instrument playing.

Interestingly, these scientific approaches describe the actions of the performer using a physically based model of the performer and measurement results. One might ask whether sources other than physically based abstractions might also be worth consideration. If one were to reduce the actions of the performer to physical actions, or physically measurable actions, and allow the performer to be viewed only through the spectacles of physically describable actions, a requirements document based on such a view would be similarly pre-specified (and thus reduced) to that provided by Schoner (see p. 99).

An additional source for describing the actions of the performer can be found in the field of instrumental pedagogy. It is the task of instrumental teachers to tell a pupil what actions to make in order to play the instrument. Such descriptions may be seen as less reliable because they are not based

[9] An ongoing project analysing the gestures of clarinet performers can be found at (Wanderley, 2007).

on existing scientific models of the performer and on physical measurement, but rather on phenomena which the teacher knows, perceives and wants to teach to students. In order to create a meaningful requirements document for a synthesiser instrument, however, it is important to meet the needs of the future users, even if those needs are derived from phenomenological descriptions of players' actions (which are true and meaningful for the player, otherwise they would not be used).[10]

5.2.2 Gesture

Since the term gesture is used widely it will be looked at more closely in this section. Claude Cadoz was one of the first researchers to investigate the role of musical gesture in music informatics systems. In his paper "Instrumental Gesture and Musical Composition" (Cadoz, 1988), he mentions that the majority of sound synthesis systems put the musician far from the usual performance practice. Investigating the role of the instrumental gesture he concludes that it is important that gestural actions include "deduced functional specificities [*sic*] that should be included in the conception of a musical informatic tool" (p. 1).

Together with Marcelo Wanderley, Cadoz describes the term *gesture* in a further article by presenting a collection of 17 definitions of gesture (Cadoz & Wanderley, 2000, p. 28). Since these definitions provide a diverse understanding of gesture, the authors conclude that the definitions are context

[10] I will come back to descriptions used in music pedagogy in section 5.3.1.

dependent and thus are “valid in their own right” (p. 33). A common feature is the “human physical behavior [*sic*]” that is directly or indirectly referenced in the definitions. The authors analyse instrumental gestures in two ways: a phenomenological and a functional approach. The phenomenological approach is understood in the paper as a descriptive analysis, describing gestures in “movement speed”, “size of the space where the gesture is done” and the “movement decomposition regarding its frequential content” (p. 33). As the context and meaning of gestures plays a crucial role in the definition of what a gesture is and what it is not, the authors avoid proposing their own definition of gesture.

Assuming that one wishes to build a device for capturing gestures, a definition and a model of gesture are necessary to create the specification for the device. If one knows that this definition is context dependent, it becomes clear that the capturing methodology implemented in the device will force the user to stay within the specific context in which the system works properly. If the player does not stay in the context or the range within which the definition and the model of gesture is valid, the measurement device may create output data which the user perceives as mistakes.

As well as the confusion that the term “gesture” has caused, one must consider that gesture describes only one subset of the actions a performer can make when playing an instrument. The reason for this is that not all actions made by a player in the performance of music are necessarily visually identifiable body movements. In my experience, muscle relaxation can be an important action while playing, but no body movement will necessarily be visible. Similarly, imagining a specific tone before it is played can be an important action when performing.

In order to use a term which is open in principle to anything a performer may want to do while performing with an instrument, I prefer to use the term "performer's actions" (defined on p. 28). Since this research is concerned with sound synthesis and the playability of synthesised sounds, I will focus on those actions that are relevant for the creation of sound in a performance.

5.2.3 Measurement of Performers' Actions

So far, it has been argued that the design of an interface and the formalisation and measurement of a performer's actions may cause problems. The question remains open—with regard to specific implementations—as to what these problems are precisely and how they might prove to be relevant to users. Since the Hypercello has been used at a very high artistic level, and its successors are very well documented, those successors will be taken here as examples to be considered in detail.

The actions selected with which to model the performer are reduced to actions which have a clear correlation to physical quantities. An action like "digging into the tone" would not have clear physical correlates and is not measured by the Hypercello. For each action to be measured, the physical correlates and the measurement must be defined and a method of measurement constructed. Presented below are actions, physical quantities, measurement and means of measurement according to my understanding of tracking technology as described by Schoner (2000b, p. 173).

- Action: stopping the string at a specific position. Physical quantity: position of finger on string. Measurement: finger position. Means of measurement: point of contact of string on a resistive strip; resistance is measured.

- Action: bow speed. Physical quantity: speed of bow. Measurement: lateral bow position. Means of measurement: electric field sensing; two oscillators on the bow with different frequencies (one at the frog, one at the tip), driving an antenna. The amplitude of received frequencies is measured, and from this data lateral bow position, bow speed and bow acceleration are calculated.

- Action: bow pressure. Physical quantity: downward force of bow hair on string. Measurement: downward force of hand on bow. Means of measurement: FSR (force-sensing-resistor) under index finger of bowing hand; resistance is measured.[11]

- Action: bow position. Physical quantity: distance of contact point of bow hair on string from bridge. Measurement: longitudinal position of bow stick. Means of measurement: oscillator with antenna on bow stick; antenna on bridge receives oscillator signal. The amplitude of the

[11] The Hyperbow (Young, 2002) differs in the measurement of downward force on the bow stick. Two-terminal foil strain gauges are mounted on the bow stick and their resistance is measured (p. 3).

received oscillator signal is measured, and from this data the distance of the bow from the bridge is calculated.

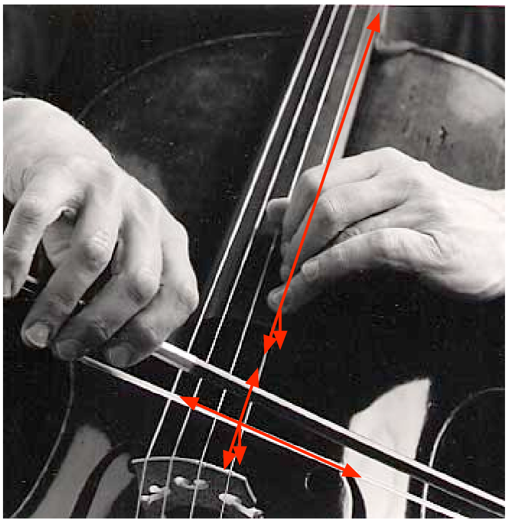

Figure 5.4: Abstraction of playing parameters.

The abstraction used to create this model is illustrated in figure 5.4 and the model is presented in figure 5.5. It should be mentioned here that the model is reductive because playing parameters such as bow angle to bridge or amount of hair are not measured. Furthermore, it is reductive because the definition of measurement does not only depend on the actions of the performer, but also on available methods of measurement (Schoner, 2000b, p. 173).

Since the goal in building the Hypercello was to “detect the player’s actions” (Paradiso & Gershenfeld, 1997, p. 73), I assume that the developers aimed to achieve a level of detail in the mapping that would allow a professional performer to make meaningful use of his actions in relation to the sound result. As mentioned in section 4.4.2, however, practice showed that

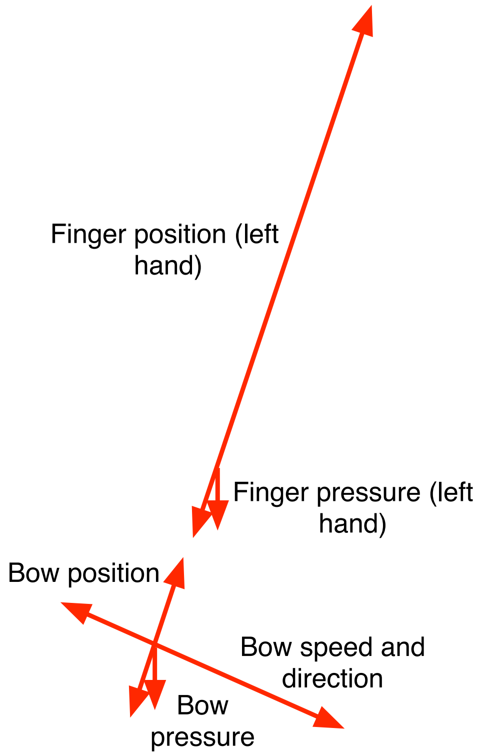

Figure 5.5: Model of playing parameters.

the expressivity of the player was not sufficiently represented in the sound result, and it was concluded that subtle actions were not captured (Levenson, 1994, p. 17).

Violinist Joshua Bell, who often plays the Hyperviolin and Hyperbow (both successors of the Hypercello), mentions his wish that the instrument be made “more organic” (Hermida, 2002, Interesting effects section, para. 6). This raises the question of what exactly these subtle actions or the organic feel are. Referring to the problem of gesture definition described in the last section, and its context-dependency, one can conclude that the definition of a “subtle action” will be context dependent too. Thus, building a measurement apparatus will always implement a definition of the object to be measured, and this definition will anticipate the framework in which a performer acts.

Since the definition of this framework is determined by the performer (and can be changed, to a certain extent, even while playing), the system developer will need a measurement device which is self-adapting to the framework in which a performer plays. If this is not the case, the measurement device will work properly only within the specific contexts to which the definitions refer. From a wider perspective one might say that such a system forces the player to subordinate her actions to the pragmatic features (explained in section 3.6) of the model implemented in the system.

Another question is that of how well the means of measurement correlate with the physical quantity to be measured. With regard to the parameter bow pressure and the physical quantity “downward force of bow hair on string", the method of measurement "FSR under index finger" can be problematic. It is conceivable, after all, that bow pressure might be applied while lifting the index finger and using other fingers to apply downward force to the string. Similarly, with regard to the parameter bow position (physical quantity: distance of contact point of bow hair on string from bridge), the method of measurement described above measures the stick of the bow, which means that it is possible to turn the bow stick closer to the antenna at the bridge or further away from it, while the contact point of hair remains the same.

In both examples, differences between the measurement results and the physical property to be measured may occur. If the performer wants to control the measurement results in order to create a specific sound, he has to focus on the measurement devices and their specific functions.[12]

Besides the known issues of latency and mistakes in measurement, another problem must be mentioned, which was already pointed out in section 3.7, namely that the process of measurement can influence the object being measured. An example of this can be found in audio signal's pitch tracking of bowed stringed instruments. In my experience, tracking mistakes and latency lead to a player's focus shifting. The focus is usually on the string, but then moves to the measuring system, because the player wants to generate specific pitches at a precise time. With Zeta MIDI string instruments some players

[12]Looking at the Hyperbow (Young, 2002, p. 3) and its method to measure downward force of the bow I assume that the possible mismatch between applied bow pressure and measurement result is reduced by the measurement of strains in the bow stick.

even stop playing because they can not accept the error rate, while others do not stop, but adapt to the specifics of the system.

I have also known players to say that they change their playing style in order to provide the acoustic material to make the pitch tracking system generate the correct pitch values. Thus, if one wishes to measure the actions of the performer, one has to accept that (in pitch tracking, for example) the process of measurement may change the object that is being measured. In part this parallels the dynamic interaction of the two systems Neumann was talking about when explaining the problem of formalisation in measurement (see section 3.7). One can conclude that the ability to formalise the measurement process with reasonable accurately is dependent on the player being measured and not simply on a "playerless" idea of the actions of a performer.

If a pitch tracker is perceived to "make mistakes" or "have latency", one may ask whether this understanding can be seen as appropriate. The reason for this is that the so-called "measurement mistakes" may result from an incomplete formalisation. In such a case they should therefore be referred to as "formalisation mistakes", because the cause of the perceived mistake does not lie in an algorithm not working properly, but in an incorrect formalisation.

In order to illustrate this more clearly I would like to refer to the example of a single note. For the player a note has a starting point and an end point. It has a specific pitch, volume and timbre and can be articulated in different ways. Viewed from a physical perspective, the note does not have a specific pitch from the beginning, but the pitch is established after the initial transient phase. Yet from the perspective of the musician the note has its pitch already from the beginning, with an articulation at the beginning of the note: when an "A" is notated in a score, the "A" is played immediately from the very appearance of the note, and not some milliseconds later.

The physical description of what happens and the musician's description are both abstractions, and there is not one "real one" and one "abstract one". It is the task of the formalisation to provide a formal structure that can in fact measure what it is intended to measure: the pitch of the note that a performer is playing and perceiving as the played pitch. If the pitch tracker jumps during the transient phase of a note one might say: "the pitch tracker made a mistake." This mistake is a mistake in relation to the pitch perceived by the performer. It is not a mistake in relation to the physical pitch in the transient phase, because the transient phase can be described as undefined in terms of physical pitch.

If the formalisation is based on the physical description of pitch, the algorithm of the pitch tracker did not make mistakes. The pitch tracker carried out the measurement as defined by the algorithm. The mistake lies in the formalisation. The formalisation was based on the physical abstraction of pitch, probably as a result of the conviction that the physical description was the real one, while the description of the musician was the wrong or insufficient one.

The difference between physical and perceptual parameters is also addressed by Jehan (2001, p. 24) who, citing Sethares (1998), writes that "physical attributes such as frequency and amplitude are kept distinct from perceptual correlates such as pitch and loudness [...]." The problem one is confronted with here is the non-congruence between perceived parameters and physical correlates. It is difficult to make a correct specification once it is clear that perceptible parameters are dependent on and intercorrelated with other parameters.

In the above-mentioned example of the Hypercello, a particular model of a musician was presented in which the performer had a certain range of physical playing parameters available. It was assumed by the developers that a new instrument with the qualities of a traditional one could be created if one had an interface to measure these parameters. With regard to the problematic issues described, one might argue that using a predefined set of playing parameters—and building a device capable of measuring the related set of physical actions—may cause problems in this context.

5.2.4 Openness, Transparency and Intimacy

In order to achieve good playability and offer the skilled performer a good level of potential for expression, a new instrument should offer sufficient openness to all the actions that a performer might want to carry out. Openness is understood here as a phenomenological property of a system: the ability of the instrument to respond—in the ideal scenario—to all actions relating to the intentions of the performer to one degree or another.

In my opinion a new computer-based instrument for use by trained players of traditional instruments should be open in several ways. It should be open to the specific actions of a performer, it should be open to being explored on an increasingly deeper level, and it should allow continual expansion and be able to differentiate and refine the actions of the performer in relation to the sound output.

In the field of computer music research, similar phenomenological qualities are also addressed, though other terms are used. I will focus on the terms "transparency" (Fels, Gadd, & Mulder, 2002, p. 109), and "control intimacy" (Moore, 1988, p. 21), since they are closely related to what I am seeking to describe with openness, playability and potential for musical expression.

Fels et al. (2002) consider that in order to be musically expressive an instrument should be transparent with regard to its input/output relation. They define the "transparency" of a musical instrument[13] as the ability of both the player and the audience to understand the mapping process of the input/output relation. Since the mapping process to which they refer is based on interfaces with a fixed set of discrete parameter outputs, two assumptions are linked with this understanding of transparency.

The first assumption is that an interface will be able to measure the essential input parameters, and the second is that the mapping process is the key element in solving the problem of communicating the expressive intentions of the player to the public. Improving an instrument in input parameters and mapping may make it more transparent in relation to the player while

[13] They refer to "both traditional instruments and technology-driven interfaces" (p. 109).

not necessarily to the public. The performer may have been working with the instrument for a long time and have understood and internalised the input/output relation. The understanding of the input/output relation may be very different for the public, however, as the public does not have the player's insight to the instrument. In other words, the transparency of an instrument is not a factor that can be seen merely in relation to the instrument. It is dependent on the instrument, but also on the player's and public's predisposition.

Moore (1988) coins the term "control intimacy". He defines this as "the match between the variety of musically desirable sounds produced and the psychophysiological capabilities of a practised performer" (p. 21). He regards the human voice as the instrument with the greatest control intimacy, because of the musically desirable sounds producible. Other instruments with a large control intimacy are, in his view, the flute, the sitar and the violin. The reasons given for this are "the wide range of affective quality in the musical sound" which can be translated from "the microgestural movements of the performer's body" (p. 22).

With the aim of achieving an "Intimate Musical Control of Computers", Wessel and Wright (2002, p. 11) propose what I would consider a low-threshold—high-ceiling approach to musical interfaces. They argue that an interface should allow for ease of use coupled with a "long-term potential for virtuosity" (p. 20). Besides this, a minimal and low variance latency is found to be crucial. In addition, event-based and signal-based (continuous control) strategies for the connection between gesture and musical result which are simple and easily understandable, are mentioned as a basis for intimacy of control in sound synthesis.

In order to improve an instrument's opportunity to allow a player to embody a computer-based musical instrument, Fels (2004) focuses the design on mapping for intimacy. This idea can be seen as closely related to the above-mentioned idea of transparency. Proper mapping designed for intimacy will "allow intent and expression to flow through the player to the sound" (p. 672). A precondition is that a sufficient degree of freedom is offered by the interface, so as to result in multiple streams of discrete parameters. Referring to the above definition of Moore's "control intimacy", Fels proposes a generalisation of intimacy that may be seen as helpful for all kinds of computer interfaces. As a key element of intimacy he mentions the feature of embodiment. A high degree of intimacy has been achieved when the user and the object are perceived as a unity, for example if the device is seen as "embodied within herself, i.e., an extension of herself" or the user perceives "herself as an extension of the object" (p. 672).

The reasons given for such requirements are in line with the current research focus on playability and potential for musical expression. In summary, there are three main strategies for achieving good quality in terms of (control) intimacy and transparency. Firstly, sufficient degree of freedom, minimal variance and low latency in the input are required.[14] Secondly, the method and comprehensibility of the mapping is crucial, and thirdly embodiment and a large potential for extending virtuosity in performance is important.

[14] Below 10ms, according to Wessel and Wright (2002, p. 13).

5.3 String Players' Requirements

System specifications that are based primarily on a mechanistic view of string playing—that is to say on a mechanical model of string playing, where a sufficient number of essential actions can be defined in advance—have been discussed in section 4.7.4. One might ask whether this concept of string playing brings with it problems which could negatively influence the further development of a computer-based string instrument. In my experience as a string player, a mechanical model and the mechanical understanding of string playing which follows from such a model can be a problem for a student learning the instrument. The violinist Yehudi Menuhin writes in the preface to Biesenbender (1992):

> I agree fully [...] about the deadly effect of teaching performance from fixed visual and deliberately frozen, formulated fractional details of position when what must needs to be 'learnt' is the *flowing* co-ordination, which incorporates a myriad of elements, themselves in constant flux of speed and proportion (p. 5).

This description of violin playing highlights the question of whether a system specification which allows a player to perform only inside a predefined framework of limited and fractional elements will cause problems because it is not open to all of the varying elements a player may want to co-ordinate. In the context of Menuhin's understanding of violin playing, I assume that the vBow (section 4.3.6) or the SuperPolm (section 4.3.2) might be problematic because the parameters measured in these systems span a reduced space in which the flowing co-ordination to which Menuhin refers can only take place in a reduced form.

Another understanding of musical performance can also be used to address the requirements of a musical instrument. As pointed out in previous sections, most of the developments described are built on the assumption that a musical instrument is a good instrument if it does exactly what the player wants it to do. In addition, an ideal instrument does not only realise the intentions of the performer, but amplifies these.

The piano teacher Neuhaus (1967, p. 3), however, describes the instrument not as a tool doing exactly—or even amplifying—what the musician wants it to do, but instead as a resistance against the musical idea. In accordance with the principle of dialectics, a musical idea is seen as the thesis, an instrument as the antithesis, and the performed music as the synthesis. Consequently, a good instrument offers a resistance against the musical idea, thereby helping the musician to achieve a sound result which is different from that desired, but offers a convincing musical quality (Poepel, 1999, p. 15).

A valid requirements document could be generated if there were one stringent description and understanding of playing a stringed instrument. However, we are confronted with different descriptions that do not necessarily fit together. It is a well known fact that there are different schools of violin playing, and it is interesting to ask what such variations mean for the development of a requirements document. In order to arrive at a deeper understanding of such differences, the following sections will look at two approaches to playing a stringed instrument. On the basis of these perspectives, I will then draw conclusions for the requirements of computer-based stringed instruments.

5.3.1 String Players' Actions

The following areas can be addressed when asking what string players do when they play (Poepel, 2008b, p. 362):

1. Individual players and their actions: one can simply ask them, "What are you doing when you play?"
2. Instrumental teaching: this is important, because it is here that the questions of how to play the instrument or what to do when playing the instrument are addressed.
3. Musicians rehearsing together: one can study what musicians achieve together and how musicians explain to one another what to do.
4. Physical analysis: one can measure the playing parameters of individual players.

Since I am concerned here with requirements and not with the abstraction of requirements (as the basis for the specifications), I will focus on the performers' descriptions of things which happen when musicians perform, but not on the physical description of players' actions. To make this point clear and to repeat what has been mentioned earlier: this thesis is written with a strong emphasis on my perspective as a performer (although naturally informed by other viewpoints). As I, and I assume many other performers, do not see the world as a physical world per se (one in which no kinds of things other than physical things exist), I believe that the world's physical descriptions are already an abstraction of it. Physical descriptions can be used to move from the requirements level to the level of specification but are not equated with the original thing itself.

Accordingly, string playing is not seen simply as a physical interaction, in which one assumes that a player is using continuous control parameters and one must discover which parameters these are and how one could ideally measure and map them to the discrete inputs of a parameter-driven synthesis engine. What I wish to do here is to investigate how string players understand their own actions, in order to be able to define requirements which meet the needs of the target users as closely as possible.

In order to meet the time constraints of the present research project, I have not asked individual players, as already done by Paine (2007, pp. 70-71), or analysed discussions in rehearsals. I have concentrated instead on instrumental teaching in asking about the actions of string players, and on performers' experiences. Material from two perspectives on string teaching will be presented. The first perspective is a rather rationalistic approach to violin playing taught by Ivan Galamian. The second one can be described as an organic method of violin playing favoured by the violinist Volker Biesenbender and following a tradition of violin playing which includes Yehudi Menuhin. I will not attempt to describe all their views on performance and the actions of the performer, but will provide insight allowing the reader to understand the differences between these approaches.

The work of the violin teacher Galamian (1962) has been used as a reference point in the development of computer-based stringed instruments (Young, 2001, p. 16; Nichols, 2003, p. 90). While Galamian regards the biomechanical view as a helpful approach to explain and teaching violin playing, he criticises the “overemphasis on the purely physical and mechanical aspects of violin technique” of past violin schools (Galamian, 1962, p. 2). He sees the key to violin playing in a mind-controlled performance with adequate “*relationships of minds and muscles* [Galamian’s italics]” (p. 2), incorporating a precise and direct response of the muscles to the mind. Thus, violin playing must begin in the mind, with an idea of sound coupled to an internalised set of corresponding physical actions.

He mentions “vowels” and “consonants” as the fundamental acoustical elements in violin playing (p. 9), a usage which can be seen as parallel to human speech. In order to give character and contour to the sound, violin tones must “have a certain admixture of percussive or accentuated elements” (p. 9). For articulation the consonant-like accentuated or percussive sounds are generated with the left or right hand. The consonant-vowel relation of a particular musical character or expression is not linked to fixed proportions. It depends greatly on the context of the performance, which may be given, for example, by the concert space. Galamian explains the importance of adjusting the relation by stating: “[the] vowel-consonant balance is not the same for the concert hall as for the studio” (p. 10).

With regard to the right arm, the hand and bow are seen as a functional unit that must be understood as a system of springs reacting in “much the same way as mechanical springs" (p. 44). Thus, the actions of the performer must always be aimed first and foremost at keeping this system in a flexible, resilient and springy state in order to avoid a tone which is "hard and ugly, the bowing clumsy and uncontrolled." Actions of the right-hand's movement include the vertical and horizontal movement of the fingers and the horizontal turning motion. The task of the wrist is to control vertical and horizontal movements of the hand. The forearm performs an open-close motion as well as a forearm rotation and similarly the upper arm executes a vertical and horizontal motion. Galamian sees the straight bow stroke at right angles over the length of the string as a basis for the creation of a good tone production. In order to play the instrument, the beginner must learn to keep the bow straight.

Three factors are crucial to tone production: bow speed, bow pressure and sounding point (bow-bridge distance). In order to create a proper tone, these three factors must be put into proper relation. Once these basic tone production requirements have been learnt, the player can go on to learn "various styles of tone production" (p. 62) in order to shape character and style of tone, for example. Concerning the three factors of bow speed, pressure and sounding point, Galamian writes:

> The understanding of this relationship is of great importance, but it must not lead to the false belief that there is only one combination of the three factors possible in any given instance. (p. 62)

This means that the proper relationship of the three factors forms a surrounding framework which the performer can vary in order to generate sound, timbre and character, and to achieve expressivity, but does not mean that a specific sound, timbre character or expression is correlated with a specific set of the three parameters. It is conceivable that a specific set of the three parameters would correlate to a specific desired tone if the three parameters were the only ones responsible for sound, timbre or character. However, since there are more than these three that are responsible for sound (for example the amount of bow hair applied to the string, or the angle of the bow relative to the bridge), the three parameters are understood to be of significant importance to the sound result, though not completely defining it.

Galamian's approach may be summarised as one which allows the actions of a performer to be separated into subsets and units which must all work properly in order to achieve a good result. If one unit is not working properly, the player must "repair" or train it. The biomechanics of the body may be seen as a preliminary form of physical description. In contrast, sound production goals such as the "character" or "style" of a tone do not offer a direct physical relation.

As cited in the section above, violinist and teacher Yehudi Menuhin describes violin playing as a flowing co-ordination of elements which are, in themselves, in a flux of speed and proportion (Biesenbender, 1992, p. 5). Similarly to Galamian, he argues that no single point in "hand, arms, neck, shoulders can ever be allowed to stiffen" (p. 5). He understands violin playing as an act of balancing with no relation to a single fixed position. "In other words, no part of the body (or mind) may be allowed to 'set', to form an obstacle to the flow of musical impulse" (p. 6). He regards it as given that in violin playing one should be "handling a living body" and compares playing to "love-making" (p. 6). For him, performing music requires a "*true* interpretation—true to the spirit behind the notes" (p. 6).

Biesenbender (1992) makes the point that violin playing may be understood and taught very differently from teacher to teacher. Moreover, different cultures may demonstrate very different approaches to musical performance (pp. 15-17). In order to achieve a better understanding of diversities in approaches to violin playing, he analyses the history of violin teaching. As mentioned in section 4.7.4, he concludes that performance instructions and models of movements, gestures and playing are dependent on the view and mentality of a specific time and world view from which the violin schools arise. Famous violin schools such as those of Leopold Mozart[15] and Guiseppe Tartini[16] were products of the Enlightenment, when one of the most widely read books amongst cultured citizens was "L'Homme Machine" (1747) by Julien Lammetrie.[17]

As Biesenbender explains, it was seen as an ideal at his time to eliminate naturally occurring elements and to replace them by regulated and controlled systems. Accordingly, violin teachers did the same and began to reduce violin playing to a set of basic elements within a mechanical system designed to en-

[15] 1719-1787.
[16] 1692-1770.
[17] 1709-1750.

able one to explain and reproduce all possible musical objects. Biesenbender concludes that such an approach does not necessarily describe what violin playing is about, but rather what methods were used when people dealt with challenges at that time.

Along with Yehudi Menuhin, Biesenbender proposes an ecological approach to violin playing as a result of the experience that many problematic students were able to play much better once the concept of the head as a "control tower" directing every muscle movement and the idea of repairing malfunctioning biomechanical units in body, arms and fingers were abandoned. These were replaced by a method according to which the human being is seen as a self-organising organism, capable of finding a personally suitable way of generating and playing music, once a musical will is there.

The question as to whether all body parts behave "correctly" is irrelevant, because a correct mechanical behaviour no longer exists. Thus, learning to play is not seen as learning a specific set of predefined and optimised actions. This approach to learning is comparable to the process of learning to walk. Small children do not learn a concept of walking by defining their body in units that must be trained separately. There is a will to get to a particular place. There is the example of walking adults. And there is the method of trial and error to experience for oneself. Using these methods will result in a personally specific way of walking.

Biesenbender criticises Galamian's approach, which suggests that "[i]nterpretation is the final goal of all instrumental study, its only *raison* *d'être*" (Galamian, 1962, p. 6). He argues that from a historical point of view, interpretation of compositions is a relatively new form in the performance of music, whereas improvisation is a much older and more generic form of musical activity. An understanding of violin playing which assumes a predefined music that is simply be be played back using a set of accordingly predefined specific actions is, for Biesenbender, a non-generic approach to music.

For him it is crucial to work with and understand the human being, the instrument and music as a whole. He sees the method of understanding a process by defining a set of essential properties—and modelling the process by subdividing it into functional units—as counterproductive in violin playing. "Methodical concepts, which orientate themselves to fixed properties and targeted conditioning of movement, programme violinistic conflicts from the outset"[18] (Biesenbender, 1992, p. 32) [Translation by Shivaun Heath].

Instead he makes several suggestions for practising in which, for example, he proposes a need for awareness of body and mind, openness to ideas and the will of a composition, sensitivity to the reactions of the instrument, and a perception of one's own will and how to use it as an impulse for the self-organising organism of instrument, player and sound.

> Proper practice is not repetition for the purpose of conditioning but rather playful discovery of properties, resistance, boundaries of the material: of sound, musical instrument and the organisation of one's body. [...] Strictly speaking there is no such

[18] "Methodische Konzepte, die sich an fixierten Einstellungen und gezielter Konditionierung von Bewegung orientieren, programmieren geigerische Konflikte gleich mit ein."

> thing as repetitive practice, for each new attempt is a one-off, improvisatory aspiration towards something, the result of which one does not yet fully know.[19] (p. 34) [Translation by Shivaun Heath]

The player is not seen here as a producer sending information coded into music to a listener. Instead, the musician is seen as an instrument with antennas transforming musical energy into sound (p. 53). A composition is not understood as a manual with instructions to be reproduced, but as a concentrated musical form to be brought to life by ever-new improvisatory striving. A teacher is not an instructor who passes on knowledge and skills in order to implement a set of functions into the student, but a gardener helping a student to find his way and grow in personal musical ability (p. 53). Playing technique is not a program of gestural functions to be stored, but is gained by putting oneself into a position of openness to contact and reception.

5.3.2 Conclusions for Requirements

The material presented above describing the actions of string players will be used in this section to generate requirements for a synthesiser played by a traditional string player who is interested in the extension of his repertoire of sounds, timbre and colours in electronic and synthesiser sounds. A

[19] Sachgemässes Üben ist nicht konditionierendes Wiederholen sondern spielerisches Erkunden der Eigenschaften, Widerstände, Grenzen des Materials: von Klang, Musikinstrument und Körperorganisation. [...] Strenggenommen gibt es kein wiederholendes Üben, denn jeder erneute Versuch ist ein einmaliges, improvisatorisches Zielen auf etwas, dessen Ergebnis man noch nicht bis ins Letzte kennt.

large number of synthesis engines already exist which are able to create musically interesting sounds. And the question now is how to build a system which will allow the string player to make use of pre-existing skills to perform synthesised sounds. What should such an instrument look like? Which requirements should it meet? One can use the descriptions provided in section 5.3.1 and conclude from them which tasks should be included in the requirements document.

According to the approach taken by Galamian (1962), the instrument should allow the production of vowel and consonant tones with both the bow and the left-hand fingers. It should offer a bow system that behaves and feels like a system of springs, and that enables the player to connect to this system with arms and body and perceive the new instrument, bow, fingers and arm as one complete spring-like system. The bow should be connected to the sound in such a way that an adequate response in terms of bow speed, position, pressure and sound result is given. The system should allow the player to articulate the tone by using consonant-vowel combinations as provided by the player herself. It should allow for the production of different characters and styles of tones. This list is not complete, but these requirements are ones that should certainly be met when a Galamian-orientated string player wishes to transfer existing skills to the playing of synthesised sounds.

Menuhin's approach would add to the list of requirements the ability of the instrument to allow the player to use a "flowing co-ordination" of elements which are "themselves in constant flux of speed and proportion" (Biesenbender, 1992, p. 5). In order to avoid conflicts in performance, the instrument should not push the player in a direction in which exact predefined actions must be made in order to achieve a musical goal. When learning from the instrument, the performer should not have to understand himself as a set of separate units serving the input of the instrument by applying fixed and formulated gestural actions. Instead, the instrument should allow for and direct the user towards a way of performing in which the self-organising abilities of the performer can be used. The instrument should offer a musically meaningful resistance, it should be interesting to explore it again and again, and it should allow the user to perceive it as an extension of the human body.

On the basis of of actions I make when playing the viola, for example, the instrument should meet the following requirements: it should allow the use and variation of the vibrato of fingers, hand and arm with an adequate sound result. It should offer different sound results when played with a relaxed or unrelaxed body and flowing breath, etc. With regard to the perceptible sound results, it should allow the player to generate variances in pitch, dynamics and intensity, in the same way as is possible with a traditional stringed instrument. It should be possible to pull the tone more or less, to lean into the sound, to play a more or less open and full tone, and to make a note smaller or larger and more or less laboured or penetrating.

I have presented many different requirements here, and yet I am sure that the list is not complete because I have only taken examples of performers' actions from the approaches to string playing described in section 5.3.1 and my own approach. While the list of tasks is not yet complete, one can already see that the requirements in related research are limited. If a basic approach to playing a string instrument includes the production of consonants with the bow as well as with the left-hand fingers, it becomes clear that a system such as the Hypercello (see section 4.4.2), which does not track the actions of a performer to generate these consonants with the left-hand fingers, will not be open to such an action. The virtual violin SuperPolm does not offer a bowing system which behaves like a system of springs at all. Both instruments appear to have been built according to a very rough idea of string playing, oriented more closely to the opportunities of interface technology than to the requirements which may be identified by studying diverse established schools of violin playing.

The approach to musical performance in general, and to violin playing in particular, that is described by Menuhin and Biesenbender raises the question of whether the building of computer-based instruments with a formally constrained measurement apparatus providing the input parameters and sensory space of a traditional stringed instrument, might not generate digital instruments that are, as a result of their in-built structure, counterproductive in relation to their requirements. If it is seen as counterproductive to teach a student using a specific set of gestures and actions in order to play the instrument properly, it can be similarly problematic if the instrument forces the player to make a specific set of gestures and actions because it does not recognise and respond to the "grey areas" around those specific actions—as, for the example, with the pitch-tracking problems described in section 5.2.3.

One might argue here that post-Galamian and post-Menuhin violin teachers such as Simon Fischer, for example, explain the sensory space of a violin and the way it can best be experienced and used by providing violin exercises specifically addressing this space and its constraints. In his book "Basics" (Fischer, 1997) he suggests a large range of exercises designed to optimise playing technique and to study the violin's response to, for example, bow speed (p. 48), bow pressure (p. 54), arm movement (p. 145), and vibrato (p. 224).

However, these exercises are not new, because Fischer writes that "[m]any of the exercises have been used widely for decades" (p. vii) and that in violin playing "new ideas usually turn out to have been thought of before" (p. vii). His idea of violin playing can be seen to follow the tradition of Galamian, because he thanks "Dorothy DeLay, whose basic exercises were not only the inspiration but also the starting point of this book" (p. vii).[20] Dorothy DeLay studied with Ivan Galamian and worked for several years as Galamian's teaching assistant.

Furthermore, the idea of sensory space to which Fischer's exercises relate, is not seen here as the violin's sensory space as *is*, but as it may or may not be described. This is linked to the methods of formalisation and the understanding of objects described in section 5.1.2. While Fischer's understanding of a performance technique in which many of the performer's actions "can themselves be broken down into several elements" (p. vi) and in which optimal physical actions are proposed and trained to "build technique" (p. vi)

[20]Simon Fischer studied violin with Dorothy DeLay (Fischer, 2010, para. 6).

is successful, the contrasting approach described by Biesenbender ("Proper practice is not repetition for the purpose of conditioning", as cited in section 5.3.1, p. 156) is successful too. I consider the two approaches to be different in their abstraction and thus their understanding of the violin and of violin playing. However, both use the *same* instrument.

The differing ideas about the performers' actions, and Menuhin's experience that violin playing includes an openness to and incorporation of very many elements—as well as an ever new adaption to those elements, which change their properties dynamically—raise the question of whether it is in fact a good idea to select a specific set of actions (from the level of requirements) when it comes to the specification.

Galamian (1962, p. vii) acknowledges that his method of violin playing is not the only right or possible one. And he is convinced that presenting violin playing in a written book is problematic, because the "very best thing a teacher can give to a student is the individualized, unique approach, which is too personal a thing to be put on paper anyway" (p. vii). If it cannot be put on paper, then it cannot be described completely in a written form, and thus does not meet one of the criteria necessary for formalisation (see section 3.3).

It is possible to define a very rough set of actions on a macro-structural level, such as pressing the finger on a string or pulling the bow back and forth across the string. However, this does not reflect the skills of trained players that I wish to address and use on a micro-structural level. Since these subtle but relevant actions are different from person to person and can vary or be adapted according to situation-specific factors while playing, and since it can even be important for the player not to separate and define them, it is necessary to find an approach to specification which does not force the developer to define, measure and formalise essential actions in advance. The following section will suggest such an approach.

5.4 A Signal-Driven Approach to Sound Synthesis

What might a specification look like if I wish to avoid both modelling the user and building an interface to serve a specific set of actions expected on the basis of a model of the user? In most of the approaches taken in related research, an interface with a fixed set of input parameters is specified in order to drive existing methods of sound synthesis. Improvement of playability is achieved by refining the model of the user and thus modifying the interface.

The approach presented here turns the tables. I take a satisfactory interface as my starting point and modify synthesis methods accordingly. As I do not wish to define the subtle actions of the performer, I must find another way to move from the requirements level to the specification. My way of doing this is to look at stringed instruments and to retain a basic element with its properties: the oscillating string.

5.4.1 The String as Information-Carrier

As indicated by the name stringed instruments, the string and its oscillation play a central role in the making of music on bowed stringed instruments. Actions such as putting a finger on a specific position on the fingerboard, playing with a specific bow speed and contact point, changing the amount of bow hair used, and so on, are all done in order to generate specific oscillations of the string. I consider the string to be the essential object which bow and fingers act on in order to approach a desired sound as closely as possible.

Following from this, I assume that the oscillating string implicitly contains all the musical information that a performer has put into a sound, since all relevant actions taken in the creation of a sound involve modifying the string.[21] Machover must have had a similar conviction. While talking about further work with the Hypercello, he mentions that "the next step [...] will be to try and get all the information back off the strings" (Levenson, 1994, p.16).

In contrast to the idea of deriving explicit information from the string, and in view of the problems described above in defining what information to get from the string and how to build measuring devices adequately, I propose to build a system allowing the use of the strings implicit information. In

[21] I am aware of the fact that the string is coupled to the body and that the information coming back from the instrument to the performer includes haptic feedback from strings, bow, chinrest and sound emitted from the body of the instrument that again affects the performer. What I am referring to here is the question of the actions a performer applies to the system in order to create a specific sound, and the information the string's oscillation contains once the actions of the performer have been carried out.

other words, I make use of a method of sound synthesis that can handle the implicit information of a bowed string. Thus, what I am seeking is to use synthesis algorithms that are mainly driven by the unanalysed audio signal of an instrument's oscillating string.[22]

5.4.2 Basic Architecture

An ideal algorithm has only the audio signal as an input signal, as outlined in figure 5.6. Given such an ideal algorithm, the formalised playing parameters in the signal-driven system do not form the pillars on which the synthesised sound of the instrument is built. Instead, the raw and unanalysed audio signal of the oscillating string is the pillar on which the sound is built.

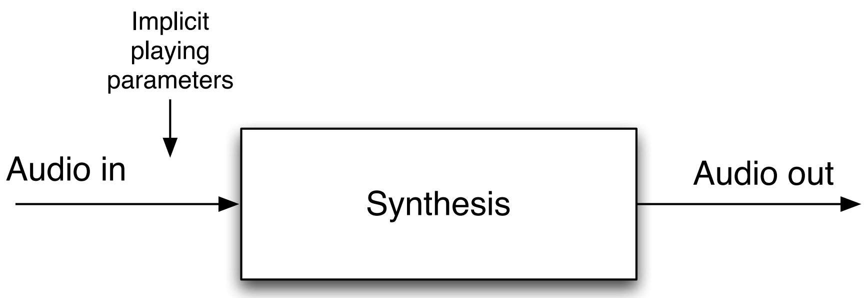

Figure 5.6: Ideal architecture of ASDSS.

Since phenomenological objects or parameters such as pitch, articulation or an action of the performer exist for the player—no matter whether they

[22]In an acoustic bowed stringed instrument, the string is coupled to the body resonating back onto the string. In this context the vibrating string cannot be treated as a discrete physical entity. Thus, the question can be asked as to what exactly the audio signal of an oscillating string is in this context. While capturing the explicitly isolated raw audio signal of the string alone will not be possible, experience has shown that an audio signal picked up close to the oscillating string (at the bridge) of an electric viola works properly.

are physically modelled, measured or measurable at all—and are incorporated into the audio signal of the instrument, it is not necessary to measure such parameters in order to provide the pillars—the explicit data of those objects and parameters— with which to feed the synthesis engine. As soon as the performer executes any sort of playing method or parameter, these are implicitly contained within the sound of the string and thus within the instrument's audio signal.

In order to carry out the transformation of the objects from the requirements to the specification, it is not necessary to go through the steps of abstraction and modelling. Therefore the following advantages present themselves:

1. With regard to the feature of reduction (see p. 36), the phenomenological method of playing an instrument does not need to be reduced to explicit, physical and functional models of playing, and can retain its complete and phenomenological properties.

2. With regard to the pragmatic feature (see p. 36), the individual model and playing style of the performer can retain its unique character and does not need to be modified to fit the needs of physical and functional models.

3. With regard to the problems in formalisation of measurement (see p. 40), the problem of a wrong or incomplete formalisation of measurement does not occur (as long as no measurement is done).

To give an example: in order to give the instrument the ability to transmit the vibrato of the performer into the sound result, it is not necessary to use a model of what vibrato is and how all its features can be measured. The reduction which comes, for example, with a model defining vibrato as a slowly oscillating alteration of the fundamental frequency, and with the accompanying "blindness" to other features that may be important for an adequate representation of vibrato, does not come into play, since the instrument's audio signal implicitly includes the perceivable vibrato without any formalisation of the concept "vibrato". In other words, a formalisation of the concept "vibrato" is not necessary in order to make the instrument open to the vibrato of the performer. Similarly, it is not necessary to formalise the bow-instrument interaction using available but reductive models. With regard to the wrong or incomplete formalisation of measurement, tracking mistakes in pitch (e.g. when playing a noisy tone with vibrato) and tracking latency do not come into play.

The openness of the instrument is not generated by the definition and formalisation of essentials. Instead, it is generated by leaving essentials (i.e. the performer's playing methods, actions and parameters used) non-formalised.[23] Since I am constructing a computer-based instrument I inevitably have to formalise as well, of course. The strategy here is to formalise a framework in which the essentials which are to be transmitted to the synthesised sound can be included implicitly. I see this framework as consisting of the instrument's audio signal, driven by the oscillation of the string.

[23] This parallels the idea of Trogemann and Viehoff (2005), who propose to create openness in a computer application by leaving essentials non-formalised (p. 146).

With a synthesiser using the parameter-driven approach, the performer must use the formalised and explicit playing parameters the instrument offers. Thus, I view the expression one can achieve with such an instrument as *formally constrained expression*. In contrast, the signal-driven approach aims to allow the use of implicit playing parameters that need not be formalised. To the extent that the instrument is transparent to such parameters, one can achieve what I would consider *formally unconstrained expression*.

The question concerning the openness and playability of a specific ASDSS implementation is not primarily connected to completeness and adequate formalisation, but to the ability of the whole system to transmit specific sets of playing parameters from the input to the output. While an ideal ASDSS algorithm would be driven only by the unanalysed instrument's audio signal, in my experience it is necessary to use feature extraction of the audio signal too. These parameters are used to modify the sound generated in the synthesis unit indirectly. Figure 5.7 presents the resulting structure.

In summary one may say that the signal-driven approach is defined by the principle that the unanalysed audio signal is the primary driving force connecting the player with the input of the synthesis engine and controlling the synthesised sound. In addition, if useful for the sound result, any other input parameter (for example an extracted feature of the audio signal) relating to an action of the performer can be used here for the indirect modification of a sound already created in the synthesis engine. Thus, the sound result is formed directly by the playing methods, performer's actions and phenomenological parameters which the performer puts into the string,

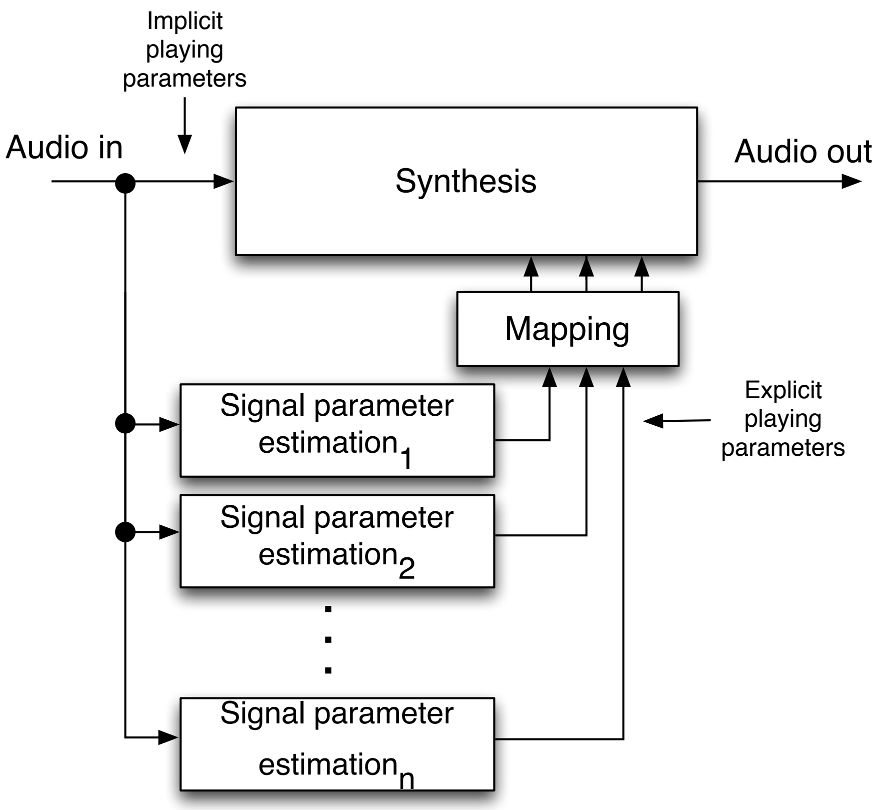

Figure 5.7: Basic architecture of ASDSS.

and indirectly by the formalised measurement data based on physical models of playing parameters. The question of the extent to which the specific algorithm is open to the player's input must be answered by looking at the individual implementations of ASDSS algorithms.

5.5 Hypothesis

Research on sound synthesis and interface construction for real-time use has generated a huge body of work. Almost all research published on sound synthesis is based on the parameter-driven approach as described in section 5.1. Given the lack of literature on the signal-driven approach described in the previous section, one might conclude that this method is not considered to have significant advantages over the parameter-driven approach—assuming that it has been considered at all.[24] In order to establish whether this approach merits further research, the fundamental question is whether the signal-driven approach can potentially show musically significant differences when compared to the parameter-driven approach. Thus, the following hypothesis has been formed:

Implementations using the signal-driven approach to sound synthesis can show musically significant differences in playability and sound result when compared to similarly complex implementations using the parameter-driven approach.

The work presented here represents an investigation of this hypothesis. It is my aim to establish whether the hypothesis will be verified by the results. Besides this, I have the following questions and aims:

- I wish to explore this method, its advantages and disadvantages.
- Are there obvious reasons why the PDSS approach might differ in major respects from the ASDSS approach?

[24]Of course one cannot draw conclusions about the judgements or motivations of other researchers in the absence of any explicit statement on their part. However, I had several discussions on ASDSS during my visits to NIME and ICMC conferences between 2004 and 2008, as well as on a research visit to the Music Technology Area at McGill University in Montréal in 2006. In these discussions I got the impression that some researchers were very interested in this concept, while the majority were somewhat reluctant and questioned the potential of the signal-driven approach, if not rejecting it per se.

- How is the ASDSS approach perceived by musicians?
- How far can I get with the signal-driven approach in terms of different synthesis variants and sounds?

The phrase “can show musically significant differences” rather than “does show musically significant differences” is used in the hypothesis because instruments are used in different personal, social and aesthetic contexts. The structure and meaning of such contexts with all their possible variances cannot be defined in advance. Thus, a musically significant difference can be validated empirically only in a specific context, and not in such a way that it will show the same differences in any context of usage.

5.5.1 Definition of Terms

In order to avoid misunderstandings concerning the hypothesis, the terms used must be defined. The term “sound synthesis” has already been defined at the beginning of chapter 5 and the term “playability” on p. 2 of chapter 1. What is understood by the “parameter-driven approach” has been explained in section 5.1, and what is meant by the “signal-driven approach” in section 5.4.2.

By “musically significant differences” I understand differences which are meaningful in the perception of music and recognised by a large enough number of people to draw statistically significant conclusions. A “sound result” is understood here as the perceptible sound at the output of the system in relation to a desired idea of sound. The sound result may differ from the desired sound idea but should offer a convincing quality in relation to the idea. Furthermore, a "similarly complex implementation" is understood here as an implementation with a similar complexity in terms of algorithmic structure, size of software and technological effort used to build the entire system.

5.5.2 Method

Since I am interested in how much I can achieve with the signal-driven approach, I have tried to modify a set of known synthesis methods in such a way that they can be driven by the raw audio signal. The implementations are presented in section 6.2. To learn more about their advantages and disadvantages in practical use, I tested such implementations in cooperation with both performers and composers. This work is described in sections 7.2 and 7.5. In order to investigate the hypothesis and to compare both approaches on a scientific basis, the following two studies were done:

1. Study with performers to evaluate parameter and signal-driven approaches. In this study, instrumentalists were given three different instruments. They were asked to perform as guided by a questionnaire, and to fill in their perceived results. For each question in the questionnaire, the instruments were rated on a five-point scale. The data was analysed in order to get an impression of the different assessments of the test subjects.

2. Study with listeners to evaluate parameter and signal-driven approaches. In this study, four instruments were selected for comparison. Eight musically relevant criteria were defined for investigation. Listeners listened to short sound examples. An A-to-B comparison of two instruments at a time was conducted with all instruments and all eight musical criteria. Listeners were asked to assess which one of the instruments was better in relation to each of the eight different musical criteria.

These studies will be described in sections 7.3 and 7.4.

5.5.3 Artistic and Scientific Approaches

Art and science meet in this dissertation. I would like to distinguish between the scientific and the artistic approach to the topic of this dissertation, because traditionally some rules and methods used in science are different from those used in the arts. A scientific outcome must be valid no matter where, when or for whom. The musical outcome must work for the parties involved with the project. What is a good sound for a composer or musician must not necessarily be a good sound for another person, even if they come from a similar cultural background. Composing or performing always includes some personal research. If results are found to be convincing from the personal and subjective point of view of the artists concerned, then that view counts, even if it is perceived to be wrong by another party. In contrast, findings in science must be independent of an individual observer.

The artistic elements in the dissertation can be seen in the development of synthesis algorithms (chapter 6); design of sounds (section 6.6); my personal subjective evaluation of the instrument, its sounds and its playability (section 7.1); and the user group *hot_strings SIG* (section 7.6). The scientific elements can be seen in the investigation of the hypothesis; the studies (sections 7.3 and 7.4); the measurement of sound results (empirically); and the conclusions drawn from the measurement results.

Chapter 6

Implementation

How can synthesis methods using the principle of ASDSS be implemented? What possibilities does it offer in a real-life situation? What limitations in variations of sounds will be found? In order to answer these questions I implemented synthesis methods based on ASDSS in the framework of an instrument. Since I am a viola player, the framework of a viola was selected. It was my aim to build prototypes and to present these to musicians to see how they responded to them. In addition, I wanted to be able to compare the prototypes with already existing methods. Besides this, I simply wanted to perform music with these new sounds myself. In general I wanted to gain experience with ASDSS sounds and evaluate at least a minimal set of ASDSS methods.

6.1 Soft- and Hardware

My approach is not concerned purely with interfaces, mapping methods or a synthesis method. Instead, it concerns the complete set of interface and synthesis algorithms. From the diagram in figure 5.7, I view the following areas of work as particularly important to the implementation and improvement of ASDSS instruments:

- the sound synthesis engine,
- the input audio signal in relation to the audio transduction, position of transducer and the transducer itself,
- the explicit parameters extracted from the audio signal,
- the mapping of explicit parameters to the synthesis engine.

These areas of work can be separated into soft- and hardware. I see work to be done in the software domain as follows:

- the development of algorithms representing the synthesis engine,
- audio signal feature extraction,
- the mapping of parameters to the synthesis engine,
- the software interface of the completed synthesiser.

In addition, hardware-related work to be done would be:

- the development of the hardware interface (in this case a bodyless viola),
- the selection of an audio transducer and placement.

In order to keep within the time frame of this dissertation, a selection of areas to work on had to be made. The criteria for selection were: most important for sound result, availability of resources, and compatibility with the time frame. On the basis of my estimation in relation to the criteria, I selected the following areas: sound synthesis engine, hardware interface and audio transducer, and audio signal feature extraction.

The development of ASDSS algorithms are described in section 6.2. Since I wanted to use extracted signal features to modify the synthesised sound indirectly, I collected available feature extraction algorithms, selected the ones that could be controlled by the performer, and improved data extraction according to the given context, in my case the audio signal from a bodyless viola. This work is presented in section 6.4.2. In the early phase of my doctoral research, I also planned to use parameters from sensor systems. I wanted to use familiar and proved systems for this, and improve them where possible. The work done here is described in section 6.4 and in Appendix D.

6.2 Signal-Driven Synthesis Methods

Since conventional synthesis engines are controlled by a set of discrete input parameters, the question arrives how synthesis engines can be driven by a raw audio signal. What is done here is to take existing synthesis methods and modify them so that they can be driven by the audio signal. This approach raises the question of which synthesis methods might be modifiable in this way and how they might be modified.

Experimenting with the use of the audio signal in synthesis algorithms, I came to the conclusion that convincing results can be produced by replacing oscillators from traditional synthesis methods with the oscillation of the audio signal. It is conceivable that other methods of modification exist. However, having worked on this question I have found that the best results to date came from using this method. The following sections will present my selection of conventional synthesis methods to be modified, and the implementation of these modifications.

6.2.1 Selection of Synthesis Methods

In previous work, I had found two algorithms to be of use, one based on subtractive synthesis, one based on simple FM synthesis (Poepel, 1999, pp. 21-22). Having established this, I then wanted to extend the simple versions of those approaches, and so looked for other methods which could be modified for use with audio signals. I took as my starting point from a list of conventional synthesis methods that I felt to be the most popular ones. This included additive synthesis, waveshaping synthesis, granulation and physical modelling. In addition, I wanted to work with variations of FM synthesis such as multiple carrier FM, multiple modulator FM, feedback FM and phase modulation (PM) synthesis.

In additive synthesis a set of (usually sine-wave) oscillators is summed together to create a waveform as achieved by the user. In order to be flexible in terms if variances of harmonic spectrum over time, the oscillators have separate envelopes for their amplitudes. The model of sound and the understanding of timbre needs explicit parameters to control the fundamental pitch, select the partials and define the envelopes. With regard to a performer and an interface these parameters must either be provided explicitly by the performer's user interface or must be defined in advance (and will then not be modifiable while performing).

Exchanging an oscillator with an instrument's audio signal would be counterproductive because the harmonic structure would no longer be defined completely by the oscillator generated partials and their amplitudes. Looking at the basic structure of additive synthesis I came to the conclusion that this method may not be suitable for the signal-driven approach. Either I would have to analyse the audio signal and drive the additive synthesis algorithms on the pillars of discrete parameter data I extracted from the audio signal; then this would lead to a PDSS implementation, which is not my focus here. Or I would replace one of the additive synthesis oscillators with the audio signal of the instrument; then the basic idea of creating a sound by adding sine oscillations would be destroyed. With respect to the idea of creating openness by leaving essentials non-formalised, additive synthesis does not seem to fit because it is based on the principle of improving the synthesis by defining and formalising essentials, i.e. by adding more oscillators and better envelope controls. Since I did not find a work-around for these problems I decided to focus on the other synthesis methods I had on my list.

I had already done basic experiments with waveshaping (Poepel, 1999, p. 21), using the audio signal instead of a sine oscillator in front of the wavetable. The result was that the synthesised sounds I obtained were very different from the input signal in terms of dynamics and harmonic structure. Since in conventional waveshaping a sine oscillation is shaped by a shaping function table I tried to solve the problem in the sound result by using a low-pass filter to eliminate higher partials of the instrument's audio signal. However, I did not find the sound results convincing. Setting the low-pass filter to lower cutoff frequencies I lost the ability to form the timbre on the instrument. Setting the filter to higher cuttoff frequencies the problem of inadequate timbre appeared again. In the present research, the following approach to waveshaping was used: the audio signal was multiplied with a sine oscillator and the oscillator's frequency input was driven by the fundamental frequency extracted from the audio signal. While it was possible to increase the influence of the instrument's audio signal and the implicit playing parameters slightly, I could not detect sufficient influence by the performer on the sound. When the gain of the audio signal was increased, the problems described above returned. Thus, the algorithm was not transparent enough to the performer's input. In terms of the important criteria, the need for an algorithm's openness to the implicit playing parameters as mentioned on p. 168, the algorithms I constructed were not found to be sufficient.

Granulation, also known as granular sampling (Lippe, 1994), is a method that can be used either with pre-recorded samples or with a live audio signal. Here, the latter was selected and useful results were obtained. The question was how this approach needed to be applied in order to generate timbres that are perceived as independent sounds, and whether these algorithms would be responsive enough to a string player making use of pre-existing skills to shape the sound. These questions are addressed in section 6.2.9.

Three approaches in physical modelling may be seen as most important. These are waveguides (Jaffe & Smith, 1995), mass-spring model (Cadoz et al., 1984) and modal synthesis (Adrien, 1991). Since physical models are essentially built on the principle of modelling physical objects, one may speculate that the use of an audio signal in this context might disturb the function of the models rather than improve it. This is particularly the case with mass-spring models because the models for vibrating strings, for example, are built to be driven by a model of a bow rather than an instrument's audio signal. Looking at the architectural principles of the three approaches, I estimated waveguides to be the one that could be modified most easily in the method I am pursuing here. I searched for ways to use the simulation of the body of musical instruments and planned to apply the pickup signal of the e-viola to the input of the waveguide algorithm. Following the experiences in my implementations, however, I eventually found myself working with sound results that were not interesting enough because the results sound very similar to those obtained through more straightforward filtering approaches. The use of a Karplus Strong algorithm driven by the audio signal provided sounds that were different enough in timbre to be perceived as individual. However, as this algorithm may be seen as a filter, I viewed this approach as a type of subtractive synthesis, and did not continue with further attempts to modify physical modelling algorithms. However, it may be possible that future research in signal-driven approaches to waveguide synthesis may find ways to provide other results.

Since I had several synthesis methods providing fruitful results I concentrated on those. The following sections will describe my approaches to them.

6.2.2 Subtractive Synthesis

Subtractive synthesis is one of the early methods of sound synthesis and usually involves filtering a noise oscillator, often with band-pass filters. The modification carried out here was to replace the noise oscillator with the audio signal. Partials of the audio signal are treated using band-pass filters. A pitch estimator controls the band frequency of the filters. Figure 6.1 presents this method. Different sounds can be obtained by setting the amplitude and bandwidth of the filters, and by determining the partials to which the band-pass filters are set (Poepel, 1999, p. 26).

If the bandwidth of the filters is set to a very low value (narrow band-

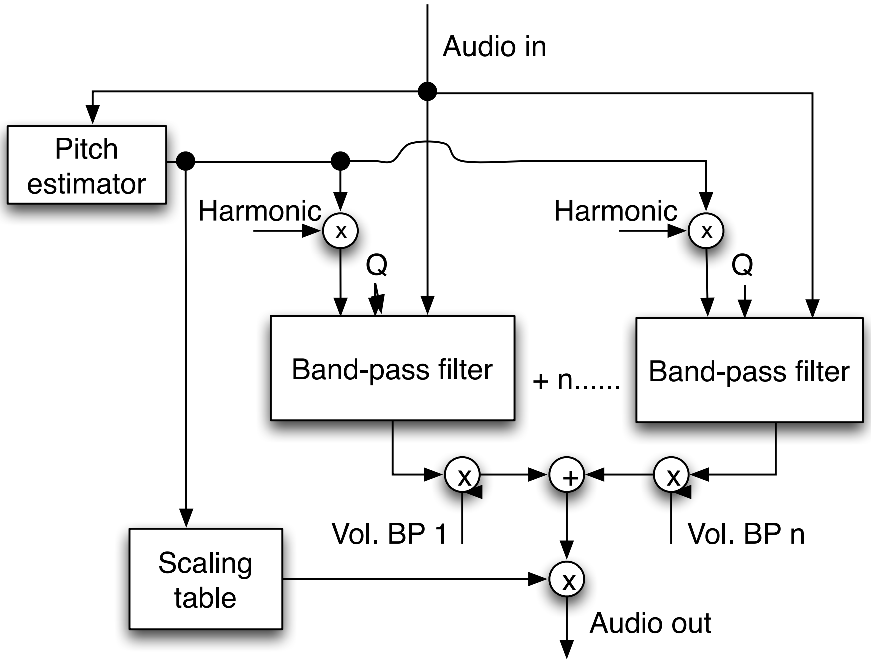

Figure 6.1: Signal-driven subtractive synthesis.

width), the difference between the sound result of a noise oscillator and that of the audio signal at the input side is minimal. From the point of view of perception, one can morph between signal-driven and parameter-driven sound synthesis by varying the bandwidth of the filters. Filtering partials out of an audio signal which is mixed with noise can be an interesting way to mix signal- and parameter-driven synthesis. An implementation of signal-driven subtractive synthesis can be found in Appendix C: see MaxMSP patch "6.2.2 SD_SubtrSynth".

With regard to the phenomenological parameters I applied when designing and improving this algorithm, I was encouraged, because I had the impression that the inner structure of a sound could be explored. In contrast to the pure audio signal of the instrument, the potential of the synthesised sound to be formed by me as a player was increased. My personal assessment of a sound using this algorithm can be found in section 7.1 on p. 233. In addition, this algorithm's sound has been used in the listeners' study described in section 7.4.

6.2.3 Simple FM Synthesis

FM synthesis was developed by John Chowning between 1967 and 1968, and presented in a publication five years later (Chowning, 1973). The frequency input of the carrier oscillator is modulated by the sum of a carrier frequency and by the scaled output of a modulator oscillator. The method can generate a rich spectrum using only two sine oscillators. The flowchart of a typical (parameter-driven) FM synthesis algorithm is presented in figure 6.2.

Taking the FM synthesis as it is presented in figure 6.2 as a starting point, the modification carried out here was to replace the modulator oscillator with the audio signal. As string players do not work with triggered predefined notes, I cannot trigger pre-programmed envelopes. Instead I use an envelope follower to drive the amplitude of the carrier oscillator. Should one wish to use ratio, as is common in FM synthesis, a pitch follower is also needed. The values from the pitch follower are then multiplied by the ratio and added to the modulator signal. As is common in FM synthesis, the ratio can be used to generate inharmonic spectra or to change the octave. It is possible to use this algorithm without the pitch tracker. If the ratio is set to zero, the pitch tracker values will be multiplied by zero and thus have no impact

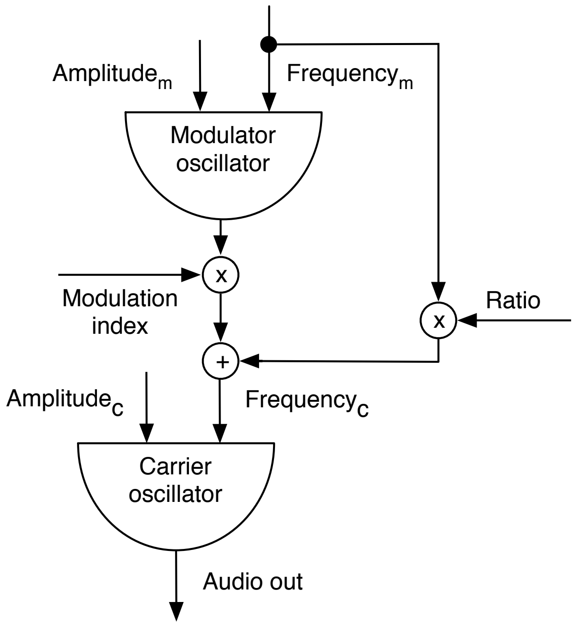

Figure 6.2: Parameter-driven simple FM synthesis.

on the sound result (Poepel, 1999, p. 27).

As found in experiments with signal-driven FM synthesis it was fun for me to play with this algorithm because the sound results were interesting—especially with ratios below 1. I could give the tone production a different character in comparison to the instrument's audio signal. While I had to adapt my playing method slightly to generate warm, aggressive, soft or sharp sounds the ability to generate a variety of sounds on the fly was encouraging. Unfortunately, I saw that the sound result becomes disproportionately harsh when played loudly. A scaling table was therefore added. Depending on the amplitude of the input signal, this lowers the signal's amplitude before it is fed into the frequency input of the carrier oscillator. A flow chart of the signal-driven simple FM synthesis algorithm can be found in figure 6.3. An implementation of the algorithm can be found in Appendix C: see MaxMSP patch "6.2.3 SD_simplFMSynth".

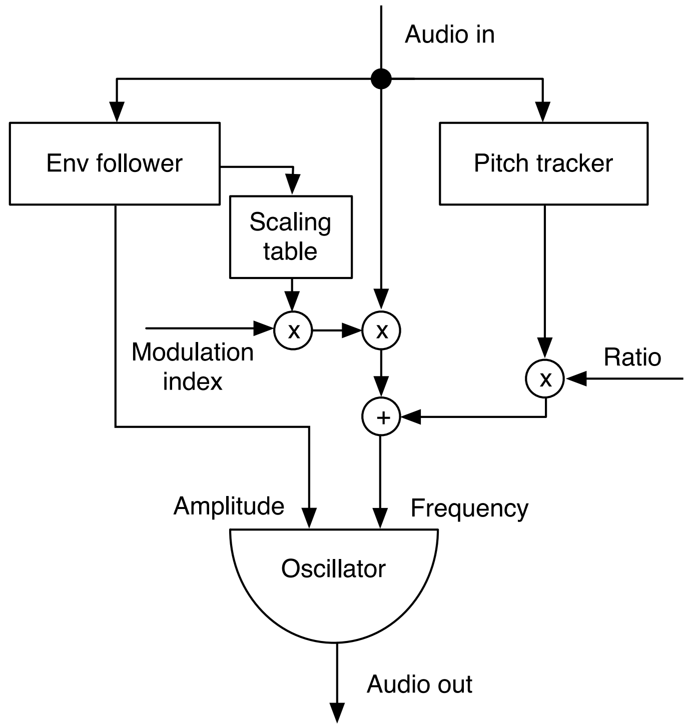

Figure 6.3: Signal-driven simple FM synthesis.

My personal assessment of two sounds using this algorithm can be found in section 7.1 on p. 232. Furthermore, this algorithm's sounds have been used in the players' and listeners' study described in sections 7.3 and 7.4. Since the modification of simple FM synthesis led to interesting sound results, the question arose as to whether and how extended versions of ASDSS simple FM synthesis could be implemented.

6.2.4 Multiple Carrier FM Synthesis

Multiple carrier FM is based on the principle that one modulating oscillator modulates several carrier oscillators. If it is possible to modify simple FM synthesis, it follows that multiple carrier FM synthesis can be implemented too.

The implementation can be carried out easily by modifying the algorithm presented in section 6.2.3 in such a way that an audio signal modulates several carrier oscillators. Figure 6.4 gives an overview of what an implementation of modified multiple carrier FM might look like. An implementation of the algorithm can be found in Appendix C: see MaxMSP patch "6.2.4 SD_MC_FMSynth". Besides the addition of similar carrier output signals, the use of different parameter settings may be interesting.

In comparison to the signal-driven simple FM synthesis algorithm described in the section above, the multiple carrier FM algorithm widens the spectrum of sounds one can define by setting the variables of the algorithm. Different ratios applied to different carriers provided results that were—from my point of view—more open for my input. Especially in fine tuning the settings of a sound this method offers good opportunities. While I have implemented a multiple carrier FM algorithm with two carriers, more than two carriers are of course possible. Using them with different scaling tables for the index calculation and different ratios I expect that this algorithm will offer a large range of slightly different sounds.

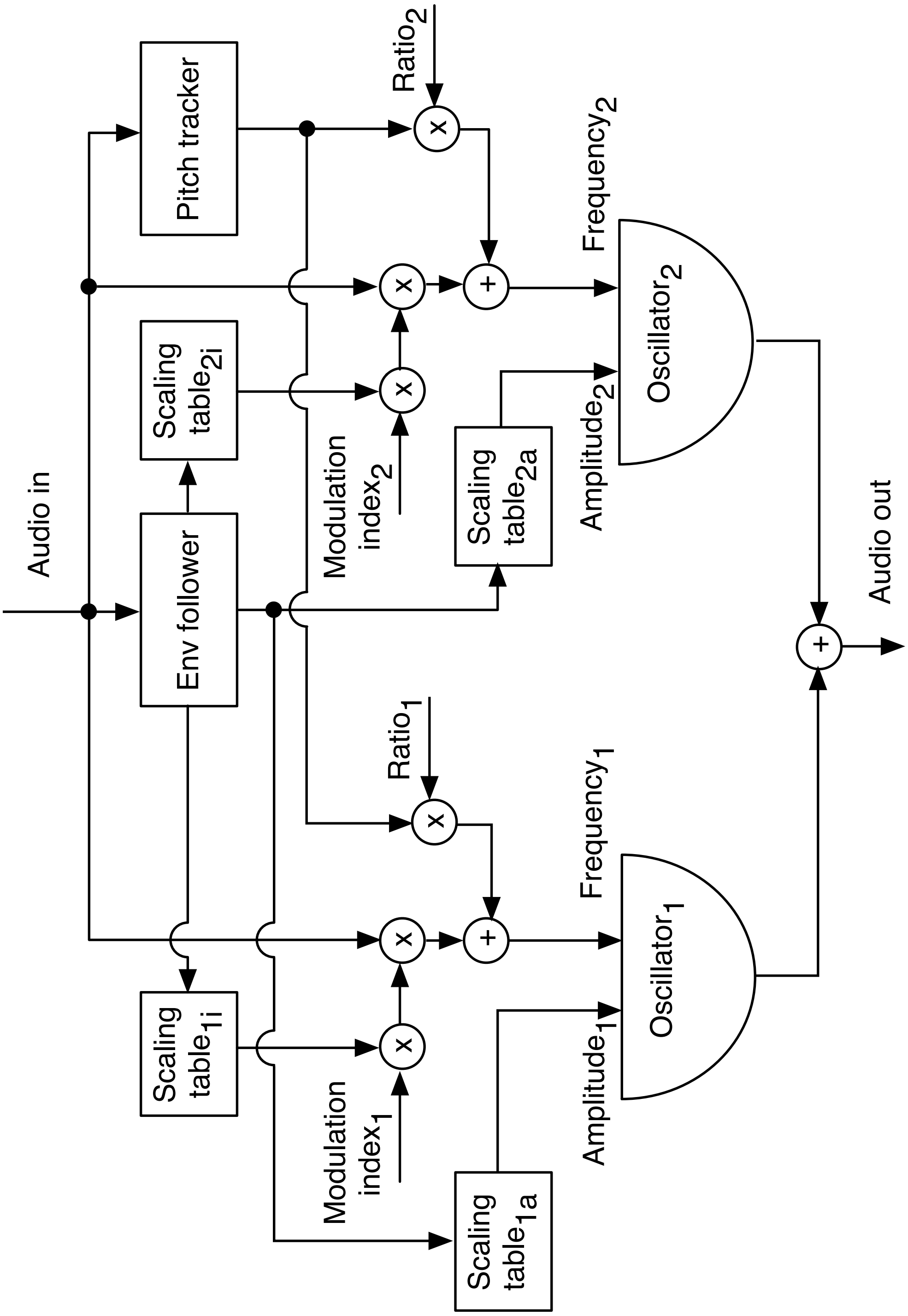

Figure 6.4: Signal-driven multiple carrier FM synthesis.

6.2.5 Multiple Modulator FM Synthesis

Interesting sound results were achieved with multiple modulator FM synthesis. In multiple modulator FM, two or more modulation oscillators are added, and the sum is used to modulate a carrier oscillator. When the doubled audio signal is used for modulation, the sound result contains a huge quantity of high partials. My proposal for dealing with this problem was to use filtered partials of the audio signal as modulator sources. Band-pass filters were therefore tuned to the partials. By doing so, I was able to select partials including the noise around the partials. For flexibility the implementation allows the following modulator sources to be selected:

- the input audio signal,
- a filtered partial of the input audio signal (pitch tracker controls centre frequency of the band),
- a sine tone (pitch tracker controls fundamental frequency).

Figure 6.5 presents an overview of the structure. It is common in multiple modulator FM for the amplitude of the modulation oscillator to be driven by a triggered envelope. As I am working with a stringed instrument interface, common note-on triggers for envelopes are not available. The conventionally used envelopes often have a significant influence on the timbre of an FM sound which I did not want to forgo. What I therefore did was to implement an envelope follower tracking the amplitude of the modulator signal, and pass the resulting parameter on, via a variable scaling table,[1] to a multiplication by the modulator signal. This allows the modulation index to be varied depending on the amplitude of the modulating source. An implementation of a variant of signal-driven multiple modulator FM can be found in Appendix C: see MaxMSP patch "6.2.5 SD_MultModFM".

This method produced sounds with satisfying results during the steady state of the tone. The transient phase, however, proved to be problematic. The problem is that if the pitch tracker cannot find a pitch during the transient phase, both modulation sources—the sine tone and the filtered partial—are generated after the pitch tracker has found the pitch. This is too late to allow a convincing sound result in the transient phase, which is, therefore, negatively effected. Accordingly, when testing the algorithm, playability was perceived to be reduced during tone attack and articulation.[2]

If the modulation source is the less filtered instrument's audio signal, the disturbance of the pitch tracker is reduced, because more of the transient phase of the tones performed on the instrument's string form the sound result. This feature may be used as an example how the definition of the essentials amplifies the problems coming with the necessity to formalise them. If one selects a specific filtered partial and a sine tone as modulation sources, these

[1]The scaling table is calculated from the parameters offset (b), scaling factor (a), and exponent (n). The formula for calculating the scaling table is: $y = a * x^n + b$

[2]The research result indicating reduced playability raises the question of whether an application including this implementation's features could potentially be musical. It is conceivable that this implementation could be mixed with other methods, thus giving a better response in the transient phase, or that additional sensory input could be used to capture data relating to performed pitch earlier.

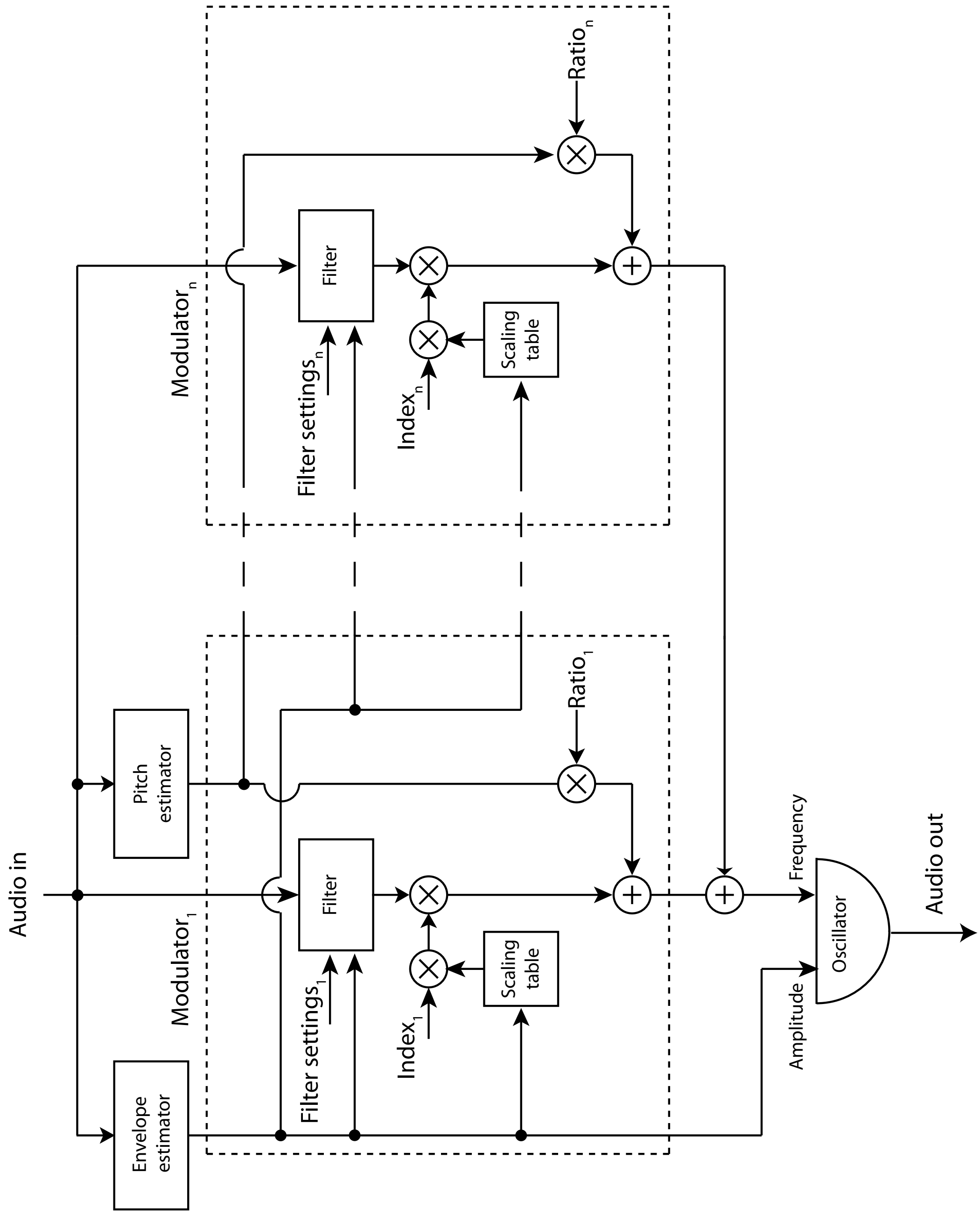

Figure 6.5: Signal-driven multiple modulator FM synthesis.

two form the pillars on which the sound is built. One remarkable feature is that the information about the harmonic spectrum the performer produces to influence the sound result is reduced to the information coming from these two sources. Another remarkable feature is that insufficient formalisation of the pitch tracker (see section 5.2.3, p. 141) disturbs the sound result to a greater degree compared to its use with the raw audio signal as in signal-driven simple FM. My personal assessment of a sound using this algorithm can be found in section 7.1 on p. 234.

6.2.6 Feedback FM Synthesis

In feedback FM synthesis, the output of an oscillator is fed back to the frequency input of the oscillator itself, or to a modulation source for this oscillator. Feedback FM can be modified to include ASDSS methods too. For example, the sum of an audio signal and the scaled output of the carrier oscillator can be used as a modulation source for the carrier oscillator. Figure 6.6 presents this implementation.

Since I work with MaxMSP, a simple feedback using a patch chord and a delay is not possible. The reason for this is that the signal processing in MaxMSP is carried out in vector blocks, but most objects do not store their output between calculation cycles. A solution to the problem of creating a feedback loop is to use objects that do store the previous vector. Thus, to implement a feedback, a workaround is to use the objects send~ and receive~

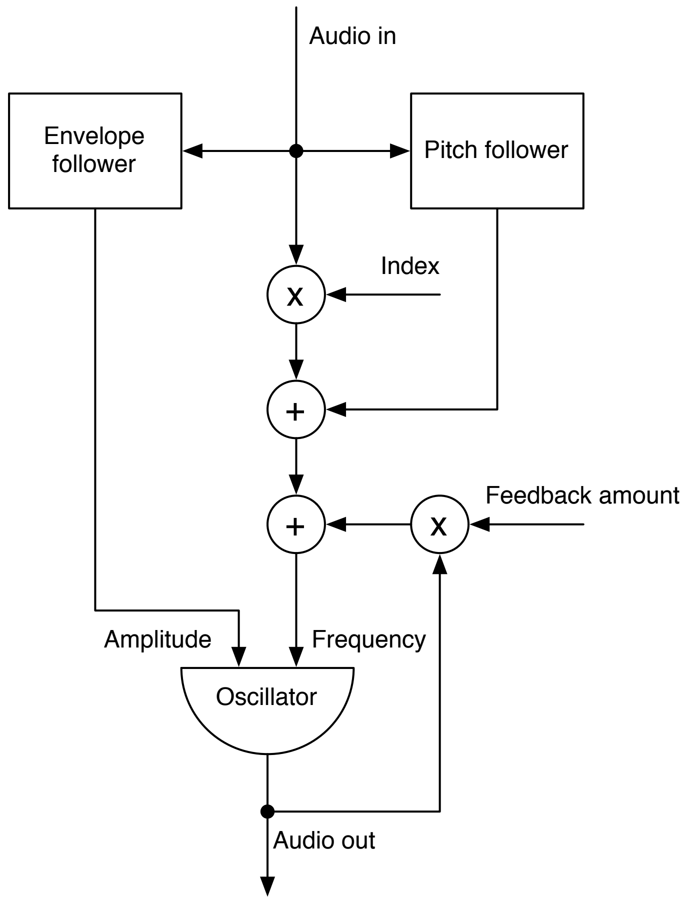

Figure 6.6: Signal-driven feedback FM synthesis.

or to use tapin~ and tapout~.[3] An implementation of a signal-driven feedback FM algorithm can be found in Appendix C: see MaxMSP patch "6.2.6 SD_feedbFM".

A problem in feedback FM is the influence of the feedback on the carrier frequency. The result is an unwanted variation in the fundamental frequency

[3]Another solution would be to write an MSP external implementing the feedback oscillator. This alternative was not considered necessary as the system worked properly with the standard MSP objects mentioned.

which is a very important parameter for the musician. Problematic sound results can emerge, varying from an irregular vibrato to a noise sound depending on the fundamental frequency and the amount of feedback. This problem again relates to the question of the algorithm's transparency for the parameters a player uses.

An alternative to FM synthesis is phase modulation (PM) synthesis. By using feedback PM synthesis, the problem of unwanted pitch variations can be eliminated. This advantage means that I did not focus any further on implementations of feedback FM. My approach to feedback PM synthesis is explained in the following section.

6.2.7 Phase Modulation Synthesis

Phase modulation (PM) synthesis is a variation of FM synthesis. Instead of modulating the frequency of an oscillator, the phase is modulated. In contrast to FM synthesis, PM synthesis does not exhibit the problems of pitch instability which feedback FM shows as a result of the fact that the frequency input is not modulated. In parallel to FM synthesis the modified variants of simple FM, multi carrier FM and multi modulator FM described above, can be implemented in PM synthesis. Figures 6.7 and 6.8 present two examples of these algorithms. Three implementations of signal-driven PM can be found in Appendix C: see MaxMSP patches "6.2.7 SD_feedbPM", "6.2.7 SD_MultModPM", and "6.2.7 SD_simplPMSynth".

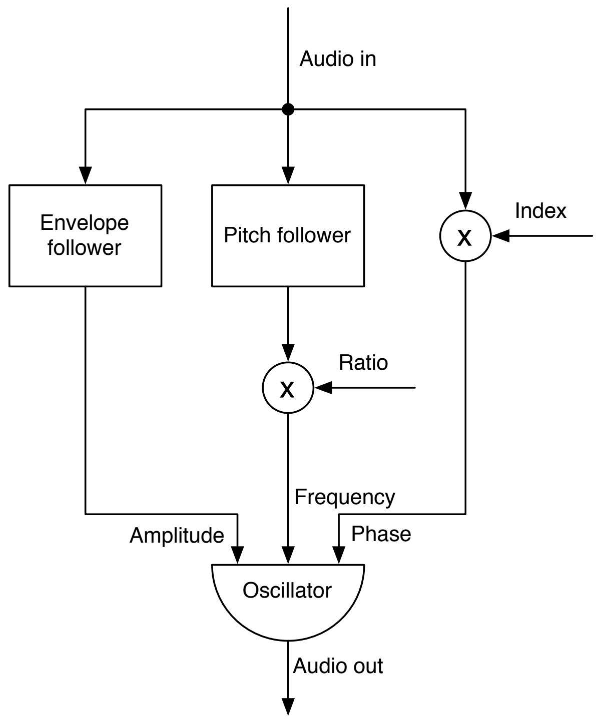

Figure 6.7: Signal-driven simple PM synthesis.

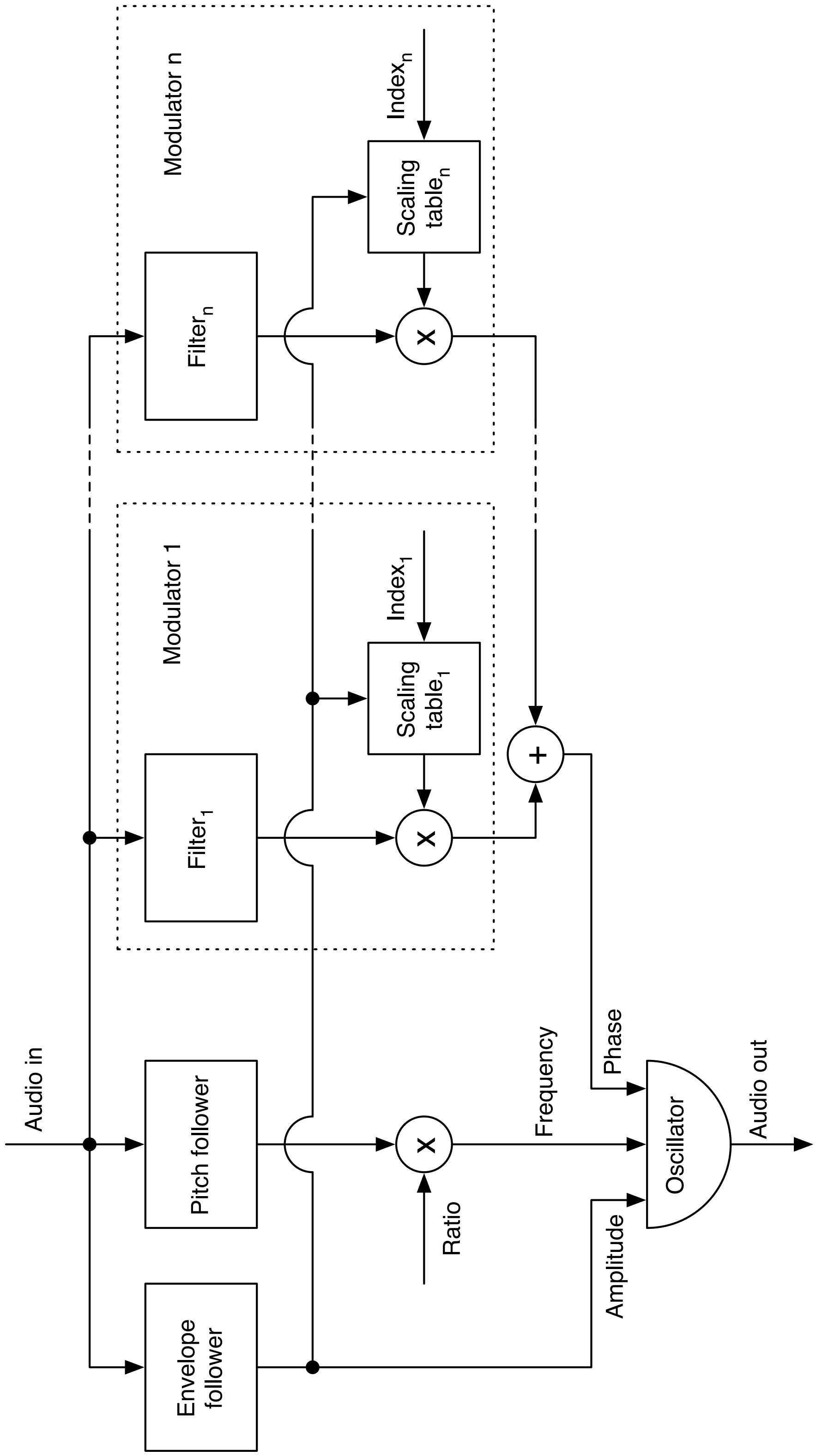

Figure 6.8: Signal-driven multiple modulator PM synthesis.

6.2.8 Combinations in FM Synthesis

One of the commercial synthesisers using PM synthesis was the Yamaha DX7.[4] This instrument was introduced in 1983 and was the first digital synthesiser that it was possible for a wider range of musicians to buy as a result of its relatively low price. The synthesiser sold very well, and its sounds were often heard in pop songs after 1983. As presented above, I was able to modify various FM synthesis methods to be driven by the audio signal. I wanted to know whether the sounds available on the DX7 could also be modified according to the ASDSS approach. The PM synthesis methods of the Yamaha DX7 and their related parameter settings were therefore implemented in MaxMSP and modified.

A virtual version of the DX7 is the FM7 introduced by the company Native Instruments[5] in 2001. The current commercially available version is its successor, the FM8. To enable me to conduct the research presented in this chapter, Native Instruments donated an FM8.[6] Like the FM7, the FM8 is almost identical in functions and sounds to the DX7. In order to build signal-driven variants of existing sounds, three FM7 presets were selected for modification: no. 3 (FM8: Tronflute), no. 13 (FM8: Pipe Organ 2), and no. 21 (FM8: Sparkle).

[4]The DX7 is officially said to use FM synthesis. In fact, however, it uses PM synthesis (Holm, 1992, p. 40). I assume that one of the reasons for this is the better response in terms of pitch stability when using feedback PM.

[5](Native Instruments, 2010).

[6]Thanks to senior sound designer Martin Jann of Native Instruments.

The DX7 offers six oscillators which are referred to as operators. It allows combinations connecting these operators to be built using multiple carrier, multiple modulator and feedback PM synthesis. The FM8 of course allows the same combinations. The selected presets show the operator matrix presented in figures 6.9, 6.10 and 6.11.

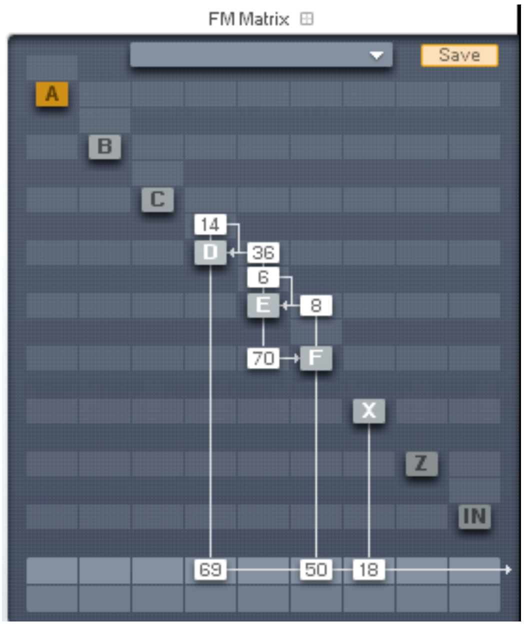

Figure 6.9: Operator matrix of FM8 preset *Tronflute*.

The boxes with the letters A-F each indicate an operator. The numbers around each operator indicate the degree of modulation the operator applies to itself (feedback), and to other connected operators. An arrow on the left, right or upper side of an operator indicates that this operator is modulated by the operator from which the arrow comes. More parameter settings for the

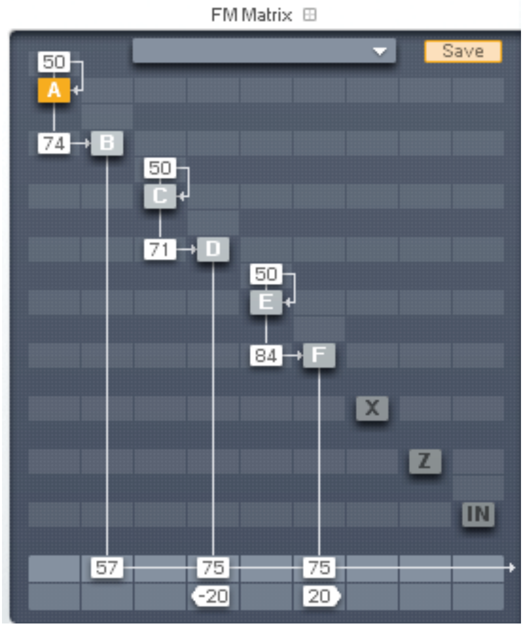

Figure 6.10: Operator matrix of FM8 preset *Pipe Organ 2*.

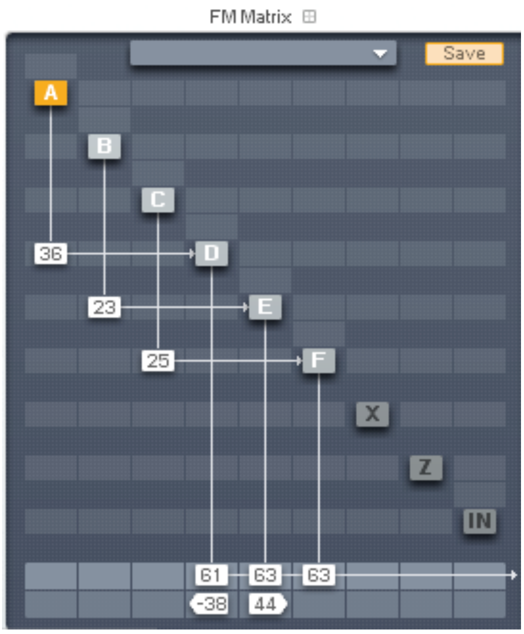

Figure 6.11: Operator matrix of FM8 preset *Sparkle*.

sounds *tronflute*,[7] *Pipe Organ 2*[8] and *Sparkle*[9] can be found in the footnotes.

This structure was redesigned in MaxMSP. As the numbers which indicate the degree of modulation (range 1-100) do not represent a linear relation to a scaling factor between 0.0 and 1.0 but an unspecified non-linear relation, I reverse-engineered these connections through FFT measurements, and compared the results using both the FM8 and a MaxMSP patch. According to my results, the numbers used in the FM8 (modulation) represent a square relation ($y = x^2$) to the scaling factor between 0.0 and 1.0. The feedback is also a square, however 100 is not related to 10, but is much smaller.[10]

Each operator in such a matrix has its own envelope driving the amplitude of the operator. Since I am working with a bowed stringed instrument as an input device which is not designed to provide note-on triggers, a predefinition of triggered envelopes would contradict the goal of good playability. These envelopes were therefore not implemented, but instead replaced with envelopes with lookup-tables which scaled the operators' amplitudes as simi-

[7] Ratios and offsets of the operators in the sound *tronflute* (figure 6.9) are set as follows: Ratio: D: 1.0; E: 1.0; F: 1.0. Offset (in Hz): D: 0.0; E: 0.0; F: 2.34.

[8] Ratios and offsets of the operators in the sound *Pipe Organ 2* (figure 6.10) are set as follows: Ratios: A: 4.0; B: 0.0; C: 2.0; D: 0.0; E: 1.0; F: 0.0. Offset (in Hz): A: 0.0; B: 0.19; C: -0.20; D: 0.17; E: 0.20; F: 0.15.

[9] Ratios and offsets of the operators in the sound *Sparkle* (figure 6.11) are set as follows: Ratio: A: 9.0; B: 6.0; C: 4.0; D: 3.0; E: 2.0; F: 1.0. Offset (in Hz): A: 0.0; B: 0.0; C: 0.0; D: 0.0; E: 0.0; F: 0.0.

[10] I do not think that I can establish this with 100% precision because the problems of vector-based signal processing mentioned earlier prevent me from achieving a proper feedback in MaxMSP. Leaving the degree of feedback on the same level but varying the signal vector size results in a change in the sound. One cannot distinguish between the incorrect degree of feedback and the incorrect vector size. Since I expect that the implementation of the FM8 allows sample-precise feedback, I conclude that I am not able to create a proper reverse-engineered scaling table.

larly as possible with reference to the envelopes of the emulated presets. It would have been beyond the scope of this research to implement all the complete sound effects and the LFO unit of the DX7. For this reason, and because a simple rebuilding of such a unit does not constitute original scientific work anyway, these units were not implemented.

The operators used in all three selected presets are sine wave oscillators. As I conclude from section 6.2.5, replacing these oscillators with the audio signal would not lead to a sound result similar to the original because of the rich spectrum an audio signal modulator generates. Again, the algorithm would not be as transparent as necessary. To reduce this problem, filtered partials of the audio signal were used. The partials were set according to the ratios of the operators. The frequency offsets in the presets of the FM8 were also implemented.

In implementing these sounds I was again confronted with the fact that the selected preset sounds of the FM8 use several triggered envelopes controlling parameters such as ratio, index and oscillator amplitudes, amongst others. Since ASDSS does not use note-on triggers, the question of how I should deal with these envelopes arose. I decided both to replace the envelopes with fixed numbers, and to use discrete parameter data from an envelope follower and map this data via scaling tables to the parameter inputs of the synthesis engine.

In my work with the FM8 sounds I initially expected that the character of a sound preset would be defined primarily by the operator settings in the matrix. In setting the envelope to fixed values, however, it became clear that the character of the sound was greatly dependent on the development of the envelopes. In my experience, the sound may derive more of its origin, identity or character from the envelopes controlling the operators in the matrix than from the timbre defined by the combination of operators.

The goal of this research is to enable a player to play a synthesised sound with pre-existing skills. This section describes my attempts to preserve the character, form and appearance of an FM8 preset sound. If it is integral to the character of the sound that it is driven by automatically varying synthesis parameters (for example envelopes and LFOs), there is no way to stop these automated aspects and still keep the character of the sound. I conclude that a concept based on triggered envelopes cannot be translated into a concept of continuous control without the loss of sound characteristics. It is conceivable that one might emulate the control movements of the automated envelopes and work on a meta-structure which allows the player to influence those movements. However, if such envelope movements are considered crucial to the character of the sound, it is clear that the playability of such a sound will be dependent on direct access to these movements.

A sound programmed in a synthesiser such as the FM8 is anticipated by its programmers to function in a timeline-based way, with LFOs and envelopes controlling the synthesis parameters. The timeline-based concept of envelopes is regarded here as a horizontal approach to the programming of a sound, because the actions controlling parameters within a sound are performed relative to a timeline started by a note-on trigger. A violin player may also conceive a sound in a timeline-based way. However, the sound will be generated from a pool of possibilities which the player can deal with and influence at every single moment. The sound of a violin is thus more strongly defined by the setting of variable parameters at any given moment.

In contrast to the horizontal concept of sound design, I consider the programming of the connection between continuous control parameters which come from the performer and the input parameters of the synthesis engine to be a vertical approach to programming the sound. As one can only obtain continuous control data from a performer's analysis, I have had to use the vertical method of conceiving a sound in programming.

In order to have a comparable basis, I deleted the envelopes of the FM8 preset sounds and set the varying envelope values to fixed numbers. With regard to the results I obtained when implementing the preset sounds in a signal-driven way, I did achieve sounds that were, in the steady state, similar to the envelopeless presets. When leaving the envelopes in the form of the original presets, I was not able to generate the sounds with the perceptible characters as known from the original presets. Thus, it was not possible in the course of this research to get to sounds known from the DX7 and play them in the manner ASDSS allows. The reductions resulting from the missing triggered envelopes and from the problems in the transient phase of the tone change the character of the sound to a considerable extent. This extent is too great to claim that one is still playing the same sound, from a perceptual point of view.

What I learned from this experience is that the DX7 is an instrument that produces sounds coupled to the construction principle of PDSS. Since the character of the sounds is dependent on the oscillators, their interconnection, the envelopes and the explicit parameters coming from the keybord, it is hardly possible to achieve similar sound results without the construction principle of PDSS. One might ask, whether the character of sounds produced by other synthesisers is similarly coupled to this construction principle.

6.2.9 Granulation

Granulation is a synthesis method in which samples of sounds are cut into small pieces. These small pieces, usually called grains, are then played back and mixed together. The playback of grains can be carried out with different speeds, amplitudes or panning positions. Truax (1987) proposes a real-time granularisation of stored samples. In addition, Lippe (1994) presents the possibility of granulating the sound of a live audio signal. Using Lippe's terms this method falls into the category of granular sampling. Since I am focusing on the use of an audio signal for sound synthesis, I have chosen to work with a real-time version of granular sampling.

I designed the algorithm and parameter setting so as to be a perceptibly distinct timbre in relation to an electric viola. Good sound results were achieved. I used the following two methods to implement real-time granulation in MaxMSP: the first is based on a delay with several read-out points. In MaxMSP I use one tapin~ and several tapout~ objects. The tapout~ objects are comparable to read heads of analogue tapes which read the audio signal from a tape with a write head provided by the tapin~ object. The read-out heads vary in speed and length of read-out time, and an envelope scales the amplitude of each read-out head. The second implementation is based on a record~ and groove~ object combination. A buffer is filled with samples, and as soon as the end of the buffer is reached, the sample write point jumps to the beginning of the buffer. Several groove~ objects read the buffer with varying speeds and amplitudes. An envelope controls the amplitude of each grain.

In real-time variants, granulation is often driven by parameters coming from an interface. I use the ASDSS principle here in the sense that I try to employ as few explicit control parameters as possible. In addition, I structure the algorithm and parameter settings in a way that keeps as many of the audio signal's implicit playing parameters "alive" as possible. Thus, the ASDSS variant presented here differs more in the goals for the design of the granulation algorithm than in basic construction principles.

The question presents itself as to how implicit playing parameters can be kept "alive" during granulation. To answer this question, two examples will be given: the bowing technique martelé and the parameter of pitch. How must the parameter settings of the algorithm be set so that martelé bowing will yield an adequate sound result? In martelé, the attack phase of the tone which can be described as beginning with a "bite" is important. To preserve the attack phase of a tone, the time duration of grains must be carefully adjusted to small values. Grains with too long a time duration cause a time shift in the attack phase, and using a large number of grains, each set to a different duration, may wash out the “bite” from the attack phase. Variation of the read-out speed of grains results in pitch variation. Using this pitch variation generates several sorts of interesting jittering or noisiness of a played tone. The pitch variation of grains, however, can be adjusted in such a way that a pitch played by the performer at the input is hardly perceptible at the output. If not adjusted carefully to small degrees of pitch variation, the parameter of pitch can be negatively affected.

The direct impact of granulation on the instrument’s live audio signal presents many possibilities for changing the shape of a sound in a fundamental way. If a performer has found a sound that does not allow specific parameters to survive, but that offers highly interesting results musically, she is free to decide whether the playability, while reduced, is still good enough to use the sound in a specific context. Three implementations of granulation can be found in Appendix C: see MaxMSP patches “6.2.9 Granulation groove~”, “6.2.9 Granulation munger~”, and “6.2.9 Granulation tapout~”.

6.3 Signal-Driven Implementations by Other Researchers

One of my goals is to inspire researchers to explore their own implementations using the basic construction principles of ASDSS. In the following sections (6.3.1 and 6.3.2) I will present two methods that use the idea and construction principles of ASDSS and have been developed by other researchers.

6.3.1 Self Modulation

Inspired by the presentation of the paper "Synthesised Strings for String Players" (Poepel, 2004), Roger B. Dannenberg developed a method called self modulation. This method was presented in a paper (Poepel & Dannenberg, 2005, p. 393). Instead of replacing one oscillator with the audio signal, both oscillators of a simple FM synthesis algorithm are replaced with the audio signal. The audio signal is fed into a delay whose delay time is modulated by the audio signal itself. In order to prevent negative delay times, the signal is biased. Figure 6.12 illustrates this algorithm.

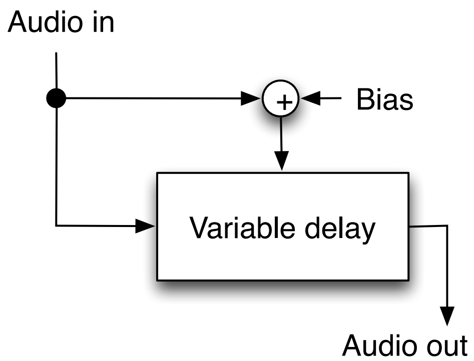

Figure 6.12: Self Modulation.

Since the sound shows a very large quantity of high partials and is thus perceived as too harsh, the input signal is filtered with a low-pass filter as described in figure 6.13. This reduces the quantity of higher partials. However, musical parameters dependent on higher frequencies can also be restricted by the filter.

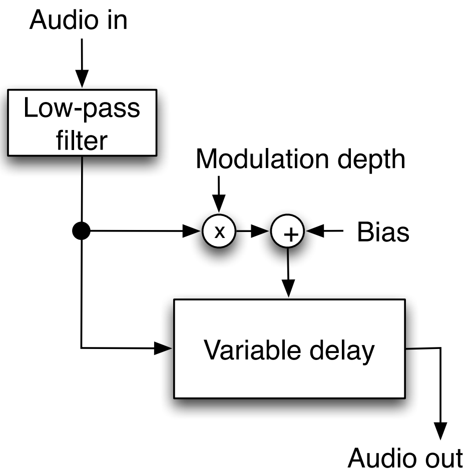

Figure 6.13: Improved Self Modulation.

6.3.2 Adaptive FM Synthesis

Lazzarini, Timoney, and Lysaght (2007) present a method called "adaptive FM synthesis". In their paper they state that this method falls into the category of ASDSS (p. 21). Figure 6.14 presents the basic structure.[11] One of the reasons for their use of this architecture is the conviction that the sound result will "retain significant gestural information contained in the original signal" (p. 21).

In contrast to self modulation (section 6.3.1), which can be seen as a single carrier modulation, they "propose a technique similar to multi-carrier FM". Their assessment, after implementing this method and testing it, was that with "this technique it is possible to have a fine control over the synthetic result" (p. 26). I conclude that the goal of an increased playability is met by their development.

[11] Image from Lazzarini et al. (2007, p. 23).

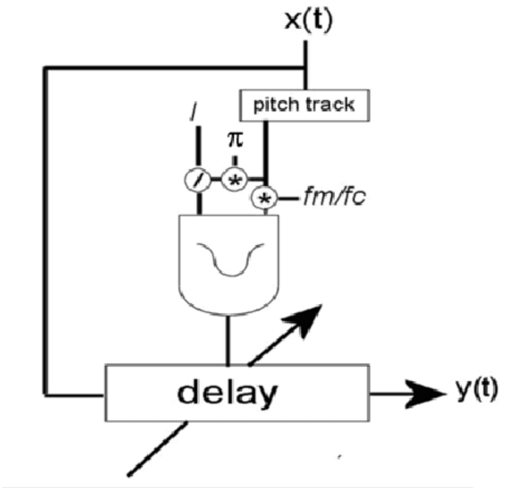

Figure 6.14: Basic structure of adaptive FM synthesis.

6.4 Sensor Systems

Since the audio signal plays a central role in this work one important factor here will be the audio transducer generating this signal. Work done in relation to this will be described in the following section. In addition to using extracted features from the audio signal, I planned to use sensor data to modify the output of the synthesis engine indirectly. When I first began working on the thesis I thought to select, from the sensors available on the market, those that help to generate the most meaningful data.

Working on this approach another solution was arrived at. I did not add ready-mades such as knobs, faders or distance or pressure sensors to the existing bodyless viola corpus, as was done in the related research described in sections 4.4.3 and 4.4.6, for example. My experience suggested that this would have caused the player to focus on the sensors, and thus to learn and use new gestures. Moreover, I do not believe that custom sensors offer the same depth of explorability, potential for learning or openness for musical ideas as strings do. I was therefore looking for a system that would adapt sensor technology to the needs of string players, rather then forcing the player to adapt to available sensor technology. The intention was that it should, as far as possible, allow the player to make use of pre-existing skills and the expressive potential of stringed instruments.

I decided to rebuild the sensor technology of the MIT stringed Hyper-instruments (see section 4.4.2). The reasons for this decision were that the system had proved to be usable on stage and had been used by world-class musicians such as Yo-Yo Ma and Joshua Bell. Besides this, I wanted to be able to compare those systems with my system. Ultimately, however, the challenge in terms of electronic engineering, and qualitatively limited results of the test system, took this work beyond the scope of the thesis. It was not pursued further. The work done on this project is described in Appendix D.

6.4.1 Bodyless Instrument and Audio Signal Transducer

In the first phase of my doctoral research I worked with a bodyless viola built by luthier Arwed Harms (2010). I asked the luthier to build a bodyless instrument which would, in relation to the size, fingerboard and neck, replicate as far as possible the feeling of my acoustic viola. The bodyless viola (figure 6.15) has the same measurements in body length, set-up, neck length, fingerboard length and width as the acoustic instrument. Accordingly, string length is also identical to that of the acoustic model. The weight is slightly greater than that of the original. Apart from the weight and sound, my perception is that the feeling of playing both instruments is similar.

Figure 6.15: Harms electric viola.

Testing different pickups with ASDSS methods showed that they can have a crucial impact on the timbre of sound and its playability. I wanted to find the pickup most suitable for ASDSS. The influence on timbre and playability of sound was found to vary depending on the synthesis method, and even on parameters settings within the same synthesis method. When selecting pickups I had two aims:

1. to find an inexpensive pickup which would offer a financially low entry threshold for string players who wanted to test my synthesis methods, and

2. to find the pickup which works best with the full range of my synthesis methods, and on stage.

First, I listed all the pickups I could find by means of a literature and web search. I then tried to acquire as many of these pickups as possible by asking companies to provide samples for test reasons. In addition, I visited dealers and owners of these pickups and recorded test sessions. The pickups tested were:

1. Olivier Pont (2010) viola pickup,
2. Willi Balsereit violin pickup,[12]
3. Shadow SH 941 (figure 6.16),
4. Shadow SH 3000 (figure 6.17),
5. Schertler Stat-V (figure 6.18),
6. Zeta Jazz Series Pickup (figure 6.19).

Besides the pickups, I also tested two instruments with built-in pickups:

1. Yamaha Silent Violin SV 120 (figure 6.20),
2. Zeta EV 24 Electric Viola (figure 6.21).

[12] Willi Balsereit is a Cologne-based luthier well known for his pickups.

Figure 6.16: Shadow SH 941.

Figure 6.17: Shadow SH 3000.

Figure 6.18: Schertler STAT-V transducer.

Figure 6.19: Zeta Jazz Series pickup.

Figure 6.20: Yamaha Silent Violin SV 120.

Figure 6.21: Zeta EV 24 Electric Viola.

The musical material recorded was open strings, scales over three octaves, and a range of bowings including legato, martelé, détaché and spiccato. Different dynamics and varying bow positions were tested, and I also improvised music to get an impression of the feel between input and output. The synthesis methods I used were modified simple FM synthesis, modified subtractive synthesis, and combinations of the two.

The result of this testing was that the Zeta EV 24 was selected, because it offered the best consistency of sound with different pitches and strings, the sound could be formed most successfully, and the system as a whole was felt to be very robust and thus stage-ready.

6.4.2 Signal Parameter Extraction

The basic architecture of ASDSS (see figure 5.7) includes the use of explicit parameters to modify the synthesised sound indirectly. Such parameters are extracted features of the audio signal. Since I wanted to be able to select from as many parameters as possible, I searched for additional parameters besides the common ones such as pitch and amplitude. My idea was to take the parameter extraction externals available in MaxMSP and to select those which can be controlled best when playing the electric viola.[13]

The parameter extractions I found were spectral flatness measure (noisiness) and spectral centroid measure (brightness). I used the MSP objects brightness~ and noisiness~.[14] My aim was to establish how well the objects' measurement represented the brightness and noisiness I played on the instrument. The objects did not correspond immediately to what I played, which meant that I had to adapt my performed noisiness and brightness in such a way the objects could analyse it. After some practice in playing to make the objects generate stable values, I was able to control the values within a small range as long as the tone was in the steady state. I tested the measurement with some string players by telling them to play more or less noisily and more or less brightly. What I discovered was that the players were not necessarily performing a more or less noisy sound in terms of the measurement algorithm, but in terms of what the test subjects understood by playing noisily. The results concerning brightness were slightly more consistent, but not stable enough to be useful.

[13]This search was conducted during a short term scientific mission at the Music Technology Area of McGill University, Montréal in August 2006. The mission had the goal of investigating which parameters can be extracted from the audio signal of an electric viola, and which can be controlled best. It was supervised by Gary Scavone (CAML, 2010).

[14]Implemented by Tristan Jehan (2010).

With regard to the question of why the ideas of noisiness were so different, talking to the musicians who tested the system revealed that they were following Shannon's noise definition rather than the physical definition of noise in acoustics. Shannon (1948) understands noise as a disturbance of a communicated information signal (p. 19). If a musician wants to communicate a musical idea, everything disturbing this communication can be understood as noise. Thus, it is the existence of the disturbance that is important. How the disturbance comes about is a secondary issue.

For these reasons I abandoned the idea of tracking noisiness and brightness, and focused on the question of the search for the best pitch tracker for the audio signal coming from the electric viola. On the basis of my results, fiddle~ (Puckette et al., 1998) was found to be the best, because of its speed and reliability. To increase the stability of the pitch parameter stream, I used a median filter which reduces sudden sharp value changes in the pitch data stream.

6.5 Signal-Driven Synthesis: Technical Observations

I have already explained one of the specifics of ASDSS, namely that since the sounds are not based primarily on discrete parameters and note-on triggers, I do not use triggered envelopes. Besides this, ASDSS is seen as distinct in several other ways. It offers a framework for what I call "formally unconstrained performance expression" (see p. 168). In an ideal ASDSS algorithm, everything the performer is used to doing in order to modify the sound will be perceptible in the sound result. In contrast to the parameter-driven approach, the question is no longer whether the interface and mapping are open to or transparent for the performer's input. Instead, the question is whether the synthesis algorithm is sufficiently open to transport these parameters. Furthermore, the problem of transparency as described in section 5.2.4 has been shifted from the areas of interface and mapping to the area of the algorithm. I see this as one of the major properties of ASDSS.

In addition, in ASDSS three observations with respect to latency, signal-dependent sound disturbances and the effect of amplitude range on the timbre seem noteworthy. The following sections will address these issues.

6.5.1 Reduction of Latency

Besides the issue of frameworks for expression, a specific feature of ASDSS can be seen in the latency problem. The latency results produced by a pitch tracker are measurable. Thus, one can calculate the latency between a trigger and the start of a tone in parameter-driven synthesis. This latency will always be there. In signal-driven synthesis it is different, since a sound is always produced, even if there is no pitch tracking data available. However, at the start of a tone, it may take a small amount of time until the correct tone is produced, because the discrete parameter of pitch may influence the overtones of the synthesised sound. The "non-correct" phase of the tone (this again depends on the ASDSS method used) may be much smaller when small intervals are performed than is the case with larger ones.

6.5.2 Signal-Dependent Sound Disturbances

When a pitch tracker is used in ASDSS, latency and tracking mistakes may affect the sound despite the fact that pitch tracking values modify the sound indirectly. Disturbances may occur. In the case of modified subtractive synthesis and signal-driven FM synthesis, the pitch tracker has an effect on timbre while finding the correct pitch at the beginning of the tone. Similarly to the latency effect described in the previous section, this effect differs depending on the pitch of the previous tone, i.e. the intervals a performer is playing.

Playing small intervals will result in a very small disturbance of timbre at the beginning of the tone, since the timbre is affected only slightly. In contrast, playing larger intervals will result in greater changes of timbre at the beginning of tones. This disturbance in timbre is thus dependent not only on the latency or mistakes of the pitch tracker but also on the signal applied to the synthesis unit.

6.5.3 Amplitude Range as a Timbre Parameter

In the particular case of signal-driven FM synthesis (see section 6.2.3), the timbre can be changed with a compressor. The amplitude affects the timbre significantly because it influences the FM index. Since many more high partials and different timbres are produced when playing loudly (or with a high index), one can generate those timbres constantly by using a compressor before the input audio signal is fed into the input of the system. Using a compressor will shift the amplitude range to specific levels where the sound results are different. Thus, scaling the audio signal at the input of the synthesis algorithm can take on the function of adjusting the timbre of a sound. Alternatively, this can be done at the volume knob on the instrument.

Essentially there is a problem of potentially competing goals. These are transparent playability versus new sonic possibilities. The two will sometimes work against each other, but the goal is to avoid restricting the use of existing expressive aspects of instrumental technique as far as possible.[15]

6.6 Design of Sounds Using Signal-Driven Methods

Once signal-driven synthesis algorithms are constructed, the next step is to design sounds making use of these algorithms. My goal in the design of ASDSS sounds was to provide sounds which would make ASDSS testable. I wanted to provide a small number of sounds for each developed ASDSS method so as to give an idea of the possibilities of these methods.

Another goal was to design sounds that offer, in my estimation, a musically meaningful timbre which sounds different to the audio signal. Besides this, the new sounds should offer high playability and a close relation to the

[15] If the volume knob is adjusted while performing, the playability may be decreased. This, however, should be no reason for a performer not to use the volume knob, as long as the advantages of using it outweigh the disadvantages.

input applied by the performer to the system. Once these new sounds had been designed, I then wanted to offer them to musicians to see how they would use them.

Among the open questions, then, were whether better playability would really be obtained; whether it would be possible to use personally specific methods of expression; whether the sounds would be perceived as interesting and meaningful by performers; and whether they would be perceived as synthesised sounds or perhaps rather as effects. In addition, these prototypes were to be used as a starting point for the user-centred design of sounds.

To achieve this, a MaxMSP patch combining the synthesis methods described above was developed. It included three versions of signal-driven FM synthesis, one version of modified subtractive synthesis, a vocoder, a single sideband modulation, and a parameter-driven one-oscillator sound synthesis followed by a small effect unit. The patch can be found in Appendix C: see MaxMSP patch "6.6 CBVA". Basic sounds were designed by me, and later by performers and composers.

To make the system clear, synthesis algorithms such as multiple modulator FM or granular synthesis sounds were implemented in separate MaxMSP patches focusing on one specific synthesis method only. As a starting point I used five of my own sounds which were already available (Poepel, 1999) and expanded these to a total of fifteen sounds to be explored. The design process of sounds for each particular algorithm will be described in the following sections. Sound examples can be found in Appendix E.1 (CD 1, tracks 1 to

29). I will describe the experiences with the sounds in chapter 7 (see p. 230).

6.6.1 FM Synthesis

The implementation of the signal-driven simple FM synthesis (Poepel, 1999, p. 21) allows the following elements to be adjusted:

- index (scaling of audio signal),
- ratio (scaling of pitch tracker values),
- offset (offset of pitch tracker values),
- waveform of carrier oscillator,
- scaling table (to influence index depending on input signal's amplitude).

These adjustments were expanded by the ability to affect ratio with the envelope. Mapping of the envelope follower to this can be switched on and off, depending on a specific range of pitch tracker values.

The signal-driven multiple modulator FM patch which was implemented includes two modulation sources and one carrier oscillator, and offers the following parameters to be set:

- selection of modulation source 1 and 2: no signal, full audio signal, band-pass filtered 1st-7th partial,

- exponential function to scale the amplitude of the modulators in relation to the audio signal's amplitude: exponent, scaling factor, offset (see footnote on p. 190).

Sound examples using signal-driven FM implementations can be found in Appendix E.1 (CD 1, tracks 1 to 7).

6.6.2 Subtractive Synthesis

As a basis for further development, I used the method presented in Poepel (1999, p. 22) which uses seven low-pass filters, each followed by a band-pass filter, with both filters adjusted to selectable partials of the input signal. The following settings are offered with which to design the sounds:

- partial to be separated by filtering,
- amplitude of filtered partial,
- offset shift of band-pass filter's centre frequency,
- scaling table of overall output volume dependent on fundamental frequency,
- bandwidth of band-pass filters,
- edge steepness of low-pass filters,
- a unit of LFOs causing the amplitudes of the filters to move fast and thus generate a bubbling sound.[16]

[16]Free-running modulators and LFOs can disturb the playability. Depending on how the parameters of the LFOs are set, however, they can give a specific colour to the sound with minimal negative effect on the playability.

Sounds were designed using the method of concept, trial and error. Combinations of filtered partials which sounded interesting were selected and refined. Sound examples can be found in Appendix E.1 (CD 1, tracks 9 and 11). A problem in designing sounds was the timbre, which—depending on the pitch of tones and the strings on which they were played—was sometimes unbalanced. I expect that this could be improved using common techniques usually called “keyscaling”. Depending on the explicit pitch values, filter settings and amplitudes of partials can be scaled.

6.6.3 Granulation

In my granulation variants, the parameters to be adjusted were the common parameters known in granulation such as length of grains, pitch variation of grains, pitch offset of grains, spatial placement of grains, grain windowing and grain offset. Good noisy sound results could be achieved through granulation when working with very small grains. The playability decreases when using larger grains because of the larger time lag with which one is confronted. However, when performing slow passages, this drawback may disturb the player less.

Depending on parameter settings, one can morph from a sound similar to the pickup signal to a sound which can be perceived as a synthesised sound. I was able to get good results with relatively small grains and a small amount of pitch variation. The spatial placement affects the playability relatively little, but can give a “big-tone feeling”. Granulation sounds can be programmed in very many ways. Once the parameters are set, some sounds offer a large potential to be explored by playing with traditional techniques. Sound examples can be found in Appendix E.1 (CD 1, tracks 31 to 43).

6.7 A Different View of Sound Synthesis

While constructing ASDSS algorithms, a remarkable experience for me was that the design of and reflection on the construction of these algorithms provided a different view of sound synthesis in a variety of respects. At the beginning I had the intention of creating the algorithms from scratch as well as building them on a foundation of known synthesis methods. Working on both, I became aware that the task of imagining an ASDSS method from scratch was difficult, because existing experience in constructing sound synthesis algorithms could only be used in very limited circumstances. When considering what the reasons for this were and where the problems lay, I came to the conclusion that the methods and experiences of how to conceive a synthesis system were problematic in the following points:

- The identity of a sound should not be defined by the synthesis unit, but rather the synthesis unit should offer a framework within which the identity of a sound can be created by the performer's application of implicit (phenomenological) playing parameters.
- One cannot build the ASDSS system by starting with a pitch-driven oscillator and adding several explicit continuous controls, parameters

affecting the spectrum, LFOs and so on, as is common in PDSS.[17]

- One cannot anticipate the sound result on the basis of the connection between explicit input parameters and the synthesis engine, because implicit input parameters should be used as far as possible.

- The envelope-based construction of a sound that one may have in mind is of no use. Instead, the shaping of the envelopes with regard to the implicit playing parameters must be left open to the player.

- The complete instrument must be open to a user's definition of how to play on a microscopic level. One must avoid defining in advance how a performer will conceive an instrument and act on it. Phenomenological playing parameters made use of by performers should be taken into account when the question of the instrument's openness or transparency is addressed.

There are, of course, similarities to PDSS in constructing a complete synthesiser. An effect unit applied to the synthesised sound is constructed similarly, for example. It is possible to map explicit parameters and interrelate these. In general, and looking at the list above, I believe that the methods and the conception differ from, and require a different kind of experience from the PDSS approach.

[17]Although I knew that the basic construction principle of ASDSS differs in this regard from PDSS, I found myself regularly slipping into the construction method of selecting unit generators and explicit parameters to control the input parameters of a formalised sound system and the related sound result precisely.

With regard to the pre-definition of performers' playing methods I would like to point out that electronic instruments such as the TB-303[18] are used in a very different kind of music and playing style from that planned by the constructing engineers. One might argue, that it is not necessary to anticipate the actions of the performer in advance. While this is indeed the case, it cannot be used as an answer to the unresolved problem of playability in bowed string interfaces, because it shifts the problem of playing to the performer instead of solving the problem in the instrument's construction. The question of how a synthesis algorithm I had constructed would react in relation to phenomenological playing parameters was very hard to conceive, if not impossible to foresee in advance. One reason for this was that the physically abstracted idea of sound and music that I knew from what I had learned about computer music, and the phenomenologically abstracted idea of sound and music that I knew from what I had learned as a viola player, were overlapping only to a certain degree.

While the construction of ASDSS methods is difficult from the point of view of conceiving a synthesis engine not based on the pillars of a formal input and formalised connections between sound result and explicit parameters, I see in PDSS a primary problem in the anticipation of the performer's actions and the need to define a parameter hierarchy or interrelation[19] of parameters (mapping). Such an anticipation may or may not suit a particular performer,

[18]An analogue Synthesiser with a stepsequencer; the instrument has cult status in the techno scene.

[19]Garth Paine mentioned the need for research on this topic during a public discussion at the NIME conference in Paris in June 2006.

the music to be performed, or the actual needs of any given moment in a performance.

An alternative to using a predefined and fixed set of input parameters and mapping system might be to have systems that offer many different types of interfaces, mappings and synthesis variants which are able to recognise the actual need of the performer and adapt to this need. This, however, would require intelligent systems that can recognise the user's needs. To my knowledge, such systems are not available in a proper working state at present.

To conclude this discussion on the view of sound synthesis, one may say that the PDSS-related understanding of sound synthesis urges developers to see the human being through the spectacles of physics, and more specifically the type of physics that sees itself as a method by which to define what sound really *is* and what musical performance *is*. One has to measure what the actions of the performer *are* and define in advance what the sound will possibly *be*.

In other words, one might say that PDSS indirectly provides an ontology of sound and performance with which not every musician necessarily will agree. From my point of view this method urges a view of human beings which defines the performer as a unit, dealing with functions and giving information to hands and arms which are, in turn, motor devices applying physical forces to the instrument. I see in this implemented musical understanding, as it were, a major problem associated with PDSS. While constructing and testing the ASDSS algorithms, I saw that there are—to a degree, at least—other possible ways to conceive and construct sound synthesis. A remaining question is, of course, how other musicians perceive these developments. The following chapter will address this.

Chapter 7

Evaluation

I have written much about the proposed advantages using ASDSS methods. The question now presents itself as to whether and how these advantages are perceived by the target users, and how these users assess systems making use of the technology described. My development is designed to be used by string players seeking to expand their repertoire of sounds. These musicians may have very different ideas about how this expansion can be achieved, and what kind of changes or reductions—which are an unavoidable part of such a project—are acceptable. I therefore expected that subjective evaluations would bring varying results. In addition to collecting users' experiences, I wanted to test my hypothesis as it was proposed in section 5.5. Therefore, in this chapter, I will present both subjective evaluations by users, which are open to any properties the users consider important, and quantitative studies looking for specific qualities in my newly developed instruments.

The subjective approach is comparable to the task string players are confronted with when looking for a new instrument. It is common to go to a luthier and to test the instruments with any individual method the performer wants to use. I present the results of my individual experiences in section 7.1, of other string players' experiences in section 7.2, and of stage-related compositions and performances in section 7.5. The two quantitative studies are presented in sections 7.3 and 7.4.

7.1 Personal Experiences

A subjective test is what many researchers do when assessing their results. I used sounds I designed (see chapter 6.6), and describe them here from my personal point of view. The method of ASDSS allows a wide range of different synthesis algorithms and indirect parameter control. Thus, one can expect that an evaluation will provide varied results according to the variety of sounds. It would be beyond the scope of this section to offer a separate evaluation of each sound I or others have designed. Instead, I will concentrate here on a small number of sounds and provide a summary which includes the aspects I found to be of interest or importance.

The sounds I selected were:

1. signal-driven simple FM synthesis; ratio = 0 (pitch tracker not in use); preset no. 2 in patch *6.6 CBVA* (see Appendix C);

2. signal-driven simple FM synthesis; ratio = 0.5 (transposing one octave lower); preset no. 3 in patch *6.6 CBVA*;

3. signal-driven subtractive synthesis; filtered 1st, 3rd, 5th, 7th, 9th, 11th and 13th partials; preset no. 6 in patch *6.6 CBVA*;

4. signal-driven multi modulator FM synthesis; preset no. 7 in patch *6.2.5 SD_MultModFM* (see Appendix C);

5. granularised sound; preset no. 24 in patch *6.2.9 Granulation groove~* (see Appendix C).

I perceive these sounds as follows:

Sound no. 1 (signal-driven simple FM synthesis, ratio = 0) offers good playability as a result of its reaction in terms of pitch variation, variations of nuance in sound, articulation, and tones with a substantial consonant portion. There is no latency other than system latency.[1] Variation of bow speed, bow pressure and contact point clearly affect the sound result, but not as much as is familiar from a traditional instrument. Scratching is no problem, bringing a tone out of a noisy tremolo works well, and bowings such as détaché, martelé, spiccato, staccato and legato provide adequate results. Double stops do not work because inharmonic sound results are generated. If played very loudly, the tone can take on an unwanted harsh quality. When playing very softly, the sound is similar to the raw pickup signal. Playing *ppp* (pianissississimo) can result in getting no sound at all, and the balance of sound across the full range of pitches is not adequate. It is hard to get

[1]In MaxMSP, system latency can be adjusted by setting the hardware buffer size and the vector size. I tested different adjustments of hardware buffer and vector size. Setting both to 64 samples, I got a system latency I felt to be acceptable. When playing fast passages, large buffer sizes such as 512 samples or above were found to be problematic in terms of a precise timing.

into the tone. The sound seems to lack the potential for modification in its inner structure, in other words, the potential provided for creating a larger or smaller, intense or less intense tone is somewhat limited. When playing at a moderately loud volume, the sound feels interesting and modifiable in timbre.

Sound no. 2 (signal-driven simple FM synthesis, ratio = 0.5) offers a similarly good playability to sound no. 1. It shows a reduction in pitch stability when playing heavily with consonants, such as in scratching, for example. The above-mentioned bowings also work well. Latency is slightly reduced, and playing large intervals can cause an uneven timbre at the beginning of the tone until the pitch tracker has found the current fundamental frequency. Vibrato is reduced in terms of intensity as compared to the traditional instrument. Double stops are absolutely impossible. It is easier to get into the tone, and the sound seems more open and warmer, while still retaining the possibility of playing very harshly. Notes on the G and C strings can sound like bass clarinet sounds and are more modifiable than the tones of sound no. 1. Compared to sound no. 1, the sound in general feels more interesting, has more possibilities to explore, and is simply more fun to play.

Sound no. 3 (signal-driven subtractive synthesis) provides the best potential for modification with the bow and the left-hand fingers in the steady state of a tone. Notes on the D string around g' allow the player to get into the tone and explore it and form it to a surprising degree. The tone can be made very soft and delicate, and also very dirty. The feeling can be described as being very close to the tone or having it very directly in one's hands. Bow speed, position and pressure affect the sound very satisfactorily. Unfortunately, the timbre balance between high and low pitches is uneven. In addition, if the pitch detector cannot find the current pitch played and starts to jump quickly from value to value, the timbre is heavily damaged. When playing this sound preset in concerts, I have found that I have to adapt very quickly to the shortcomings in the consonant phase of tones, and have started to play notes with reduced consonants. In the steady state, the sound feels very interesting and invites exploration. The adaptation in the consonant phase reduces the playability. However, as mentioned earlier, it may be impossible to eliminate adaptation entirely while still maintaining the goal of increased possibilities. In my experience, the possibilities the sound offers during the steady state outbalance the reduced playability in the transient phase by far.

Sound no. 4 (signal-driven multi-modulator FM synthesis) offers the highest potential for sounding different from the pickup audio signal. The sound offers an interesting but small field to be explored. The attack phase of the tone played seems to be cut off, which means that the sound is bad for articulation. The latency is close to that of a synthesised sound driven by features extracted from the audio signal. Different bowings are brought out poorly. The pitch reaction, including with vibrato, is acceptable. It is interesting to study how the timbre can change in relation to volume, bow speed, position and pressure. The ability to modify tone, timbre, intensity and character, however, feels restricted.

Sound no. 5 (granularised sound) is different from the other sounds in that latency is always present in all aspects of playing. Depending on the actions applied to the input side, the sound seems to be somehow unclear or washed out. Timbre changes at the input result in timbre changes at the output, but not always adequately so in relation to the desired results. The pitch of the sound has changed, thus the parameter of pitch has not “survived” in this preset. The tone feels very close at hand and very malleable. With regard to the character, there is an amusing nervousness in the tone. I returned again and again to this character. Double stops are possible. Spiccato feels very washed out in the sound result. In my view, the sound is potentially interesting for somewhat funny or ironic passages.

Considering these specific sounds and my experience with ASDSS sounds in general, my impression is that one achieves a range of sounds falling within a specific “sound area” distinct from the sound variety of known synthesis possibilities. This area is defined by the interface and the input coming from the player. In other words, the sounds often have a string-like shape. It is logical that a string player wanting to achieve a sound result close to what she is inputting will not create a sound like a crashing car, for example. The string-specific playability does not come without cost. Since the goal is for the sound results to be coupled adequately to the input of a string player, the playability comes at the cost of a reduction in sounds that stay within the general framework of string players’ sounds, shapes, envelopes and dynamics.

While testing and refining sounds, I became aware that it was possible to create sounds offering good playability, timbre and potential to be explored. Sounds that felt close at hand and seemed to offer a large potential for exploration provoked my interest most, because I was able to find qualities in subtle variance and timbre of sound which I had not found in the parameter-driven string synthesisers I had played so far. However, I must state clearly that the goal of creating sounds with a similar timbre and potential for variances, like the sounds of the FM8, for example, has not been achieved. On the other hand, ASDSS sounds worked very well in concerts (see section 7.5), the basic construction principle of ASDSS is very open, and other researchers have already used it to develop their own methods (see section 6.3). Thus, I see good reason to hope that a broader range of sounds might become available through further development.

7.2 String Players Testing the Instrument

What do other players say when they play with or test the system? Whenever I had the opportunity, I presented and distributed the system to string players. This was done, for example, at several talks on musical interfaces. The instrument was offered for testing afterwards. Additionally to this, it was tested in informal meetings with string players and composers, and in the group hot_strings SIG (this group will be described later in section 7.6).

The main questions were as follows:

- What do players do when they test the instrument?
- What was interesting and important to the players?
- What did players say?

- Were there any things that the majority of players said?

One of my observations was that many performers initially try the instrument as they normally would when testing a traditional instrument. This means that they play some scales, arpeggios and different bowing techniques, followed by some passages from their (mostly classical) repertoire. They test how the instrument reacts in relation to what they put into it. As it becomes clear that the instrument reacts slightly differently, they start to explore the differences. This was often done by playing within a range where the instrument reacts as expected, and then moving gradually into a playing style where they expected that the reaction would not be adequate. Examples of this are playing higher and higher, or playing slow passages and repeating them faster and faster. In addition, precise bowings and articulations were often made less and less precise in order to see what would happen. Or the instrument was treated very roughly with martelé bowings in fortissimo in order to see how the instrument would react in an extreme situation.

The DVD in Appendix E.2 E.2 contains videos (chapters 6 - 9) documenting string players testing an ASDSS instrument. In chapter 6 of the DVD one can see particularly well how the player Gerardo Vitale tests the instrument's behaviour at the boundaries of its precise reactions. In chapter 7 of the DVD, Mari Kimura (2010), who is well known for performing with violin and electronics, can be seen testing my instrument.

In summary my observations are that a player first tests how the core of the instrument and a selected sound feels; then variances of this core are tested, followed by tests to establish where the boundaries of the core lie and how one can deal with them and the instrument's behaviour at them. I conclude that both the zone of regular sounds and the boundary with non-regular sounds, as well as the properties of both, are seen as important factors for the performers who have tested the developments.

In most cases I had to press the musicians for statements about the instruments, because many performers had little experience of bowed stringed synthesisers and were therefore uncertain about an estimation. For this reason I am able to present only a small collection of statements in response to my request to hear their perceptions of the instrument. Test subjects often mentioned that they felt the instrument was too heavy. With regard to playability and sound, different opinions were heard. While some players with a more traditional approach mentioned that the sound should be closer to a traditional instrument, performers who were already used to electronic stringed instruments were quite satisfied with sound and playability. The following collection of statements may give an impression of the different opinions found (Poepel, 2008a, p. 185):

- "This does not at all feel like a normal violin since I am playing here (points to the Harms viola) and the sound is coming from there (points to the loudspeaker)."
- "The system is impressive since it is such a simple idea but with such a huge effect on playability."
- "It is nice but it makes no sense for me to play a bass clarinet-like sound or a pan flute on the violin."
- "Such an instrument will never feel like a traditional instrument. The feedback loop between the bridge and body does not exist. Therefore the oscillation of the body will not affect the bridge and the string."

One of my research goals was to develop sounds which were felt by target users to be usable and interesting. My assumption was that the users would complain about the drawback of ASDSS: the fact that conventional methods of sound synthesis cannot be used and the possible sound variances might therefore be considered far too small. However, this was not the case.

I would like to present one example of sounds being designed in a user-centred way (Beyerle, Heineken & Zupková, personal communication, December 5, 2004). The performers and I began with a combination of signal-driven FM and signal-driven subtractive synthesis. The performers found it to be a problem that the speaker was not mounted directly on the instrument so as to be able to hear the sound directly, as is the case with a normal instrument. The sound was also felt to be inadequately characteristic of a string sound, and to sound too much like a wind instrument. In addition, it was mentioned that the sound was lacking in warmth. For these reasons, Hatto Beyerle and Babette Heineken were convinced that the sound would not allow the communication of musical ideas in the way to which string players are used. In order to make music in a string-specific way, they felt that the sound should be more string specific. The following statement was made by Hatto Beyerle: "Beside [*sic*] the importance of a string-specific playability it is important to have a string-specific sound. The question of the playability or feel of instrument cannot be separated completely from the sound" (Poepel & Marx, 2007, p. 283).

In order to offer the performers as many sound options as possible in this user-centred sound design project, I needed to improve the system. First, the position of the pickup was modified to achieve a sound which was warmer, natural and round, but less even between different strings. Then the bandwidth of the band-pass filters, which generate the partials, was opened. After this, a small portion of two versions of parameter-driven simple FM was added. Finally, a small amount of single sideband modulation was included.

My assessment was that the sound was now closer to the pickup signal and was likely to be perceived as uninteresting because of its similarity to the audio signal. The performers, however, did not think that the sound was uninteresting at all. They perceived the sound far away from the pickup signal or the traditional instrument. While I viewed the ASDSS sounds developed so far as rather similar, the performers felt the differences between these sounds to be very large. Thus, it became clear here that the identity criteria of a sound and the quality criteria for an interesting sound differed between the performers and me.

I had a different experience when working with the violinist and composer Günter Marx, who has been playing electronic string instruments for more than 25 years. While he perceived the playability of ASDSS sounds to be increased as compared to MIDI sounds performed with his Zeta MIDI violin, it became clear that he had already adapted to the formally constrained and more limited playability of the Zeta MIDI violin. A sound such as that developed with the string players above was perceived as too close to a violin sound. Because of his experience with electronic sounds, his quality criteria focused more on the question of how different an ASDSS sound is from a traditional sound, and how variably the sound can react in response to the input. One interesting example is that of a sound with an inharmonic spectrum in which the inharmonicity is influenced by the amplitude of the input signal. Another example of such a sound can be found on the accompanying DVD of Appendix E.2 (in chapter 11), where the question of the playability of ASDSS sounds is discussed and demonstrated with such a sound.

The feedback from the users can be used to improve the instrument, especially in reaction to the requirements document as suggested in figure 3.2 of section 3.8. Besides obvious considerations such as the visual user interface of the software, weight of instrument and sound source close to the ear, the requirement of equipping a performer for string-specific communication of musical ideas can be of importance. Extensions of the sound repertoire require different sounds from that of traditional strings. This raises the question of the extent to which a sound can differ from the instrument's audio signal while still preserving the methods of communication to which string players are accustomed. Conversely, there is the question of how similar a sound can be to the instrument's audio signal without losing its attractiveness to the performer because it is *too* similar to the instrument's audio signal. On the basis of my experience, which I gained by loaning my ASDSS instruments to performers and asking them about the usefulness of the sounds, I believe that the answer to the difference-similarity question depends on the individual player and his previous experiences and preferences, and therefore cannot be answered in a general way.

One important experience is that a performer may judge a synthesised sound to be interesting because of its high ability to communicate musical ideas, despite its—from my point of view—relatively close similarity to the instrument's audio signal (see the example of sound design described above on p. 240). Another point I regard as interesting is the fact that the ability to communicate was considered dependent on both the playability *and* the sound. Thus, when it comes to optimising the musical (tangible) interface, the question concerning the relevant actions of the performer and how they can be tracked appears not to be sufficient on its own. It may be of use to look additionally at the question of the conditions which a sound must meet in order to allow—via the actions of the performer—the communication of the musical ideas of string players. To my knowledge this question has not yet been addressed in the research field of new interfaces and computer-based instruments for string players. However, since the goal of communicating in a familiar way was mentioned as important by some test subjects, it may be of use to add it to the requirements of a bowed stringed synthesiser.

7.3 Empirical Study with String Players

As mentioned before, besides loaning the instrument to performers and asking for their subjective assessments, I was also interested in scientific methods of evaluating instruments. In analysing related research, I have seen that very few developments were presented with empirical evaluations incorporating target users (see section 4.7). One reason may be that it is difficult to get string players to test the system in a formal way without being paid the fees musicians would usually receive for a rehearsal. Another reason might be that developers did not consider the possibility of different results from the ones they themselves expected. In addition, the lack of a tradition or method for the evaluation of new instruments may be viewed as a reason for a rather reluctant attitude to participation in such studies. Though established methods for evaluation are rare, there are some publications presenting these. Wanderley and Orio (2002) use tools from HCI (human-computer interaction) to evaluate musical interfaces. An example which applies this method in an extended form can be found in a study by Isaacs (2003) in which a Korg Kaosspad KP 2 and an accelerometer used for the production of sound are evaluated.

In this section an empirical study is presented. My background in the design of empirical studies is based on what I learned while studying Audio Design, and on the experience I gained with previous studies. In the present context I used the models and approach to empirical studies proposed by Beller (2004). The main question, the questionnaire, as well as the test setup were designed by me.[2]

In the study, performers compare three bowed stringed synthesiser instruments on the basis of pre-established questions, and assess how well the

[2] In order to get outside opinions on the design of my study, I discussed the study with several people, for example musicologist Dr. Jin Hyun Kim, lecturer in systematic musicology and media science (she also was conducting research into musical interfaces) at the University of Cologne, Germany; with electrical engineer Dr. Eric Lee (at that time PhD student in media computing at RWTH Aachen University, Germany, developing and evaluating musical interfaces); and with computer scientist Prof. Dr. Jan Oliver Borchers (my second supervisor), head of the Media Computing Group at RWTH Aachen University. Sample size and methods of data analysis in particular were discussed with Dr. Eric Lee and psychologist Dr. Carolin Demuth, at that time PhD student in psychology at the University of Osnabrück, Germany. The significance of results was discussed with Dr. Carsten Röcker, psychologist and postdoctoral researcher at the Media Computing Group at RWTH Aachen University, Germany.

instruments behave in relation to the features detailed in those questions. Two of the instruments are based on PDSS and one uses the method of ASDSS. The study was conducted in January 2005 and was published in a peer-reviewed paper (Poepel, 2005).

7.3.1 Question

In this study, the player-assessed potential for musical expression of three instruments is compared. My question is whether a significant difference in the assessment of the potential for musical expression can be found when comparing one example of the signal-driven approach to one example of the parameter-driven approach. In addition, I ask about the performers' perception of differences between continuous and note-on triggered (MIDI-driven) control of a synthesised sound.

7.3.2 Premises

If students at a music academy are asked whether they want to be expressive, the majority answers "yes". Most of them mention musical expression as one of their primary goals (Juslin, 2003). Thus, for the target users of my instruments, musical expression can be seen as an important goal in the performance of music. For this study I borrowed knowledge about musical expression from the field of music psychology. I refer here to music psychology because the questions of what can be understood by musical expression and how musicians create musical expression is investigated there on a scientific basis. The findings in the literature I cite are based on studies with traditional musicians. I therefore expect a considerable overlap between the test subjects and the target users.

Juslin (2003) mentions several factors influencing musical expression (p. 278). One of these is the instrument. Others are the ways musicians perform, the ways listeners perceive, the context in which a performance is presented, and of course the way in which the piece of music has been composed. According to Juslin (2001, 2003), musical expression is realised by communicating emotion via music. This is done by coding expressive intentions, for which specific cues are used. The cues include tempo, sound level, timing, intonation, articulation, timbre, vibrato, tone attacks, tone decays and pauses (Juslin, 2001, p. 316).

As I am investigating an instrument's potential for musical expression, I focus on the instrument-related (Juslin, 2003, p. 278) and performer-related (Juslin, 2001, p. 361) factors relevant to the model of musical expression. In the present context I regard the following cues as useful to the question of an instrument's potential for musical expression: pitch, dynamics, timbre, articulation and timing. Given the fact that the newly designed instruments are intended to make use of the existing skills of a string player (the specific target users), it is desirable that they should have an input-output accuracy that is as high as possible in relation to the actions performed to generate these cues. Therefore, the new instruments must be tested with particular attention to the playing methods performers use to produce the cues.

7.3.3 Design of the Study

The first task was to select instruments for comparison. The criteria for this selection were comparability and availability for the study. In order to have a hardware basis that—to ensure comparability—was as similar as possible, instruments which extracted parameter data from the audio signal were selected.[3] I chose the following three instruments:

1. Instrument A: Zeta MIDI viola bridge, mounted on a traditional viola (see figure 7.4, p. 264), MIDI synthesiser Kawai K5000, program A044;
2. Instrument B: Harms electric viola (see figure 6.15), pitch and amplitude tracking, mapping, parameter-driven simple FM synthesis;
3. Instrument C: the same electric viola, pitch and amplitude tracking, mapping, signal-driven FM synthesis.

Instruments B and C use simple FM synthesis, the first one based on the parameter-driven approach and the second on the signal-driven approach. To provide a comparability between the signal- and the parameter-driven instruments, a sound was chosen which equals the known latency problems

[3]Some parameter-driven synthesis instruments make use of gesture tracking methods. One might argue that a comparison of signal-driven synthesis with parameter-driven synthesis should incorporate parameter-driven synthesis making use of gesture tracking. It is, however, possible to extend the signal-driven approach to incorporate gesture tracking data in order to modify the sound indirectly as well. Since tracking data was not used in the signal-driven approaches developed here, and furthermore the test needs to use—for reasons of comparability—a comparably complex technical framework, the tests were conducted only with instruments deriving parameters from the audio signal.

of the Zeta MIDI system (Yoo & Fujinaga, 1998, p. 13) and which is based on an FM-like timbre but is slightly adjusted to a string timbre. All three instruments use a pitch tracker, an envelope follower and a synthesis unit. In order to be able to distinguish more easily between the acoustic sound of the viola's wooden body and the synthesised sound, parameters are set to transpose the sound result one octave lower than the sound performed on the interface.

The theoretical construct I am concerned with is "the instrument's potential for musical expression." In order to make it measurable, the construct must be operationalised (Beller, 2004, pp. 29-30). The process used is similar to the method of moving from the requirements to the specification level as described in section 3.4. As I am dealing with models and thus abstractions of the production of musical expression, I am aware of the fact that these abstractions will necessarily generate a degree of blindness, and that the models used include the three features mentioned by Stachowiak (1973, p. 131) and described in section 3.6. However, since I am seeking to use a scientifically accepted method for empirical work, I accept these shortcomings. I would like to state clearly that the results obtained can only be seen as valid as long as one remains within the pragmatic feature and the features of mapping and reduction which correspond to the model of musical expression used in my study.

One might argue that the reductive parameterisation or the coding of expression which Juslin uses in his model is compromising, since the model itself is perhaps adequate only to a certain degree, and may be more differentiated and expanded by future research. However, in order to work on a scientific basis it is necessary to use scientifically accepted models such as are available at the time. Since the model of Juslin is widely accepted, the use of this model is seen to be adequate.

Operationalisation must define the indicators which are to provide measurement data. On the basis of the models of musical expression described above, the indicators in this research are to be found in the field of performer-related skills and the assessed instrument response in relation to the cues of pitch, dynamics, timbre, articulation and timing. I therefore conclude that the indicators lie in pitch accuracy, dynamics accuracy, timbre accuracy, articulation accuracy and timing accuracy. What is measured is the participant's assessment of the instrument's accuracy with respect to its response to actions performed by the participant to generate these cues. More specifically, the following actions are seen as relevant (playing techniques are given in brackets) (Poepel, 2005, p. 229):

1. timing accuracy: tempo, timing, pauses (e.g. pizzicato, collé, spiccato, short notes);
2. pitch accuracy: intonation, vibrato (e.g. different notes with pauses, legato tones, glissando, different vibratos);
3. dynamics accuracy: sound level (e.g. crescendo, decrescendo, pp, mf, ff, sfz);
4. articulation accuracy: articulation, tone attacks, tone decays (e.g. détaché, martelé, spiccato, scratching);

5. timbre accuracy: timbre (e.g. digging into the string, pulling the sound with the bow, variation of bow-bridge distance).

A questionnaire was constructed which first asks about the background of participants in order to be able to analyse the measurement data with respect to the predisposition of participants. Factors asked for include gender, age, professional or amateur status, experience with electronic musical instruments, music preference before or after 1950, music performed mainly before or after 1950, interest in electronic devices and interest in electronic sounds.[4] After this the questionnaire asked about the cue accuracies and, on a more general level, about the instrument's transparency for musical expression when playing a given musical phrase. 18 tasks were given to be played with each of the three instruments, in order to compare them. 13 tasks were related to the cues, and 5 to musical phrases. Each task was followed by one of the following three questions:

- Is the instrument transparent in relation to the playing method used?
- Is the relationship between gesture and electronic sound adequate?
- Is the intended sound result well represented in the perceived sound result?

A five-point scale (1: not usable, 5: very good) was used and the participants were asked to evaluate every instrument in relation to every task.

[4]These factors were selected on the basis of my experience with factors I had used to cluster results in an unpublished study on musical interfaces I had conducted previously. The selected factors were the ones I found to be of use.

Once the 18 tasks were complete, the participants were asked in question 19 to evaluate whether they thought the instrument was useful for musical expression, and in question 20 whether they found it was useful for their individual personal expression. The design of the study included an interview on the participants' experiences with the instrument.

The questionnaire can be found in Appendix A.1. It was hypothesised that instrument A and B would lie closely together because they are both parameter driven, and that instrument C would differ from them because it uses the signal-driven approach.

7.3.4 Procedure

The duration of each test was about 90 minutes. First, the participants were introduced to the test and to the instruments (20 minutes). After that, the tasks were undertaken, and after each task the ratings were made (60 minutes). Finally, the interview on experiences with the instrument was conducted (10 minutes).

13 trained string players, 7 women and 6 men, took part in the study. Among the participants were 8 professionals (earning their living with the instrument) and 5 amateur musicians. The musicians came from professional orchestras in Cologne, from the Jazz-Violin class of Cologne Music University and from the Cologne and Bonn amateur music scene.

The tests were conducted in the author's office at the Academy of Media Arts in Cologne. The synthesised sound was projected by two Adam P-11 active monitor speakers, placed at a distance of approximately one meter from the participants.

7.3.5 Analysis

In order to analyse the data, the mean values of the answers to the 20 questions were calculated. In addition, answers were put into groups according to the factors asked about at the beginning of the questionnaire. The mean values of the ratings were calculated for the following subgroups: women and men, professionals and amateurs, musical preference before or after 1950, and more or less interested in electronic sounds.

To verify the difference between the three instruments, an analysis of variance was done (ANOVA). A significant difference between the three was found. This was indicated by the ANOVA values of $F = 29.76$, critical $F = 3.26$ and $p < 0.01$. Validity was assessed by comparing the mean of the 18 tasks with the mean of questions 19 and 20 which questioned whether the participants found the instrument to be usable for musical expression (see table 7.1).

	18 tasks (mean)	perceived expressivity (mean)
Instrument A	1.90	1.77
Instrument B	3.59	3.77
Instrument C	3.83	4.00

Table 7.1: Values to estimate validity.

In order to make reliability testable, cue-related tasks were run twice. In order to calculate reliability, the correlation of the twice-run tasks was calculated. The correlation lies between 0.40 and 0.74 with an average of 0.58. On the basis of these results one must acknowledge that reliability in some tasks was low. Therefore, and because of the small number of only 13 test subjects, the values gained from this study and presented in the following section can be seen as possible tendencies, but not as exact results concerning the potential for musical expression of the tested instruments. The interviews were analysed by looking at what participants said, how this may be relevant to the question of musical expression, and whether there were statements which were found more than once.

7.3.6 Results

As presented in figure 7.1, the overall results of all 18 tasks (questions 1 to 18) show that the MIDI viola was found by the 13 participants (in black) to have a much smaller potential for musical expression than both FM violas, while the signal-driven and parameter-driven FM synthesis violas were ranked close together with a slight preference for the parameter-driven one. Four participants evaluated the potential of instrument C for musical expression (signal-driven) to be the highest, while two participants evaluated that of instrument B (parameter-driven) as highest. The potential of both instruments was rated equally by seven participants.

An interesting tendency was that test subjects with a low interest in electronic sounds assessed the instruments differently from those with a high interest. As can be seen in figure 7.1, the former (green) assessed the difference between instruments B and C to be greater than the latter (blue).

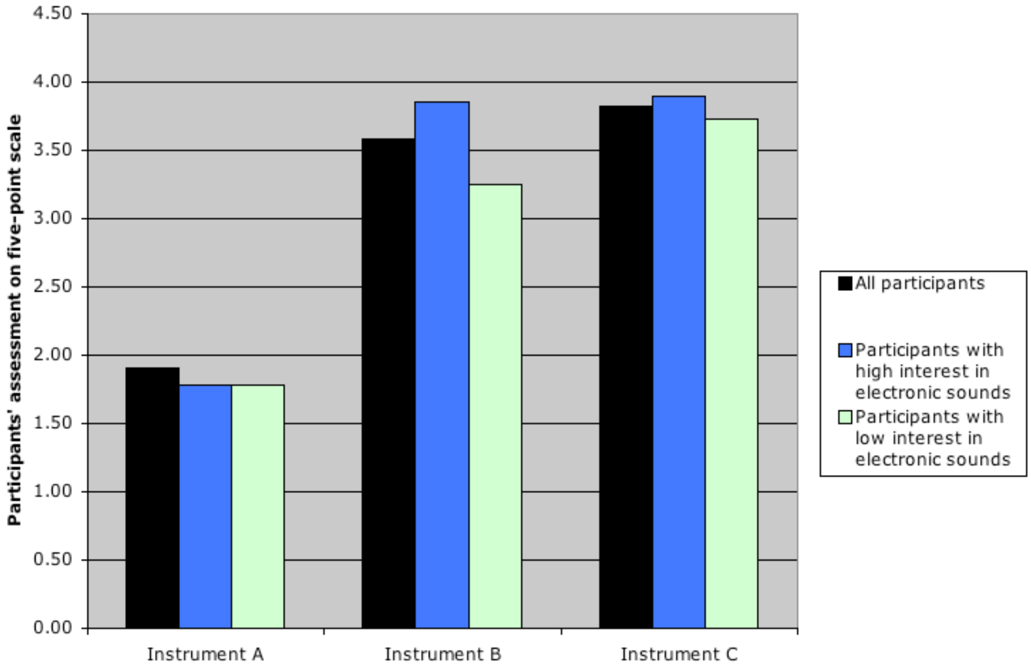

Figure 7.1: Overview of instrument ratings by participants. Level of interest in electronic sounds can influence the perception of sounds.

Analysing the tendencies in assessment concerning instruments B and C in relation to the expressive cues (pitch, timing, dynamics, articulation and timbre), it is interesting that participants assessed dynamic accuracy as differing most significantly, as presented in figure 7.2.

Figure 7.3 presents the calculated mean values of the groups selected according to the background factors of participants. An issue that might be worth further investigation is the relatively high rating for the Zeta MIDI viola by women. It would be interesting to find out whether a larger sample of participants would show a significantly different perception of the Zeta MIDI viola by women and men.

In the interview, five participants mentioned that instrument C provided a

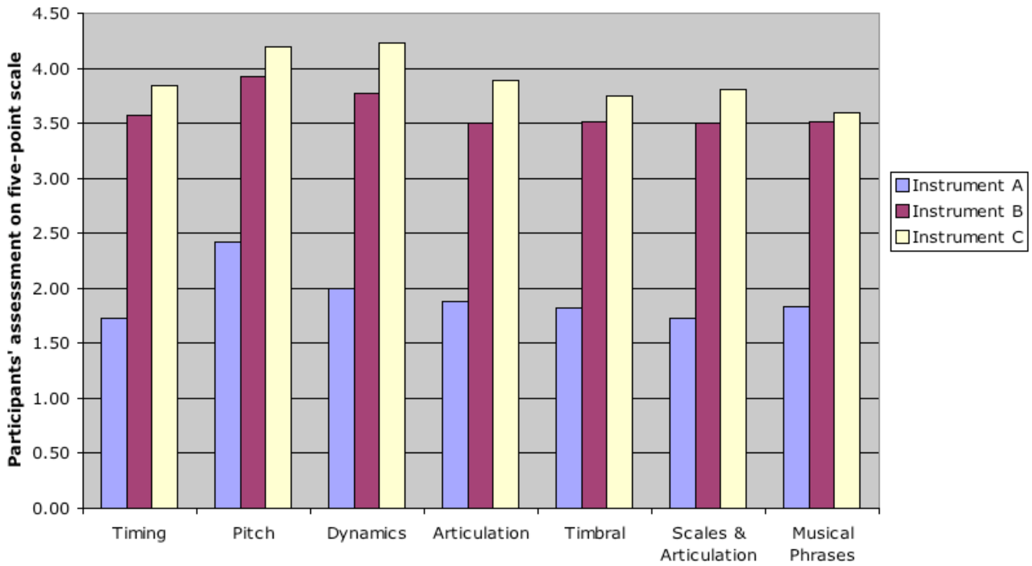

Figure 7.2: Differences in cue groups.

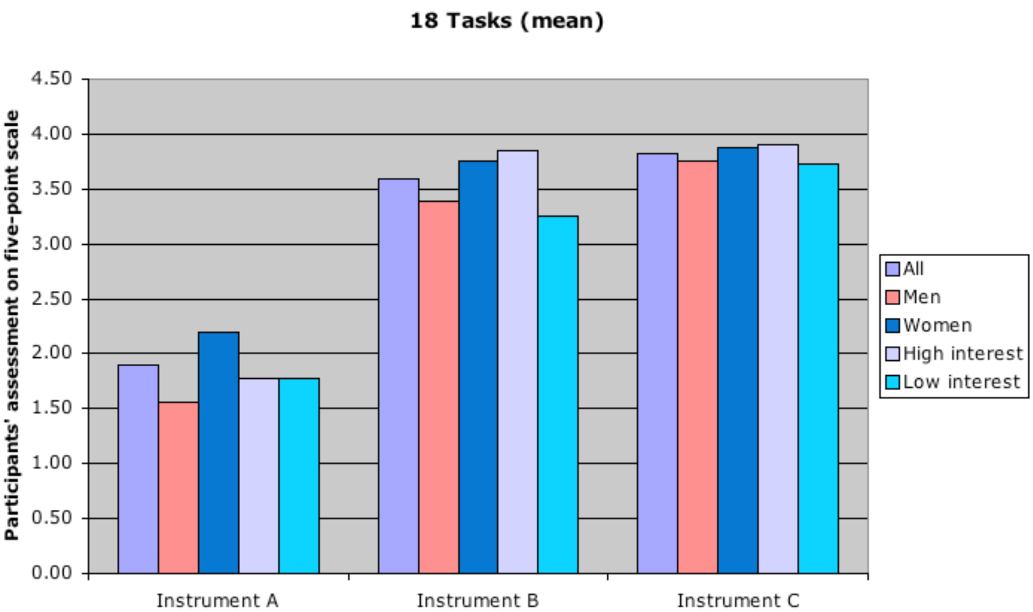

Figure 7.3: Comparison of selected background factors.

higher accuracy with regard to instrument-specific playing techniques. Other statements by the participants on instrument C were:

- I can play the bowing technique collé more effectively.
- I can go into the tone more effectively.
- I feel that the tone is more effectively in my grip.
- I can pull the tone more effectively.
- I can play legato between different tones more effectively, the instrument is more transparent for my musical expression.

The greater weight of the Harms viola as compared to a traditional viola was often mentioned to be unpleasant. With regard to the reaction of the instrument in fast spiccato passages, some participants preferred instrument B, whereas others preferred instrument C. Four participants found that instrument C reacted better than instrument B in terms of vibrato. Four participants found that instruments B and C were lacking a sufficient equal balance of timbre in relation to low and high notes or strings. Four performers criticised the "bubbling" effect (caused by fast pitch-tracker jumps) when playing very noisily or close to the bridge (sul ponticello). Two participants with a high level of interest in electronic sounds were not quite satisfied with the sound result of instrument C. They stated that they had expected the kind of synthesised sounds to which they were used. In fact, the sound of instrument C was perceived to lie between that of a string instrument and a synthesizer. The lack of and need to play double stops was mentioned by two performers, and two participants spoke about the need to have the system not in a laptop but in a mechanically stable 19-inch rack-mount unit in order to be able to use it easily on stage.

While it was hypothesised that instruments A and B would produce more similar results than instrument C, the results do not confirm this assumption. On the basis of the results data, one might conclude that instrument C is assessed as having the highest potential. However, the low reliability in some tasks and the relatively small number of participants mean that the small difference cannot be used here to talk about a significantly greater potential.

It may be asked why the results failed to show a significant difference between signal-driven synthesis and parameter-driven methods. I assume that in signal-driven synthesis the quality of openness will only emerge over a longer period of time. One might hypothesise that the potential for exploration was not recognised by participants because of the short amount of time they had with the instrument. A statement by one participant appears to underline this view. Demonstrating in the interview what he had learned about the sounds, he spoke about the ability of the parameter-driven FM synthesis (instrument B) to react to varying bow positions. As bow position is not tracked here and thus does not influence the sound, it became clear that the fine nuances of reaction I was interested in could not necessarily be experienced within the time the participant was working with the instrument in this study.

While I did learn a great deal from the results and the statements of players in this study, I did not meet the goal of answering my initial question in a clear and precise way. I therefore decided to conduct a second study with a larger group of participants and a more reliable measurement method. This study is presented in the following section.

7.4 Empirical Study with Listeners

In order to be able to work with a larger group of participants, I decided to conduct a listener-based study. To plan and carry out the study I collaborated with the audio and video engineering degree course at the University of Applied Sciences in Düsseldorf. The cooperation was organised in the context of a dissertation written by student Thomas Gwosdz (2007) and supervised by Dieter Braun and me.

7.4.1 Question

The primary question to be answered was set by me. As I was not planning a study in which performers would assess the qualities of an instrument, I could not ask about an instrument's potential for musical expression in relation to individually varying inputs given by performers. Instead, I asked about differences in the output of different synthesis methods when applying the same input signal. Accordingly, the question was:

Are there significant differences in musical expression for a listener when comparing parameter-driven and signal-driven synthesis methods using the same input signal?

7.4.2 Premises

With regard to the question of musical expression, I used the same basic assumptions from music psychology described in section 7.3.2. In addition, I used the experiences of listening tests gained in the audio-video engineering degree course at the University of Applied Sciences in Düsseldorf. According to these experiences,[5] the study should ideally take approximately 45 minutes, or slightly longer if necessary. With longer studies it was found that the concentration of the listener decreases.

As instruments need to be compared, Dieter Braun suggested using a paired comparison. With each task, every instrument was to be compared to every other instrument. When comparing instrument A to instrument B, the comparison needed not only to provide the A-B transition, but also the B-A transition. For this reason, the structure A-B-A-B was selected. The duration of a sound example should be no longer than 15 seconds in order not to lose the focus on the general character of a sound, and no shorter than seven seconds in order to gain a long enough impression and to validate the sound example.

7.4.3 Design of Study

With regard to the selection of the instruments, I did not want to use the Zeta MIDI viola because of my conviction (based on the results presented

[5] Dieter Braun is a professor of acoustics at the University of Applied Sciences in Düsseldorf and has been supervising empirical audio studies there for more than 15 years.

in section 7.3.6) that this instrument would be rated very low. As there was a specific audio signal as the source for the test, this signal was used as the first instrument. In order to compare parameter-driven synthesis with signal-driven synthesis, I again selected the parameter-driven FM synthesis as described in section 6.2.3 and the signal-driven FM synthesis (see section 6.6.1), but this time in the same octave as the audio signal, in order to facilitate comparison. As I was not only interested in evaluating one signal-driven method, I added the method of signal-driven subtractive synthesis (section 6.6.2).

The following four instruments were selected for comparison:

1. Instrument A: pickup signal (Zeta Jazz Series pickup, no use of any tracking method),
2. Instrument B: signal-driven FM synthesis, (ratio = 0, thus no use of pitch follower, use of envelope follower),
3. Instrument C: parameter-driven simple FM synthesis (use of pitch and envelope follower),
4. Instrument D: signal-driven subtractive synthesis (use of pitch follower, no use of envelope follower).

Though I could have set the ratio of the signal-driven FM synthesis (instrument B) to one, I chose not to do so because I wanted to compare one signal-driven method using an envelope follower and one using a pitch follower, to see whether there would be perceivable differences, and if yes, how strong these differences in perception would be.

Another question was that of the audio signal fed into the synthesis methods. It is clear that a performer will work with the audio feedback coming from the synthesis engine and speakers, and that the playing method will be adapted immediately to the specifics of the sound result. To retain this feedback, it would have been necessary to play the same musical phrase with each of the synthesis methods separately. This, in turn, would have caused the problem of different signals having been applied. Thus, the need for equality in comparison would have been negatively affected. It is impossible to achieve both a regular feedback loop from the instrument to the player, adjusting the input to the output, and an equal input guaranteeing equality of comparison. As I was aiming here to conduct a properly scientific study, I preferred to use the same source for all synthesis methods, and accepted the problem of a limited feedback loop between player and sound.

As I was again investigating the issue of musical expression with regard to a specific sound, I used the same models and understandings of musical expression as I had done in the player-based study. The theoretical construct in this case is the "musical expression perceived when applying the same input to different synthesis methods". To develop the indicators I again used the cues relevant to the coding of musically expressive ideas, which are pitch, dynamics, timbre, articulation and timing. Thus, the indicators were the assessed ability of the synthesis method to represent the cues as compared with another synthesis method.

Comparing four instruments to one another requires six comparisons (A-B, A-C, A-D, B-C, B-D, C-D). In order to calculate reliability I chose to carry out one comparison twice. To determine validity, one synthesis method was compared with itself. This meant a sum of eight comparisons for one task. Calculating that an average test sound would take 11 seconds, followed by a break of 0.5 seconds, meant that one comparison would take 4 x (11 seconds + 0.5 seconds) = 46 seconds. Then the participants would need a time slot of five seconds in which to record their assessment and to make a comment. Doing this eight times would mean that one block of comparisons would take on average six minutes and 48 seconds (408 seconds in total). Adding a break to relax and to regain concentration would mean another 30 seconds. Keeping within the time frame of approximately 45 minutes, eight blocks of comparison could be planned.

One parameter relating to each of the five cues under investigation was determined for five of the eight blocks. These were:

1. articulation: clearness (martelé bowing),
2. pitch: vibrato (intensity),
3. dynamics: crescendo – decrescendo,
4. timbre: change of timbre (changing bow position to morph from a soft to a harsh sound),
5. timing: rhythmical precision (given rhythmical passage).

The sixth block asked about the musical expression of a given phrase. Two blocks were left. I used the first to ask about the continuity of a legato in order to investigate the effect of the pitch tracker in legato. The second block was used to ask again about precision in spiccato passages, as the topic was mentioned in the interviews of the player-based study, though no clear preference for a specific instrument was found (see section 7.3.6). A ninth block was added which was not intended to be used. In Dieter Braun's experience (personal communication, March, 2007), the concentration of the listeners could be sustained more successfully in the eighth block if a ninth was due. The intention was to tell the participants that block nine would be skipped once block eight was finished.

An example of a question asked in a two-sided comparison (A-B) is: in which of the two sound examples is the rhythmical precision better represented? The questions were asked in a way which made it possible to use the ordinal measuring level. Ordinal measurement enables the arrangement of a rank order of the synthesis methods in relation to a specific question within a block. It does not, however, provide information on how large the differences between the objects are.

The questionnaire did not refer to instrument A, B, C or D. Since even these letters may influence the participants, the two symbols diamond and square were used. A five-point scale was again selected. Asking about intensity in a comparison of two sound examples marked with diamond and square, the answers on the scale would be:

1. Diamond is much more intense than square.
2. Diamond is slightly more intense than square.

3. Diamond and square are equal.
4. Square is slightly more intense than diamond.
5. Square is much more intense than diamond.

The questionnaire consisted of a cover sheet and nine pages. The cover sheet asked for the participant's profession, age, gender, and whether electronic music was listened to often. The nine pages each covered one block. Each block started with a description of the parameter to which it was related, followed by a short explanation of what was being investigated in the block. After that, a definition was given of the parameter which was to be asked about. Finally the question was asked. The participants had to rate each comparison by marking one of the five boxes. In addition, they had the option of providing comments about the sounds in a separate box. The questionnaire can be found in section B.1 of the Appendix.

The sound files were created by first recording and selecting musical phrases I found to be usable for each block. The instrument used to generate the sounds was a traditional viola equipped with a Zeta pickup from the Zeta Jazz Series (see figure 7.4). The sound files were then applied to the selected synthesis methods and the newly generated sounds were recorded. After this, a random sequence of the defined pairs was generated for each block. The sound files were put into the order of this sequence, each one followed by a pause of 300 ms. Each row had the order diamond – square – diamond – square, and after each one a pause of five seconds was set. A CD track was created for each block, and a CD was created which included all block tracks. The CD tracks can be found on CD 2, tracks 1 to 8, in Appendix E.1.

Figure 7.4: Instrument used to record the sound files.

It was planned to carry out the listeners study in many different places, in order to achieve a high number of participants. The study was therefore conducted using headphones to ensure an identical listening situation in each of the different places. The headphones used were Stax SRS 2020 electrostatic headphones, which offer extremely precise sound reproduction. A common commercial CD player was used, and plugged directly to the headphone amplifier.

It was hypothesised that instrument C (parameter-driven) would be rated lower than the others, and that instrument A (pickup signal) would be rated highest. It was expected that instruments B and D would be somewhere in between.

7.4.4 Procedure

The procedure was planned by me, and the practical work with the participants was carried out by Thomas Gwosdz. 43 participants completed the study. The average age was 27.5 and the range of age was 14 to 66. Among the participants were 30 men and 13 women. 35 participants had experience of and were able to play a musical instrument, and 29 stated that they listen to electronic music. The musical professionals were: 15 students of audio/video engineering, 4 professional musicians, 3 school pupils. The occupations of the other participants were office assistant, secretary, chemist, businessman, care worker and two pensioners.

After an introductory explanation of the procedure of the study, the participants were asked to fill in the first page of the questionnaire. Then the questionnaire was explained and it was asked whether the definitions of the parameters used were understood and clear. The timbre change via bow position (bow close to bridge, bow over the fingerboard) required the most explanation. Then the headphones were put on and track 1 was started and left running without interruption until it came to an end. If further questions arose they were answered in between the blocks.

Because of the limited number of headphones, the maximum number of participants in one cycle was four. Including the introductory explanation and questions, a cycle took 75 minutes on average. The music examples only took 47.6 minutes. Figure 7.5 presents a typical setting with a cycle in progress.

Figure 7.5: Participants in the study.

7.4.5 Analysis

To analyse the data the computer application *Winhör* was used. This programme was created for the purpose of analysing data from acoustical comparison studies.[6] The check-box responses of the questionnaire were converted into numerical values according to the list on p. 262. Median, quartiles and statistical frequencies were calculated according to the ordinal level of the questions.

The median is the value that divides the distribution of measurement values into two equal populations, i.e. half the measured values fall below

[6]It was developed by the students Volker Banken (1993), Kai Schriewer (1995) and Andreas Pyka (1997) at the acoustics laboratory of the University of Applied Sciences in Düsseldorf, Germany.

the median, half above. The quartiles divide the measurement results into four equal populations. The second quartile is equal to the median. Looking at the first and third quartiles allows one to see the range of answers without extreme values. These are often caused by participants losing concentration while filling in the questionnaire, or other similar errors. Thus, these extreme responses are usually not looked at.

	Blk 1	Blk 2	Blk 3	Blk 4	Blk 5	Blk 6	Blk 7	Blk 8
Val.	100%	100%	100%	95.3%	95.3%	93%	100%	97.7%
Rel.	83.7%	83.7%	83.7%	70%	70%	76.7%	93.0%	74.4%

Table 7.2: Analysis of validity (Val.) and reliability (Rel.).

The absolute frequencies are calculated by counting the number of occurrences of each possible measurement value. Figures B.11 to B.26 in Appendix B.2 present the analysis of median, quartiles and frequencies (see histograms) of all eight blocks. Validity was determined by comparing one synthesis method with itself. Reliability was calculated on the basis of twice-compared pairs in each block. The analysis data for validity and reliability can be found in table 7.2. The abbreviation Blk stands for Block.

The calculations of frequencies, quartiles and median presented in Appendix B (figures B.11 to B.26) were analysed again and classified into the following groups:[7]

1. clear decision for an instrument (Clr),
2. taste decision with a tendency towards a clear decision (Twt),

[7]These terms are standard terms used at the acoustics laboratory of the University of Applied Sciences in Düsseldorf for statistical analysis of listening tests using a paired comparison.

3. taste decision without a tendency (Twot),
4. equal distribution with a tendency (Ewt),
5. equal distribution without a tendency (Ewot).

To be classified as a *clear decision (Clr)* the data must provide the following characteristics in the diagram of median and quartiles:

- Median must be outside the range between 2.5 and 3.5 (equal class).
- While one of both quartiles Q1 and Q3 may lie between 2.5 and 3.5, the line connecting Q1 and Q3 must not include the range between 2.5 and 3.5.

To be classified as a *taste decision with a tendency towards one of the two instruments compared (Twt)*, the following characteristics must be met:

- The histogram must show two poles. Two non-neighbouring lines must be higher than the other lines.
- One of the poles must be considerably higher than the other.
- The line connecting Q1 and Q3 must include the range between 2.5 and 3.5.
- The median must be outside the range between 2.5 and 3.5.

I refer here of a “taste decision” because the analysis shows that there are two clusters of participants who voted for each of the two systems. I conclude that a preference for one system is not based on its advantage but on the taste preference of a participant for one of the two systems compared.

To be classified as a *taste decision without a tendency (Twot)*, the following characteristics must be met:

- The histogram must show two poles. Two non-neighbouring lines must be higher than the other lines.
- The poles must be equally high.
- The line connecting Q1 and Q3 must include the range between 2.5 and 3.5.
- The median must be inside the range between 2.5 and 3.5.

This class was found relatively rarely in the results.

To be classified as an *equal distribution with a tendency towards specific instrument (Ewt)*, the following characteristics must be met:

- The histogram must not show two poles.
- One line in the histogram must be considerably higher than the others.
- The line connecting Q1 and Q3 must include the range between 2.5 and 3.5.
- The median must be inside the range between 2.5 and 3.5.

To be classified as an *equal distribution without a tendency towards a specific instrument (Ewot)*, the following characteristics must be met:

- The histogram must not show two poles.
- There must be no line which is considerably higher than the others in the histogram.
- The line connecting Q1 and Q3 must include the range between 2.5 and 3.5.
- The median must be inside the range between 2.5 and 3.5.

In order to calculate an overall rating, a system of point allocation was developed as presented below:

- clear decision for an instrument (Clr): 3 points;
- taste decision with a tendency towards a clear decision (Twt): 2 points;
- taste decision without a tendency (Twot): 1 point;
- equal distribution with a tendency (Ewt): 1 point;
- equal distribution without a tendency (Ewot): 0 points.

The analysis of block data for classification into the five categories can be found in table 7.3. This table uses the abbreviations Cmp (Comparison), Categ (Category), V (Validity), R (Reliability), Pts (Points), Blk (Block), and the abbreviations of categories as given above. By summing up the data in table 7.3, one can generate a table presenting the ranking of the

	Cmp 1	Cmp 2	Cmp 3	Cmp 4	Cmp 5	Cmp 6	Cmp 7	Cmp 8
Blk 1	Articulation: clearness							
Pair	A-C	A-B	C-D	B-B(V)	D-B	B-C(R)	C-B(R)	A-D
Categ	Clr A	Clr A	Ewt D	100%	Twt B	Clr B	Clr B	Twt A
Pts	A:3	A:3	D:1	-	B:2	B:1.5	B:1.5	A:2
Blk 2	Continuity of legato							
Pair	C-D(R)	D-B	A-B	C-D(R)	A-D	B-B(V)	A-C	B-C
Categ	Ewt C	Twt B	Twt A	Ewt C	Clr A	100%	Twt A	Twt B
Pts	C:0.5	B:2	A:2	C:0.5	A:3	-	A:2	B:2
Blk 3	Precision in spicatto							
Pair	C-D(R)	A-C	B-D	B-C	D-D(V)	D-C(R)	A-B	D-A
Categ	Clr D	Clr A	Twt D	Ewot	100%	Clr D	Clr A	Twt A
Pts	D:1.5	A:3	D:2	0	-	D:1.5	A:3	A:2
Blk 4	Musical expression							
Pair	B-D	A-C	B-C(R)	B-B(V)	B-C(R)	A-B	C-D	D-A
Categ	Clr D	Clr A	Ewt B	95.3%	Ewt B	Clr A	Clr D	Twt A
Pts	D:3	A:3	B:0.5	-	B:0.5	A:3	D:3	A:2
Blk 5	Pitch: vibrato							
Pair	A-D	D-C	A-B(R)	B-C	B-D	A-C	A-B(R)	B-B(V)
Categ	Clr D	Twt D	Clr B	Clr B	Twt D	Ewot	Ewot	95.3%
Pts	D:3	D:2	B:1.5	B:3	D:2	0	0	-
Blk 6	Dynamics: crescendo - decrescendo							
Pair	A-B(R)	C-D	A-C	A-B(R)	D-A	A-A	B-C	D-B
Categ	Clr B	Clr C	Ewt C	Clr B	Clr A	93%	Ewt B	Clr B
Pts	B:1.5	C:3	C:1	B:1.5	A:3	-	B:1	B:3
Blk 7	Timbre: change of timbre							
Pair	C-D	A-B	A-D	D-D(V)	B-C	D-B(R)	A-C	B-D(R)
Categ	Clr C	Clr A	Clr A	100%	Clr B	Clr B	Clr A	Clr B
Pts	C:3	A:3	A:3	-	B:3	B:1.5	A:3	B:1.5
Blk 8	Timing: rhythmical precision							
Pair	A-C	C-D(R)	A-A(V)	D-B	A-B	B-C	D-A	C-D(R)
Categ	Clr A	Clr C	97.7%	Clr B	Clr A	Clr B	Clr A	Ewot
Pts	A:3	C:1.5	-	B:3	A:3	B:3	A:3	0

Table 7.3: Analysis for classification into categories.

	Block 1 (points)	Block 2 (points)	Block 3 (points)	Block 4 (points)	Block 5 (points)	Block 6 (points)	Block 7 (points)	Block 8 (points)
	artic.	legato	precis.	expr.	vibr.	dynam.	timbre	timing
1	Pickup (8)	Pickup (7)	Pickup (8)	Pickup (8)	SD SS (7)	SD FM (7)	Pickup (9)	Pickup (9)
2	SD FM (5)	SD FM (4)	SD SS (5)	SD SS (6)	SD FM (4.5)	PD FM (4)	SD FM (6)	SD FM (6)
3	SD SS (1)	PD FM (1)	SD FM (0) PD FM (0)	SD FM (1)	Pickup (0) PD FM (0)	Pickup (2)	PD FM (3)	PD FM (1.5)
4	PD FM (0)	SD SS (0)	—	PD FM (0)	—	SD SS (0)	SD SS (0)	SD SS (0)

Table 7.4: Overview of ranking in blocks.

instruments in relation to the eight questions asked in the blocks. This data is presented in table 7.4. In the table the abbreviations SD FM (signal-driven simple FM), PD FM (parameter-driven simple FM), and SD SS (signal-driven subtractive synthesis) are used. Pickup stands for the pickup signal. An overview of the overall ranking and results will be provided in the following section.

7.4.6 Results

In order to calculate an overall ranking of the instruments, the points from all instruments and positions were summed up. The smaller the position sum the better, the larger the position sum the worse. Doing this produced the following ranking of instruments: A – B – D – C. The data calculated is presented in tables 7.5 and 7.6.

		1st position	2nd position	3rd position	4th position	Position sum
A	Pickup	6x	—	2x	—	12
B	SD FM	1x	5x	1x	1x	17
C	PD FM	—	1x	3x	4x	25
D	SD SS	1x	2x	1x	4x	24

Table 7.5: Overall results.

It can be concluded from this data that the listening test provides evidence that the signal-driven synthesis method can offer musically significant differences in sound result. The parameters I asked about are musically significant ones, as shown by the research outcomes in music psychology which were described above. The sound results were perceived to be different when using the same input signal.

Overall position		Instrument	Position sum	Point sum
1.	A	Pickup signal	12	52
2.	B	Signal-driven simple FM	17	33.5
3.	D	Signal-driven subtractive synthesis	24	19
4.	C	Parameter-driven simple FM	25	9.5

Table 7.6: Overall ranking.

It is possible that these differences may be considered to be advantages of the synthesis method. One can see, for example, a clearly perceived advantage of instrument D (signal-driven subtractive synthesis) over instrument C (parameter-driven FM) in terms of musical expression (in relation to the sound example). Or one can see a perceived advantage of the signal-driven methods over the pickup signal and parameter-driven FM in terms of vibrato.

A comparison of expected results and actual results reveals similarities and differences. Similarity is found in the overall ranking of the instruments and, for example, in the rankings in block 4 (expression) and block 1 (clearness). Differences are found in the fact that the pickup signal was not evaluated in all cases as having the highest validation in relation to an expressive cue.

With regard to dynamics, signal-driven and parameter-driven synthesis were perceived to have a wider range of dynamics than the pickup signal. Furthermore, vibrato was perceived to increase in intensity when using the methods of signal-driven subtractive synthesis and signal-driven FM. I assume in both cases that the incorporation of spectral changes caused the perception of these factors to be more intense, in one case as timbre vibrato and in the other as timbre dynamics.

Looking at the comments of the participants, it should be mentioned that a number of test subjects pointed out the “bubbling” of or “beeps” in sounds. This is caused by the pitch tracking problem (mostly in the transient phase of a tone). While the statements may be interpreted as indicating that unwanted pitch tracking data affects the signal-driven sound results less disturbingly than the parameter-driven ones, it is clear that the "bubbling" and "beeps" can be perceived as a problem in signal-driven synthesis too. On the other hand, in the context of a real-life situation, one can expect that a player will adapt his playing style to the measurement problem, and try to provide input material which prevents the trackers jumping until the correct fundamental frequency has been found. As mentioned earlier, adapting in this way will reduce the playability. However, if a performer accepts the drawback of reduced playability in order to make use of the advantage of a new sound, there is a worthwhile motivation for the player to try to avoid the unwanted "beeps".

In terms of point sum and position sum (see table 7.6), the distance between instrument B (signal-driven FM) and D (signal-driven subtractive synthesis) is greater than the distance between instrument D and C (parameter-driven FM). One might ask why the two signal-driven systems are not closer together. I assume from the statements of the test subjects that the use or non-use of the pitch tracker plays a role here. On the basis of my experience I expect that if I had used the pitch tracker in the signal-driven FM instrument too, instrument B would have been perceived as (and thus ranked) closer to instrument D.

In my view, a clear indication that ASDSS can increase the potential for musical expression is given in the result found in block 4 (expression). As I recorded just one short sequence relating to a specific musical style, the result here is of course valid only in this instance. Therefore, the result does not suggest that this increased potential would necessarily also be available in other expressive contexts. However, since the romantic phrase used can be seen as the kind of phrase traditional string players often play, I conclude that the potential for musical expression found here is valid in a wider context. In my view, the results of this study provide encouragement for the further investigation and development of signal-driven synthesis methods, because the results give a clear indication that ASDSS can offer expressive potential for musicians.

7.5 Compositions and Performances

Although the last two sections (7.4.5 and 7.4.6) have focused on musically relevant theoretical constructs, the research presented in this thesis has as its primary goal the production of outcomes which can be used in practical musical contexts. I therefore wanted to use ASDSS methods in stage-related composed and performed music. Since my scientific work allows only the measurement of specific indicators valid within a specific framework, experiences from the field of performance and composition are of great interest. They help to evaluate the instruments from a more general perspective.

The signal-driven methods presented in this thesis have been used in two compositions, several improvisations and three concerts.

The compositions were:

- Alexander J. Harker: *Torque*, 2006, for electric violin and live MaxMSP, written for Cornelius Poepel's electronic violin system (see Appendix E.1, CD 1, track 45),

- Carter Williams, *Chemical Sunset*, 2007, for two electric violins using Audio Signal-Driven Sound Synthesis and video (see Appendix E.1, CD 1, track 46, and E.2, DVD, chapter 5).

And the concerts were:

- duo digiStrings (Torsten Harder, cello and e-cello; Cornelius Poepel e-viola), June 4, 2005, Neustrelitz, Germany, and June 5, 2005, Berlin, Germany (see Appendix E.2, DVD, chapter 12),
- concert series "Nocturne", first concert by the group hot_strings SIG, June 26, 2007, Academy of Media Arts, Cologne, Germany (including the first performance of *Chemical Sunset* by Carter Williams).

In addition, a study etude was composed by composer and violinist Günter Marx, who is the leader of the Dortmund Philharmonic Orchestra. This etude includes many ASDSS sounds which the composer designed himself. A video documentation of this can be found in Appendix E.2 DVD, chapter 10. All three composers have extensive experience of electronic music and live electronics. Therefore, the first simple conclusion one can draw is that the method was found to be interesting enough to be chosen from among the vast number of possibilities for the context of an individual composition. Thus, for the composers concerned, it must offer meaningful opportunities.

It might be interesting to hear why they thought the method was worth investigating. First of all it should be mentioned that all three composers have many years of experience in violin-playing. Composer Alex Harker highlighted the ability to modify an FM sound spontaneously as required, simply by using the common playing techniques of a violin player. He wanted to be able to use malleable electronic sounds, and he simply liked the range of variations between clean and dirty that one can apply to the sound by changing discrete parameters.

Carter Williams was interested in the slight variations that one can apply to the sound by making subtle varying actions. He extended my ASDSS algorithms, adding a slow change in a set of discrete parameters which alter the synthesised sound. Asked for a personal statement describing his experiences, he wrote:

> As a composer and performer I am always looking for ways to extend the sound possibilities of instruments. My exploration of new sounds most often takes the acoustic instrument as its point of departure, and when I work with electronics I seek to augment and transform these sounds in such a way that the processing of the sounds does not mask or destroy the special characteristics of the sound. Often it seems that the input to many systems is arbitrary, as the synthesis engine has its own distinctive quality which overwhelms the often subtle nature of the acoustic signal. As such I am very interested in the development of the ASDSS system.
>
> My experience with this system was in general very positive. My piece *Chemical Sunset* uses various combinations of ASDSS methods (principally FM and MSSM) applied to the signals of two electric violins. The system was able to generate subtle sounds which exhibited a good range of control and variation based on various playing techniques, such as pizz., sul pont. or scraping

the strings with the wrapping of the bow. However some aspects of playability were unsatisfactory. The most significant problem I encountered was the limitation of modules which required pitch tracking. Here this was particularly a problem because of the desire to use double stops in the piece; latency and 'tracking mistakes' were the other major issues. The solution in this case was to drive the pitch variables automatically based on a pre-programmed sequence of events. Also it was often difficult to control dynamics. For instance performing gradual decrescendi and crescendi required more than normal amounts of control from the violinists giving the sounds an extremely delicate feel and fragile quality, which while in some contexts musically interesting, proved at times to be a limitation.

For me the areas for future work are twofold: the physical instrument interface and the synthesis modules. For me the biggest problem was that these sounds where designed to be used with bodyless electric violins—an instrument which, in my subjective opinion, presents many obstacles for the player. For me the most important disadvantage of the electric violin is that the player does not have the same range of timbres when one considers the unprocessed pickup signal. In my view this is due to the lack of a resonating body and severely limits the range of tone colours which are possible. As such any signal which the player can send into the synthesis system is a reduction of the range of sound modulation that the violinist is used to from the acoustic instrument. Because of this in my recent work with this method I have sought to use a traditional violin with the audio signal fed through a microphone. Despite the disadvantages of a microphone (feedback, bleed through from other instrument, ambient noise, etc.), for me the flexibility and playability gained by retaining the violin as an already mastered interface with the system outweighs any disadvantages. Of course one major difference in this approach is that the original sound of the violin will be necessarily always present in the final sound result. Even if the signal is not amplified its acoustical level cannot be masked. For example transposition effects with the electric violin are no problem, but when an acoustic instrument is used the original pitch level is unavoidably also present as it is impossible to render the violin silent, because of the presence of a sound board and body which project the sound. Consequently the sound modules em-

> ployed must take into account that a synthesis must include and blend with the violin's natural sound. My solution has been to employ a form of granular sampling in which multiple modules produce transpositions of the sound, thereby creating an artificial spectrum. This has analogies with the approaches of the spectral composers such as Grisey and Murail who sought to orchestrate spectral analyses of various instruments. Such a method, while certainly not the only possibility, does allow on the one hand a much more direct use of the sound generated by the performer and on the other hand a much more transparent way for the sound designer to manipulate the spectrum and therefore the brilliance and harmonicity of the resulting sounds.
>
> In conclusion this has been a very fruitful area of investigation. Moreover I believe strongly in such a performer centred approach to development and I think that such an approach once perfected has the potential to open up the world of computer music to a much wider range of musicians.

As Günter Marx was very interested in the ASDSS sounds, and as I was interested in studying how a professional performer would use them, a collaboration was begun. Several meetings were scheduled in order to explain the system and to design sounds. One notable observation was that the playability was found to be very good by the violinist.[8]

However, the question of how interesting a sound was felt to be played a much more important role. As mentioned earlier, the reason he wished to use electronic sounds was in order to have sounds that were completely different from the ones a violin can produce. Sounds such as the ones used in my studies presented in sections 7.3 and 7.4 were of no interest to him because they were too close to a traditional instrument.

[8]See the interview with Günter Marx on his perception and experiences of the method and sounds of ASDSS (Poepel, 2008a, pp. 193-194).

In order to design sounds he was interested in, the algorithms we took as a starting point were extended by adding numerous possibilities for mapping features extracted from the audio signal to discrete parameter inputs which I added to the algorithms. A problem often addressed by the violinist was the software interface of my MaxMSP patch, which was not very user-friendly. The development of user-friendly software interfaces is a considerable problem in software engineering. Though designing the software interface was outside the scope of this research, it could make a productive focus for future work.

In the concerts with the duo *digiStrings*, the sounds were used in improvisational parts of several compositions by Torsten Harder. For these concerts, VST plugins of my signal-driven synthesis engines were developed for use in the real-time sequencer *Ableton Live*.[9] The signal-driven FM synthesis sound with the ratio = 0.5 was perceived and mentioned by members of the audience as close to a warm bass clarinet sound. I had most enjoyed playing with parameter-driven subtractive synthesis sounds. In the rehearsals and concert I felt that the communication between my playing and the sound was good. The sound could easily be made intense, or given a certain kind of liveliness. I felt that I was in control of the tone and that it was more effectively "in my grip". Thus, I could react more effectively to the concert atmosphere and to the mood communicated by my partner. With regard to tracking problems, I found myself playing with the pitch tracker mistakes and resulting bubbling sounds in some passages of the concert.

[9] Ableton (2010).

I had many discussions with composers to whom I had presented my ASDSS algorithms. One statement by the composer Sergio Luque, given after he was introduced to my ASDSS system, is presented here as a summary of a frequently heard assessment: “I never heard this kind of sounds with so much information in it. All the little things you do are in the sound. If you would want to generate this electronically, you would need an enormous complex software” (Luque, personal communication, November 20, 2007).

7.6 User Group hot_strings SIG

Like many other researchers, I developed my system for the world of string players. In the course of testing my system and discussing the general idea of using electronics and computer systems for string performance, I became aware of the fact that many research results and developments (examples are described in chapter 4) are not known or used by the target users. Figure 7.6 explains this issue.

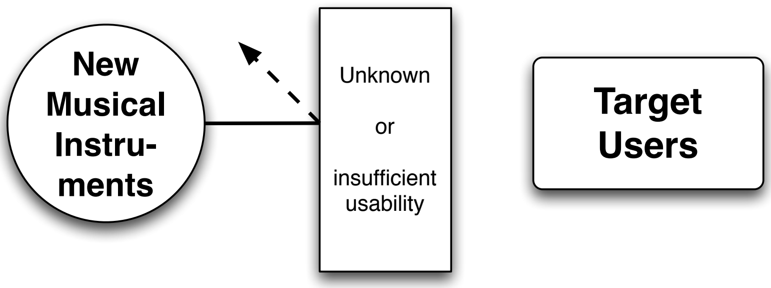

Figure 7.6: Missing link between researchers and target users.

The wishes and needs of users in relation to products are often different from what developers imagine. Besides this, developers and musicians need compositions for newly developed instruments, and composers have their own ideas about how music should develop in the future. These insights led me to the idea of bringing performers, researchers, instrument developers and composers together in order to exchange knowledge, wishes and experiences, and to initiate collaborative projects.

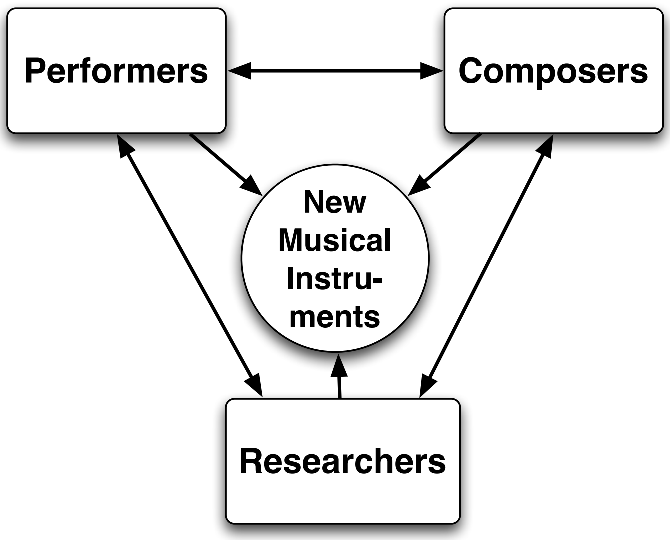

Figure 7.7: Performers, researchers and composers developing together.

Together with the violist Hatto Beyerle I founded such a group in December 2004. It was given the name hot_strings SIG. “Hot” stands for highly interesting instruments and “SIG” is an abbreviation for Special Interest Group which is often used in technology development groups. Thus, the name combines the instrumental and technical interest. The group, its actions and the results of its work were described in a paper (Poepel & Marx, 2007).

In relation to the work of this thesis, I was interested to see what the needs of the users were and how the users perceived my ASDSS system. In addition, I was interested in a user-centred design[10] for my instruments. The group was organised in such a way that meetings were held twice a year (until July 2007). 32 members primarily from Europe, joined the mailing list. The meetings had an average of seven participants. Two excerpts from a video documentation of the 5th meeting, which was held on October 23, 2006, at IRCAM, can be found on the accompanying DVD (see Appendix E.2, chapters 10 and 11, and figures 7.8, 7.9 and 7.10).

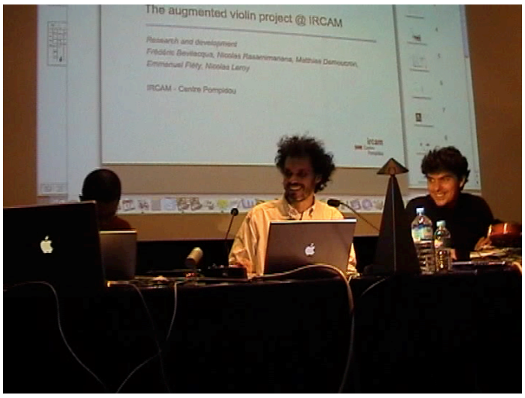

Figure 7.8: 5th SIG meeting at IRCAM: Nicolas Rasamimanana, Frédéric Bevilaqua and Matthias Demoucron talking about the IRCAM augmented violin project.

I do not wish to repeat here what has already been published in the paper (Poepel & Marx, 2007) about the group, its activity and the results of its work. However, I would like to summarise the experiences relevant for this research. One common experience was that string players who have undertaken the challenge of working with instruments which are different from traditional ones are often very individualistic. Thus, ideas about what

[10]The method of user-centred design has been described by Greenbaum and Kyng (1992). The design process involves cooperation between developers and users.

Figure 7.9: 5th SIG meeting at IRCAM: Günter Marx demonstrating signal-driven sounds.

Figure 7.10: 5th SIG meeting at IRCAM: Lenka Zupková presenting hardware extended sounds.

makes a good instrument, an interesting sound or a useful system can differ substantially.

The physicist and violin maker Martin Schleske (2010) mentioned in a meeting that, for him, the absence of a body would take away so much of the character and meaningful identity of a stringed instrument that he would not be able to use the sound. It might be interesting to develop instruments that keep the body but add small portions of ASDSS sounds in order to achieve expanded and useful timbres. In contrast, and as mentioned earlier, Günter Marx was interested in sounds that sounded very different from what he was used to from the traditional violin, but which he could still use his violins to play. The composer and violinist Lenka Zupková explained at the 5th meeting that she was generally satisfied with her electronic violin, but that she was seeking more warmth in the sound and a bigger potential for expressivity.

It was difficult, almost impossible even, to encourage a larger number of traditional string players to attend the group meetings and work with the instruments. When presenting videos of the related research work and developments made in the field, the enthusiasm was relatively limited. While the instruments were found to be scientifically interesting, the sound results and the expected playability were not so convincing that the attendees were eager to try out such instruments for themselves. An exception to this was the Hypercello.

I conclude that the needs of the string players who already had experience of electronic instruments tended towards the expansion of their existing instruments in accordance with their individual requirements. In addition, there was a need for support with the fine tuning of sounds and with making the system stageworthy in terms of robustness. From my experiences with traditional string players who are open in principle to new sounds and instruments, I conclude that one must first offer an instrument which builds a good "bridge" between the traditional instrument and its sounds, and the new instrument and its sounds. This bridge must address issues of easy handling and mechanical stability, and have at least a small repertoire and the ability to offer (amongst others) some traditional string-like sounds. The interface aspect, which often covers questions such as that of the actions the performer carries out, must be widened to include the questions of the expectations and needs the target users have, if the new developments are to be adopted by this target group.

Some researchers and musicians in the computer music scene argue that traditional musicians who do not use their computer-based instruments are not open to the new opportunities emerging in music. However, even if the majority of traditional string players might be somewhat reluctant, in my experience there is a community that is open for new developments. Unfortunately, the gap between their own field and the new field is still too great for them to bridge. Thus, the research into musical interfaces should take into account this wider view on interface questions as well.

The group hot_strings SIG organised a concert in order to present the results of collaborations, as described in the previous section. The concert included several compositions for strings and electronics. ASDSS sounds were used in the previously mentioned composition by Carter Williams, and in an improvisation. The composition focused on the subtle nuances applied by the performers to the sound. Little movements of the hands or arms could affect the sound in a more effective (and sometimes more problematic) way than expected. The improvisation allowed the use of a wide range of personal sounds and expression which the performers had a great deal of fun using. Figure 7.11 presents an image of the group at this concert. A documentation of the rehearsals can be found on the accompanying DVD, chapter “Images”, images no. 8-18 (Appendix E.2).

Figure 7.11: Performers of the first concert by hot_strings SIG.

7.7 Hypothesis and Results

The results of the evaluation indicate that the method of ASDSS can offer a meaningful potential for musical expression. With regard to the main hypothesis as presented in section 5.5, the results of the listeners' study (section 7.4.6) verifiy that implementations using the signal-driven approach to sound synthesis can show musically significant differences in sound result when compared to similarly complex implementations using the parameter-driven approach. As this was a listeners' study, the result is valid for the perspective of listeners. With regard to musically significant differences in playability, which can be assessed from the perspective of listeners only to a certain degree, the statements of string players who have used my developments indicate that performers using these signal-driven implementations can—depending on the performer's disposition—perceive a positive difference in playability in comparison to implementations using the parameter-driven approach.

The question which follows logically is what this means for the field of musical instruments and sound synthesis. This question will be addressed in the following sections. As I would like to incorporate discussions that have arisen doing my work on this thesis, I will begin with a discussion on the method and validity of the listeners' study.

7.7.1 Validity of Comparison

In discussing the listeners' study (section 7.4) with other researchers, the question arose as to whether the study was trying to compare approaches that were simply too different. As Dan Overholt put it,

> Overall, while I agree that ASDSS can be said to have more expressive power than PDSS (due to the use of the audio signal), I just look at ASDSS and PDSS as 2 different things - comparing them directly can be a bit "apples vs. oranges". (D. Overholt, email communication, October 5, 2007)

Overholt's argument is based on the conviction that ASDSS does not fall into the category of sound synthesis but of sound processing. Thus, the study did not compare different methods of sound synthesis, but methods of sound processing with methods of sound synthesis. The core problem—that of improving the interface, mapping and synthesis algorithms—is not addressed. Overholt argues further that the playability and sound result of an instrument based on sound processing will always be better for a string player than one based on parameter-driven sound synthesis, because of the use of the audio signal.

Firstly, it can be stated clearly on the basis of my results that the ASDSS sounds were not always perceived as better in all the parameters I tested. In block 6 (dynamics), for example, parameter-driven FM was clearly perceived to have a wider range of dynamics than signal-driven subtractive synthesis. And in block 2 (legato), parameter-driven FM was found to offer a slightly better continuity in legato than signal-driven subtractive synthesis. Thus, the belief that the parameter-driven version will always provide worse results is wrong.

Secondly, one may ask whether the construction principles of synthesisers should always follow the paradigm of the parameter-driven method as it was introduced by Max Mathews. In other areas of construction it is very common to mix methods. It is common, for example, to experiment with car engines that use different principles. And it is common to compare these for suitability in daily use and petrol consumption. Since I am building my instruments for musicians, what matters is the sound, the playability and the potential for meaningful musical action and expression, as perceived by the target users. As long as the users show musically meaningful preferences and do not complain about the sound being just an effect but not a synthesised sound, I believe that the comparison is valid. The signal-driven and parameter-driven methods both have the same goal, but use different construction principles.

I understand that, if one's definition of sound synthesis is based on the parameter-driven approach, one might want to complain about methods which cannot be compared. However, since according to my definition of sound synthesis (see p. 113) the method of ASDSS falls into this category, and since my categorisation parallels the perception of many of the target users, I do not think that I am comparing oranges and apples, but different apples with each other. The questions one might want to ask here could be: "What makes something an apple?", "Who is the apple for?" and "How, then, could it be produced?"

I conclude from this discussion that one of the challenges of the ASDSS approach is to conceive and consequently to construct a computer-based instrument in a different way. One might argue that this approach does not solve the problem as described above. I believe, however, that the core problem addressed here is the instrument and the performer using it, and not common shortcomings of traditional construction principles.

7.7.2 Discussion of Results

What do the results of my studies mean? What can one conclude from the results? While it was possible to verify the hypothesis with the results of the listeners' study, I would like to make clear that the results do not indicate that the signal-driven synthesis method is better in a general way than the parameter-driven one, or that it has greater potential and should therefore be the one on which future investigations focus. While the results confirm that there is potential, one cannot say how much potential ASDSS has in comparison to PDSS.

According to the measurement data results of the players' study (section 7.3), the players' assessment of the difference in potential for musical expression was relatively small (figure 7.1). As mentioned earlier, I believe that the problem of reduced playability may not necessarily have been perceived immediately when testing the instrument. This observation was very different from my expectations. I imagine that for a better investigation of this point, time would need to be invested in playing and getting to know the instrument. It would be interesting to see what happened if there were a study with a longer test period for players.

Working with performers, I had hoped early on in my work that once a downloadable version of ASDSS sounds was available, a number of string players would use my software for their own musical purposes. Particularly after the publication of the first paper on this topic I expected questions and requests from users. This was, however, not the case. The reasons I found were that traditional musicians generally need a system which is very easy to install. Even the need to use a pickup, laptop and speakers, and to get this system running, is far too much for many of them. While many string players were open to the task of testing the system and probably to using it, only a small number of users were really making full use of the system.

In order to create an instrument which is usable by a larger number of target players, problems such as a better software user interface and an overall system which is easy to handle first need to be solved.[11] What I have learned from this experience is that if I want a target group to use my musical interfaces and new digital instruments, the question of the musical interface does need to deal with the real problems the users encounter, even though such issues may seem mundane and of little academic interest.

While I wish to emphasise my conviction that the ASDSS approach deserves further investigation, I do not wish to dispute that the parameter-driven approach for use with stringed instruments seems not to have been

[11] Although these factors can impact significantly on playability, the primary focus of this study was on the instrumental interface, rather than the software interface. As noted earlier (see p. 281), the software interface will provide one focus for future development.

developed very far, and still has potential. I conclude from my experiences with string players such as Günter Marx that with a continuously parameter-driven Zeta OSC[12] violin, one could generate a substantial amount of musically interesting material. If one uses bowing systems that generate also parameter data, such as the Hyperbow (see section 4.3.5) or the *K-bow* (McMillen, 2008, p. 347), I believe that it will take time for the player to reach the "dead end" of the instrument which I addressed earlier (see section 2.1, p. 11).

[12]OSC (2010).

Chapter 8

Future Work

Many questions have arisen from the work presented in this thesis, and a range of possible improvements to the instruments presented has been identified. These possible improvements will be addressed first here, followed by the more general issues.

In order to give potential users a tool with increased usability, future work should include building software using the methods described, with a conventional effects unit and a manual and graphical user interface meeting the standard of comparable synthesisers. Besides this, it would be helpful to have VST plugins containing the methods. As some test subjects (especially those who were jazz violinists) mentioned that they would need a 19-inch rack-mount unit, it would be of interest to build one using ASDSS methods. If one could offer these applications, the opportunity for string players to use the system would be increased.

Many challenges for future work present themselves with regard to reducing the shortcomings observed so far in the present system. To overcome the problem of it not being possible to use double stops, it would be interesting to use a bridge such as the Zeta Jazz pickup (see figure 6.19), and to pick up each string signal separately and feed a separate ASDSS algorithm which might also address issues of "unevenness" across the strings. Depending on the separation of the different strings' audio signals, it should then be possible to play double stops. Pitch tracker mistakes can probably be reduced by using additional explicit parameters, such as the finger position on the finger board or the point where the finger hits the string. Using combinations of such measurements could improve the stability of the pitch tracker's output.

The number of ASDSS methods and sounds should, of course, be increased. The work of Lazzarini et al. (2007) is a valuable contribution in this field and more of this kind of work might be possible. In addition, the sounds designed so far can be improved or varied. Some sounds do still have a less than ideal balance of timbre in relation to the performed tone in varying pitches and amplitudes. This could be improved by using the concepts of key and velocity scaling from conventional synthesisers. Another option would be to investigate in greater depth how a compression of the audio signal and a shifting of the dynamics to specific dynamic bands would affect the sound or the balance of timbre, and to use these research results to offer the user new options in sound design.

The attack (transient) phase of the tone was mentioned as problematic because of the change in timbre that can occur. An alternative one might consider testing would be to build a hybrid using the pickup signal in the attack phase and the ASDSS signal afterwards. With regard to the haptic feedback found in the interface presented here, the feedback comes from the vibrating strings (essentially the pickup signal) and not the haptic feedback of the synthesised sound. It might be helpful to feed the synthesised signal via a mechanical oscillation to the bridge or other parts of the instrument's body, in order to increase the amount of haptic feedback concurrent with the synthesised sound.

Methods for the evaluation of synthesisers are still rare. Although a number of evaluations in the field of musical interfaces were published in 2008,[1] one may still ask how to evaluate in such a way as to offer the developer insight into how the instrument will be perceived by its target users. The question of how a long-term evaluation might be organised seems to be of particular importance. On the basis of my experience, I anticipate that the questions of what musical material to use (problem of aesthetically relevant material) and which method of research (quantitative, qualitative) to use would be of importance in such evaluations. Better differentiated methods for evaluation should be developed, both in relation to traditional musicians and to those who focus primarily on contemporary or experimental music.

In order to achieve better-differentiated models of musicians, it may be

[1] At the NIME conference in 2008, for example, Stowell, Plumbley, and Bryan-Kinns (2008), Kiefer, Collins, and Fitzpatrick (2008), or Geiger, Reckter, Paschke, Schultz, and Poepel (2008) presented evaluations of computer-based musical instruments.

helpful to consult literature on instrumental pedagogy. Corpus linguistics[2] might be a means by which to gain a deeper insight into the actions, meaning and qualities of string players. On a more general level, one might consider using literature on instrumental performance to study the interaction between a human being and a technical artefact. While it is not possible to study how a modern device such as a mobile phone was used in former times, string instruments offer a long and well documented history. In building a research field which one might call *Interaction Archaeology*,[3] musical instruments can be of use through the documentation of methods of dealing with them over hundreds of years.

With regard to the instrument itself, it might be fruitful to conduct studies with a larger population on the question of what qualities string players feel they already have, and which they are seeking to use with new instruments. I suggest developing prototypes of instruments and conducting field studies with string players using these prototypes. In my experience, direct work and discussion with the target users is of great importance in reaching the goal of making instruments with which users can be satisfied.

Formalisation is an issue that was addressed often in this research. Some researchers regard formalisation as the task to make an object's description

[2]According to McEnery and Wilson (2001) corpus linguistics can be described "as the study of language based on examples of 'real life' language use" (p. 1). Corpus linguistics may be applied here, for example, to the study of text corpuses produced by violin schools and focusing on the description and understanding of the actions of performers and the performers themselves.

[3]This term parallels the term *Archeology of Media* coined by the media theorist Siegfried Zielinski (2010).

more and more precise until it is precise enough to be rendered in a computer language. By contrast, I demonstrated that parts of the objects themselves may be lost in the course of this process, because abstractions must be defined. It seems that objects or actions exist in musical performance that need an imprecise description in order to create space for the player to perform or generate them. Thus, researchers of computer music need to work on the question whether objects may in fact exist that cannot be talked about or defined in words and physical descriptions. And the next question would be that of how the necessity for formalisation in computer-based instruments can be addressed without damaging these objects or excluding them from of a music system.

Looking at the objects that can be described and measured, one can ask what happens in the microstructure of a performer's actions when a personal style of playing is developed. With reference to the PDSS approach, one can ask how an interface might need to be able to adapt to such a personal style of performance. One might consider different sets of sensor systems that can be adjusted to the style of the individual player. Or one might have a larger set of mapping methods that can be specified to adapt to the personal style of a performer.

With regard to the model of and assumptions about music which one uses when following Max Mathews' MUSIC-N approach, it might be worthwhile to consider the pragmatic features of this model in greater detail. It would be interesting to establish where the assumed generalisation of this model has its limitations, and what this might mean for a musician. The investigation of this model and its features might offer useful results for computer-based music systems, especially when designed for real-time performance.

When discussing my system with other researchers, the question arose as to whether it might be possible to use the system with other instruments. Of course ASDSS can be used with any kind of instrument that can generate an audio signal. It can, for example, be implemented as VST plugins and be used with different inputs. However, one does not get an ideal result simply by using the software developed here with other instruments. In my experience, the algorithms need to be tuned to the different instruments because they differ on the macroscopic level in timbre, articulation and methods of playing. Therefore, they generate different audio signals which must be treated differently in the algorithms in order to retain the essential features of the instrument selected in each case.

In the present work, performers were addressed who already have skills and want to use these for the performance of synthesised sounds. It is interesting to ask whether the openness of an instrument and the ability to explore it more and more thoroughly might play a meaningful role for players who do not come with the pre-existing skills of a traditional musician. One could ask for example: is it possible that systems incorporating mechanically oscillating objects connected to an algorithm can create a meaningful quality even if the performer does not come from the background of a traditional instrument? One might for example, build computer-based instruments that are not played by knobs and faders but by oscillating mechanical objects such as strings, metal tongues (as in the case of a thumb-piano or mbira) or thin batons. Or an instrument could be included in a console for sound diffusion and performance of a mix of tape-based and live electronic music. A mixing console could, for example, include something like a fader in which a wound string is scraped. There are many possibilities for combining oscillating objects and sound algorithms.

The string instruments presented in this thesis are designed for string players searching for new musical possibilities. However, only a few of them use electronic stringed instruments. Thus, besides the future need to offer instruments that prove to be worth using, I perceive a need to offer accompanying workshops for the purpose of getting to know and mastering the instruments. In addition to the design of the instrument itself, the interface aspect thus includes the work on presenting the new instruments, their opportunities and limitations, and related compositions and performers.

Appendix A

Materials: Performers' Study

A.1 Questionnaire

This study was conducted in Cologne, Germany, and the language of the questionnaire was therefore German. I provide the German version of the questionnaire below as well as an English translation. The translation is by Shivaun Heath.

Empirische Studie - Drei elektronische Streichinstrumente

Erst mal vielen Dank für´s Mitmachen! Ich hoffe das Probieren am Instrument macht ihnen Freude. Die Studie ist selbstverständlich anonym. Sämtliche Daten werden nur innerhalb dieser Studie benutzt.

Ein paar Fragen an die Person:

Geschlecht (M/W) ____

Alter ____

Würden Sie sich eher der Gruppe der Laien oder der Profis am Streichinstrument zuordnen ?
(L / P) ______

Haben Sie Erfahrung mit elektronischen Streichinstrumenten? (Eine Zahl auf der Skala zwischen 1 und 5, 1 bedeutet: >ich habe überhaupt keine Erfahrung mit elektronischen Streichinstrumenten<, 5 bedeutet: >Ja, ich bin Experte<) ______

Wo sehen Sie Ihre musikalische Präferenz: in der Musik vor 1950 oder in der Musik nach 1950 ?
(davor / danach) _________

Wann wurde die Musik komponiert die Sie meistens Spielen, vor 1950 oder nach 1950?
(davor / danach) _________

Interessiert Sie das Material Elektronik und Geräte aus diesem Material z.B. HiFi- Anlagen, Waschmaschinen, Computer, Handys etc. ? (Zahl auf der Skala zwischen 1 und 5, 1: >nein, überhaupt nicht< 5: >ja interessiert mich sehr<)______

Haben Sie Interesse an elektronischen Sounds, Synthesizerklängen, verstärkten Streichinstrumenten, spielen über Effektgeräte? (Zahl auf der Skala zwischen 1 und 5, 1 bedeutet: >nein, überhaupt nicht< 5 bedeutet: >ja interessiert mich sehr<)______

Und nun entspannen Sie Sich, überlegen Sie beim beantworten der Fragen nicht lange, sondern antworten Sie spontan, so wie Sie es einfach empfinden.

1

Figure A.1: Questionnaire used in study with performers, page 1.

Translation of questionnaire, page 1:

Empirical Study - Three Electronic Stringed Instruments

First of all thanks very much for participating in this study! I hope that you will enjoy trying out the instruments. The study is of course anonymous. All of the data collected in the study will be used only for the purpose of this research.

A few questions about yourself:

Sex (M/F) ____

Age ____

Would you classify yourself more as a professional or as an amateur string player?
(P/A)

Do you have any experience of electronic stringed instruments? (a number on a scale between 1 and 5, 1 meaning: "I have no experience at all with electronic stringed instruments", 5 meaning: "yes, I am an expert")

What music do you prefer: music from before 1950 or music from after 1950?
(before/after)

When was the majority of the music that you play composed, before 1950 or after 1950?
(before/after) ________________

Are you interested in electronics in general and electronics devices such as hi-fis, washing machines, computers, mobile phones, etc.?
(a number on a scale between 1 and 5, 1." no, not at all", 5: "yes, I'm very interested")
____

Are you interested in electronic sounds, synthesiser sounds, amplified stringed instruments, playing through effects units?
(a number on a scale between 1 and 5, 1 meaning: "no, not at all",5 meaning "yes, I'm very interested") ________

And now relax, try not to think too long about your responses to the questions, but answer spontaneously, simply giving your impressions.

Der Instrumententest

Der Test besteht aus 20 Fragen.

Sie finden jeweils eine Vorgabe was gespielt werden soll. Dann finden Sie die Frage zu dem eben gespielten und gehörten. Weiterhin finden Sie die Buchstaben A, B und C. A, B und C stehen für die drei Instrumententypen die getestet werden sollen. Bei A, B und C sollte jeweils eine ganze Zahl zwischen 1 und 5 eingetragen werden.

Die Zahlen bilden eine 5-Teilige Skala zwischen Nein (1) und Ja (5) und haben ungefähr folgende Bedeutung:

1: Nein, überhaupt nicht (unbrauchbar)

2: Nein, eher nicht (mangelhaft)

3: Na ja, erkennbar aber nicht nur sehr bedingt brauchbar (ausreichend)

4: Ja, mit Einschränkungen (befriedigend)

5: Ja, sehr gut

Die Fragen

1) auf einer Tonhöhe: cresc. / decresc. Dynamik, pp, mf, ff, sfz

Ist das Verhältnis >Bewegung (oder Spieltechnik) – elektronischer Klang< in Bezug auf die Dynamik stimmig?

A: ____ B:____ C:____

2) auf einer Tonhöhe: Ton halten (Spannung im Ton, Entspannung), scharfer Ton , weicher Ton, Strichart Collé (Saite mit dem Bogen zupfen)

Ist das beabsichtigte Klangergebnis im tatsächlichen Klangergebnis gut repräsentiert ?

A: ____ B:____ C:____

3) Glissando, unterschiedliche Töne mit Pausen (Intervalle bitte variieren), Vibrato

Ist das beabsichtigte Klangergebnis im tatsächlichen Klangergebnis in Bezug auf die Tonhöhe gut repräsentiert

A: ____ B:____ C:____

2

Figure A.2: Questionnaire used in study with performers, page 2.

Translation of questionnaire, page 2:

The Instrument Test

The test consists of 20 questions.

In each case there is an example to be played. Then you will find a question about what you have just played and how it sounded to you. You will also find the letters A, B or C. A, B and C stand for the three instrument types which are being tested. For each letter A, B and C, a whole number between 1 and 5 should be entered.

The numbers represent a five part scale between No (1) and Yes (5), and have the following approximate meanings:

1: No, not at all (unusable)
2: No, not really (poor)
3: Yes and no, recognisable but not very usable (fair)
4: Yes, with some reservations (satisfactory)
5: Yes, very good

The Questions

1) on one pitch: cresc./decresc. dynamics, pp, mf, ff, sfz
Does the physical gesture (or playing technique) correspond to the electronic sound in terms of dynamics?

A: ____ B: ____ C: ____

2) on one pitch: holding the note (tension in the tone, relaxation), sharp tone, gentle tone, coll bow stroke (plucking the string with the bow)
Is the intended sound result well reflected in the actual sound result?

A: ____ B: ____ C: ____

3) glissandi, different pitches with rests (please vary the intervals), vibrato
Is the intended sound result well reflected in the actual sound result in terms of pitch?

A: ____ B: ____ C: ____

4) sf decresc, cresc, kurze Töne, pizz, langsamer Springbogen

Ist das Verhältnis Bewegung (oder Spieltechnik) – elektronischer Klang (Fokus auf die Ansprache/Reaktion des Instruments) stimmig?

A: ____ B:____ C:____

5) Glissando, unterschiedliche Töne mit Pausen dazwischen, variierende Intervalle

Ist das Verhältnis Bewegung (oder Spieltechnik) – elektronischer Klang stimmig?

A: ____ B:____ C:____

6) zum Klang der elektronischen Instrumente passend, Musik eigener Wahl: langsame, getragene Melodie mit starkem Ausdruck, schnelle Passage aus einer Komposition oder Improvisation

Ist das Instrument für die angewandte Spielweise (Tonhöhe/Dynamik /Reaktionsgeschwindigkeit/Artikulation) und den musikalischen Ausdruck transparent?

A: ____ B:____ C:____

7) unterschiedliche Kontaktstellen, kurze Töne, sf, decresc, wechsel Kontaktst.cresc, pizz

Ist das beabsichtigte Klangergebnis im tatsächlichen Klangergebnis gut repräsentiert (Reaktionsgeschwindigkeit)?

A: ____ B:____ C:____

8) auf einer Tonhöhe aber in unterschiedlichen Lautstärken: spitzer Ton, stumpfer Ton, Pizz, Détaché, Spicatto, Sul Ponticello, Kratzen

Ist das Instrument für die angewandte Spielweise (Ansprache/Artikulation) transparent?

A: ____ B:____ C:____

9) Tonleiter durch zwei Oktaven martelé dann in je 8 Tönen gebunden

Ist das beabsichtigte Klangergebnis im tatsächlichen Klangergebnis gut repräsentiert (Ansprache/Reaktionszeit)?

A: ____ B:____ C:____

3

Figure A.3: Questionnaire used in study with performers, page 3.

Translation of questionnaire, page 3:

4) sf decres., cresc., short tones, pizz., slow ricochet (jet)

Does the physical gesture (or playing technique) related to the electronic sound (focus on the response/reaction of the instrument)?

A: ____ B: ____ C: ____

5) glissandi, different pitches with rests in between, varying the intervals

Does the physical gesture (or playing technique) correspond to the electronic sound?

A: ____ B: ____ C: ____

6) music of your choice which suits the sound of the electronic instrument: long, extended melodies (expressivo), fast passages from a composition or improvisation

Is the instrument transparent for the playing techniques used (pitch/dynamics/responsiveness/articulation) and for musical expression?

A: ____ B: ____ C: ____

7) different bow-bridge distances, short tones, sf, decresc., change bow position, cresc., pizz.

Is the intended sound result well reflected in the actual sound result (responsiveness)?

A: ____ B: ____ C: ____

8) on one pitch, but with different dynamics: sharp tone, dull tone, pizz., dtach, spicatto, sul ponticello, scratching

Is the instrument transparent for the playing techniques used (responsiveness/articulation)?

A: ____ B: ____ C: ____

9) scales over two octaves martelé then slurred in groups of eight notes

Is the intended sound result well reflected in the actual sound result (especially attack/responsiveness)?

A: ____ B: ____ C: ____

10) gebrochene Dreiklänge durch drei Oktaven, Oktav-Glissando (von oben nach unten)
Ist das Instrument für die angewandte Spielweise (Tonhöhe) transparent?

A: ____ B:____ C:____

11) in den Ton reingehen, daran ziehen: lange Töne halten, lange crescendi und decrescendi.
Ist das beabsichtigte Klangergebnis (und Klanggefühl) im tatsächlichen Klangergebnis gut repräsentiert ?

A: ____ B:____ C:____

12) Perpetuum Mobile, Andante e Rondo ungarese
Ist das Verhältnis Bewegung (oder Spieltechnik) – elektronischer Klang stimmig?

A: ____ B:____ C:____

13) Andante e Rondo ungarese
Ist das beabsichtigte Klangergebnis (und Klanggefühl) im tatsächlichen Klangergebnis gut repräsentiert ?

A: ____ B:____ C:____

14) oben beginnend: Tonleiter durch zwei Oktaven 4er Bindungen, dann Spicatto
Ist das Instrument für die angewandte Spielweise (Artikulation/Tonhöhe) transparent?

A: ____ B:____ C:____

15) Ton so lange wie möglich halten, Martelé (langsam) mit scharfem Attack,
Ist das Instrument für die angewandte Spielweise (Ausdrucksfarbe/Artikulation) transparent?

A: ____ B:____ C:____

4

Figure A.4: Questionnaire used in study with performers, page 4.

Translation of questionnaire, page 4:

10) arpeggios over three octaves, octave glissandi (from high to low)

Is the instrument transparent for the playing techniques used (pitch accuracy)?

A: ____ B: ____ C: ____

11) digging into the tone and maintaining the intensity: long sustained tones, long crescendi and decrescendi

Is the intended sound result (and feeling) well reflected in the actual sound result?

A: ____ B: ____ C: ____

12) Perpetuum Mobile, Andante e Rondo ungarese

Does the physical gesture (or playing technique) correspond to the electronic sound?

A: ____ B: ____ C: ____

13) Andante e Rondo ungarese

Is the intended sound result (and feeling) well reflected in the actual sound result?

A: ____ B: ____ C: ____

14) starting above: scales over two octaves, slurred in groups of four, then spicatto

Is the instrument transparent for the playing techniques used (articulation/pitch accuracy)?

A: ____ B: ____ C: ____

15) holding a note for as long as possible, martelé (slow) with a sharp attack

Is the instrument transparent for the playing techniques used (articulation/expression)?

A: ____ B: ____ C: ____

16) warme Töne , kalte Töne

Ist das beabsichtigte Klangergebnis im tatsächlichen Klangergebnis/Klanggefühl repräsentiert?

A: ____ B:____ C:____

17) eine aggressive Musik freier Wahl

Ist das beabsichtigte Klangergebnis im tatsächlichen Klangergebnis gut repräsentiert ?

A: ____ B:____ C:____

18) Eine lebendige, füllige und blutvolle Musik eigener Wahl

Ist das Instrument für die angewandte Spielweise (Ausdrucksfarbe/Artikulation) transparent?

A: ____ B:____ C:____

19) Vom bisherigen Eindruck her:

Ist das Instrument gut durchlässig / brauchbar für musikalischen Ausdruck?

A: ____ B:____ C:____

20) Vom bisherigen Eindruck her:

Ist das Instrument gut durchlässig / brauchbar für Ihren persönlichen musikalischen Ausdruck?

A: ____ B:____ C:____

5

Figure A.5: Questionnaire used in study with performers, page 5.

Translation of questionnaire, page 5:

16) warm tones, cold tones

Is the intended sound result well reflected in the actual sound result/feeling?

A: ____ B: ____ C: ____

17) aggressive music of your choice

Is the desired sound result well reflected in the actual sound result?

A: ____ B: ____ C: ____

18) lively, full blooded, sanguine music of your choice

Is the instrument transparent for the playing techniques used (articulation/expression)?

A: ____ B: ____ C: ____

19) based on your impressions so far:

Is the instrument sufficiently transparent to/useful for musical expression?

A: ____ B: ____ C: ____

20) from your impressions up to this point:

Is the instrument appropriately transparent to/useful for your personal musical expression?

A: ____ B: ____ C: ____

Appendix B

Materials: Listeners' Study

B.1 Questionnaire

This study was conducted in Cologne and Düsseldorf, Germany, and the questionnaire was therefore in German. A translation of the questionnaire is provided below. The translation is by Damaris Schultz-Pöpel.

Hörversuch

Datum:

Uhrzeit:

Studiengang bzw. Beruf:

Alter:

Geschlecht:

Ich spiele ein Instrument: ja ...... nein .....

Ich höre öfters elektronische Musik: ja ...... nein .....

In diesem Hörversuch werden verschiedene Klänge gegenübergestellt, die mit einer elektronischen Viola erzeugt wurden.

Der Hörversuch ist in neun Blöcke aufgeteilt. Jeder Block beginnt mit einer Frage, nach der Sie die darauf folgenden Hörbeispiele beurteilen sollen. Bitte geben Sie Bescheid, sobald Sie nach Durchlesen der Frage für die Hörbeispiele bereit sind (Bei Verständnisproblemen stellen sie bitte Fragen). Zu Beginn jedes Blocks wird vom Vorführenden die Nummer des Blocks angesagt.

Die Beispiele, die mit der Raute und dem Quadrat bezeichnet sind, werden in der Reihenfolge Raute – Quadrat – Raute - Quadrat angeboten. Ein Paar beinhaltet eine Phrase der elektronischen Viola, die von unterschiedlichen Synthesizern wiedergegeben wird. Eine Phrase dauert ca. 9 Sekunden. Nachdem ein Phrasen-Paar zweimal vorgestellt wurde, haben Sie zwei Sekunden Zeit sich für eine Antwort zu entscheiden. Zu beurteilen ist nur, worauf sich die Fragestellung bezieht. Machen Sie in jedem Fall ein Kreuz, auch wenn Sie nicht 100%ig sicher sind, weil Ihr Fragebogen ansonsten nicht ausgewertet werden kann. Zusätzlich haben Sie die Möglichkeit, im Feld „Bemerkungen" besondere Auffälligkeiten zu beschreiben. Bitte machen Sie von dieser Möglichkeit Gebrauch!

Vielen Dank für die Mitwirkung bei diesem Hörversuch!

Figure B.1: Questionnaire used in study with listeners, page 1.

Translation of page 1:

Listening Experiment

Date:

Time:

Course of studies or profession:

Age:

Gender:

I play an instrument: yes.... no....

I frequently listen to electronic music: yes.... no....

In the following listening experiment, different sounds generated by an electronic viola are compared.

The experiment is divided into nine parts. Each part begins with a question with which you are asked to evaluate the listening examples which follow. Please let me know when you have read the question and are ready to start (if you have any questions, please ask). At the beginning of each part the experiment guide will announce the number of the part.

Examples represented by a diamond and a square will be played in the following order: diamond – square – diamond – square. One pair of examples consists of one phrase played on the electronic viola, reproduced by different synthesizers. Each phrase lasts approximately nine seconds. After a phrase pair has been played twice, you will have two seconds to decide on an answer. You are only asked to evalute what the question refers to. Please always make a cross, even if you are not 100 % sure of your answer. Otherwise your questionnaire cannot be evaluated. Additionally you have the opportunity to decribe any particular problems in the "comments" box. Please make use of this opportunity!

Thank you very much for participating in this listening experiment!

Block 1

Klarheit

In diesem Block werden die Beispiele der Strichart Martelé auf das Kriterium des klaren Hervortretens der einzelnen Töne untersucht.

Definition: Klarheit eines Tons im Martelé ist die Möglichkeit, den Beginn, die Tonhöhe und das Ende eines Tones deutlich zu erkennen.

Fragestellung: **In welchem Beispiel treten die einzelnen Töne klarer hervor?**

◇			□		
klarer			**klarer**		
viel	**etwas**	**gleich**	**etwas**	**viel**	**Bemerkungen**

Figure B.2: Questionnaire used in study with listeners, page 2.

Translation of page 2:

Part 1

Clearness

In this part, examples using the bowing technique 'martelé' are examined with regard to the clear prominence of the single notes.

Definition: the clarity of a note in 'martelé' is the possibility of recognising clearly the start, pitch and end of a note.

Question: **In which example is the note more clearly prominent?**

Block 2

Legato

In diesem Block werden die Beispiele eines Legatos auf das Kriterium der kontinuierlichen Verbindung zwischen den Tönen untersucht.

Definition: Unter Kontinuität der Verbindung wird hier der geschmeidige, bruchlose und gebundene Übergang von einem zum nächsten Ton ohne Neuartikulation verstanden.

Ein gelegentlich zu hörender Piepston ist durch die jeweilige Synthesevariante bedingt und in ihre Bewertung einzubeziehen!

Fragestellung: **In welchem Beispiel ist eine bessere Kontinuität der Verbindung vorhanden?**

◇			□		
besser			**besser**		
viel	**etwas**	**gleich**	**etwas**	**viel**	**Bemerkungen**

Figure B.3: Questionnaire used in study with listeners, page 3.

Translation of page 3:

Part 2

Legato

In this part, examples of a legato are examined with regard to a continous conjunction between the notes.

Definition: the continous conjunction is understood here as a smooth, unbroken and linked transition from one note to the next, without new articulation.

The occasionately heard peep is caused by the respective synthesis version and you are asked to include it in your evaluation.

Question: **In which example did you hear a better continuous conjunction?**

Block 3

Deutlichkeit

In diesem Block werden die Beispiele einer schnellen Passage auf das Kriterium der Deutlichkeit untersucht.

Definition: Unter Deutlichkeit wird die klare Erkennbarkeit und Präzision der einzelnen Töne verstanden.

Fragestellung: **In welchem Beispiel ist eine bessere Deutlichkeit der musikalischen Passage gegeben?**

◇			□		
besser			**besser**		
viel	**etwas**	**gleich**	**etwas**	**viel**	**Bemerkungen**

Figure B.4: Questionnaire used in study with listeners, page 4.

Translation of page 4:

Part 3

Precision

In this part, examples from a fast section are examined with regard to precision.

Definition: precision is unterstood as the clear perception and articulation of the individual notes.

Question: **In which example is a better precision in the section of music found?**

Block 4

Ausdruck

In diesem Block werden die Beispiele einer langsamen Passage auf das Kriterium des musikalischen Ausdrucks untersucht.

<u>Definition</u>: Unter musikalischem Ausdruck wird die Hörbarmachung eines inneren Vorgangs (z.B. von Emotionen) verstanden.
Bei diesem Beispiel handelt es sich um getragenen bis traurigen Ausdruck.

Fragestellung: **In welchem der beiden Beispiele kommt dieser getragene bis traurige Ausdruck besser zur Geltung?**

◇			□		
besser			**besser**		
viel	**etwas**	**gleich**	**etwas**	**viel**	**Bemerkungen**

Figure B.5: Questionnaire used in study with listeners, page 5.

Translation of page 5:

Part 4

Expression

In this part, examples of a slow section are examined with regard to musical expression.

Definition: musical expression is understood as an inner process (e.g. an emotion) made audible. In this example it is a slow, melancholic expression.

Question: **In which example is the slow and melancholic phrase better expressed?**

Block 5

Vibrato

In diesem Block werden die Beispiele eines streicherüblichen Vibratos auf das Kriterium der Intensität des Vibratos untersucht.

Definition: Unter Intensität wird die Möglichkeit des intensiven Ausdrucks in einem Ton verstanden, welche durch Vibrato erzeugt wird.

Fragestellung: **In welchem der Beispiele ist ein intensiveres Vibrato zu hören?**

◇			□		
intensiver			**intensiver**		
viel	**etwas**	**gleich**	**etwas**	**viel**	**Bemerkungen**

Figure B.6: Questionnaire used in study with listeners, page 6.

Translation of page 6:

Part 5

Vibrato

In this part, examples of a typical string vibrato are examined in regard to the intensity of the vibrato.

Definition: the intensity is understood as the possibility to generate an intensive expression in a note by using vibrato.

Question: **In which example do you hear a more intense vibrato?**

Block 6

Crescendo - Decrescendo

In diesem Block werden die Beispiele eines Crescendos und Decrescendos auf das Kriterium der dynamischen Veränderung untersucht.

<u>Definition</u>: Unter An- und Abschwellen der Lautstärke wird die wahrgenommene kontinuierliche Veränderung der Dynamik eines Klangs verstanden.

Fragestellung: **In welchem Beispiel ist ein stärkeres An- und Abschwellen der Lautstärke zu hören? (stärker im Sinne von größer oder umfangreicher)**

◇			□		
stärker			**stärker**		
viel	**etwas**	**gleich**	**etwas**	**viel**	**Bemerkungen**

Figure B.7: Questionnaire used in study with listeners, page 7.

Translation of page 7:

Part 6

Crescendo – Decrescendo

In this part, examples of crescendo and decrescendo are examined with regard to the dynamic change.

Definition: the rising and ebbing away of volume is understood as the perceived continous change in the dynamic of a sound.

Question: **In which example can you hear a stronger rising and ebbing away of the volume? (stronger in the sense of larger and more voluminous)**

Block 7

Klangveränderung

In diesem Block werden die Beispiele eines Kontaktstellenwechsels auf das Kriterium der Veränderung des Klangs untersucht. Die Kontaktstelle ist die Stelle, an der der Bogen auf die Saite aufgesetzt wird.

<u>Definition</u>: Unter Veränderung des Klangs wird die klangfarbliche Veränderung während eines Tons verstanden.

Fragestellung: **In welchem Beispiel ist eine stärkere Veränderung des Klangs von weich zu scharf vorhanden?**

◇			□		
stärker			**stärker**		
viel	**etwas**	**gleich**	**etwas**	**viel**	**Bemerkungen**

Figure B.8: Questionnaire used in study with listeners, page 8.

Translation of page 8:

Part 7

Change of Timbre

In this part, examples of different bow positions are examined with regard to sound modification. The bow position is the place where the bow touches the string.

Definition: the change of timbre is understood as the modification of the tone colour within a note.

Question: **In which example is a stronger modification of the sound from soft to hard present?**

Block 8

Rhythmische Präzision

In diesem Block werden die Beispiele auf das Kriterium der rhythmischen Präzision untersucht.

<u>Definition</u>: Unter rhythmischer Präzision wird die exakt Wiedergabe eines gegebenen Rhythmus verstanden.

Fragestellung: **In welchem Beispiel ist die Präzision des gespielten Rhythmus besser zu hören?**

◇			□		
besser			**besser**		
viel	**etwas**	**gleich**	**etwas**	**viel**	**Bemerkungen**

Figure B.9: Questionnaire used in study with listeners, page 9.

Translation of page 9:

Part 8

Rhythmical Precision

In this part, examples are examined with regard to rhythmic precision.

Definition: rhythmical precision is understood as the exact reproduction of a given rhythm.

Question: **In which example do you hear the precision of the rhythm better?**

Block 9

Tonhöhenveränderung

In diesem Block werden die Beispiele auf das Kriterium der Gleichmässigkeit in der Tonhöhenveränderung untersucht.

Definition: Unter Gleichmässigkeit der Tonhöhenveränderung wird die wahrgenommene kontinuierliche Veränderung der Tonhöhe verstanden.

Fragestellung: **In welchem Beispiel ist die Gleichmäßigkeit der Tonhöhenveränderung besser?**

◇			□		
besser			**besser**		
viel	**etwas**	**gleich**	**etwas**	**viel**	**Bemerkungen**

Figure B.10: Questionnaire used in study with listeners, page 10.

Translation of page 10:

Part 9

Change of Pitch

In this part, examples are examined with regard to constant change of pitch.

Definition: the constancy in change of pitch is understood as the percieved continuous change of pitch.

Question: **In which example is the constancy in the change of pitch better?**

B.2 Results

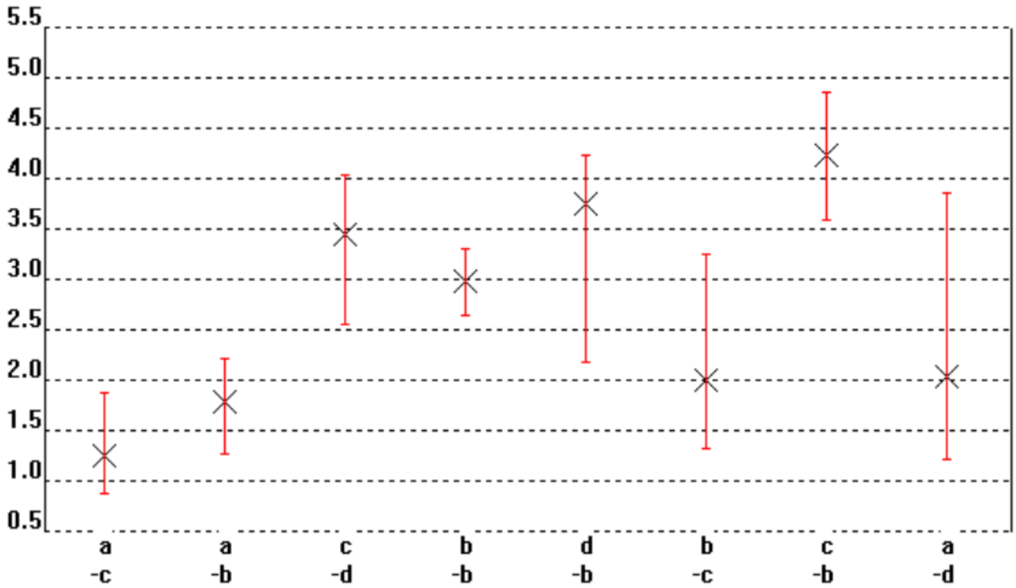

Figure B.11: Median and quartiles of block 1 (clearness).

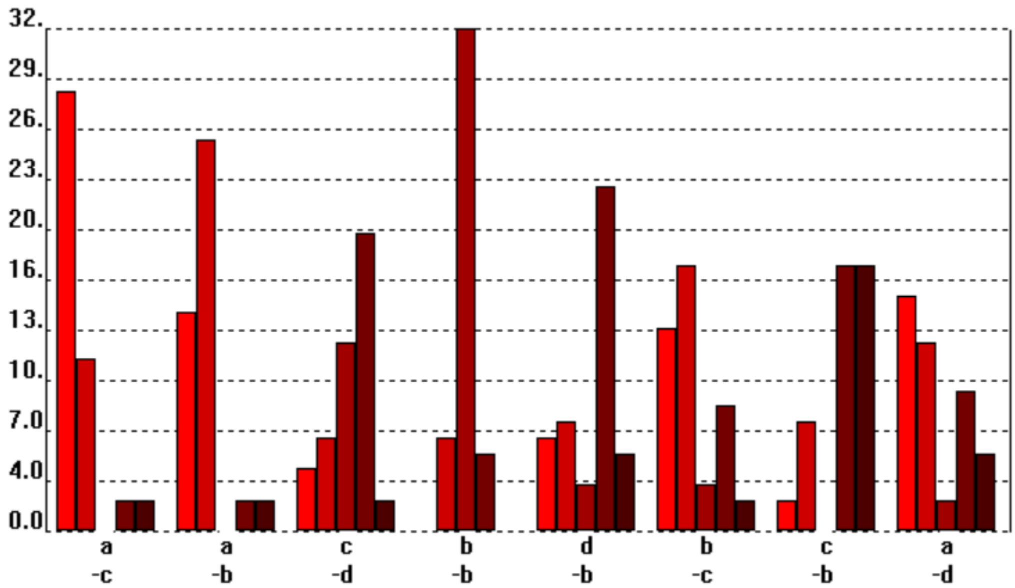

Figure B.12: Histogram of block 1 (clearness).

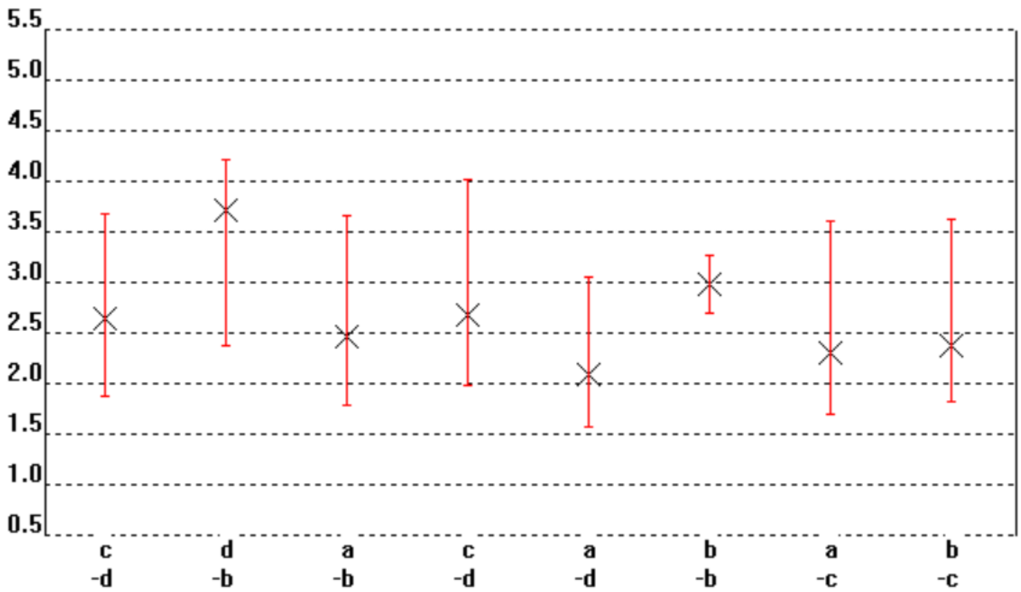

Figure B.13: Median and quartiles of block 2 (legato).

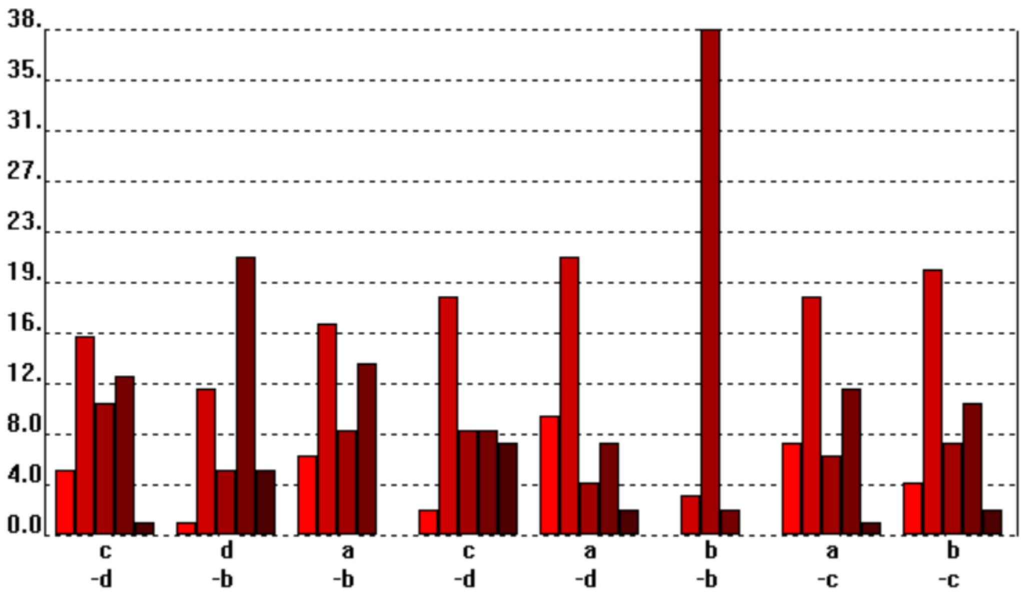

Figure B.14: Histogram of block 2 (legato).

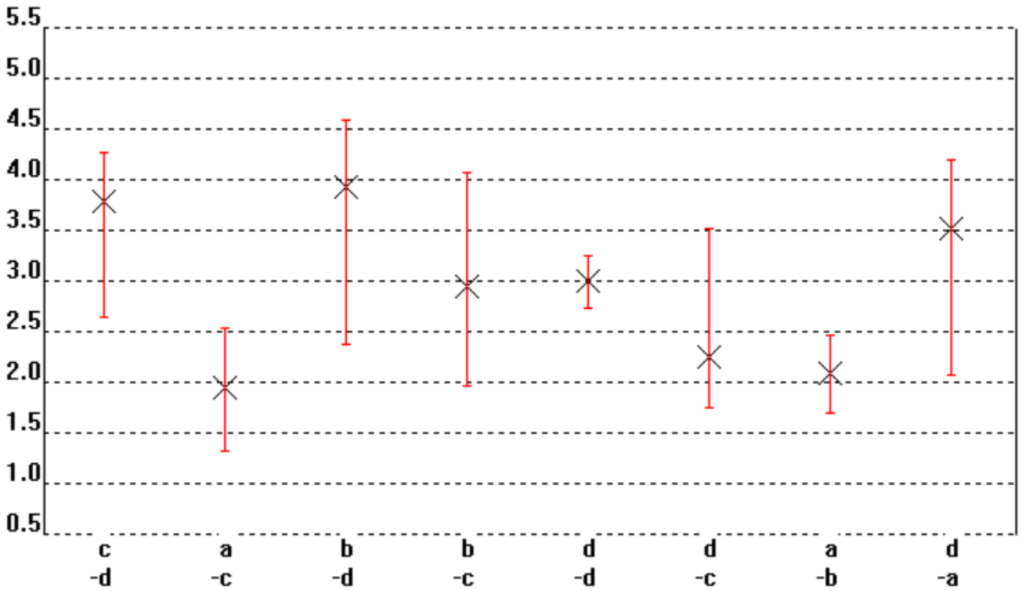

Figure B.15: Median and quartiles of block 3 (precision in spicatto).

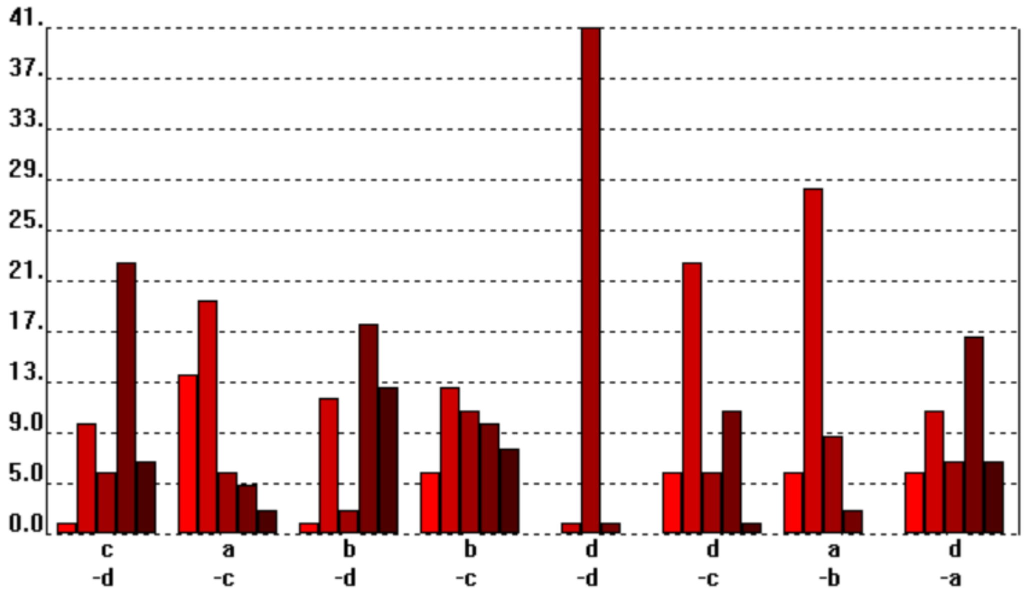

Figure B.16: Histogram of block 3 (precision in spicatto).

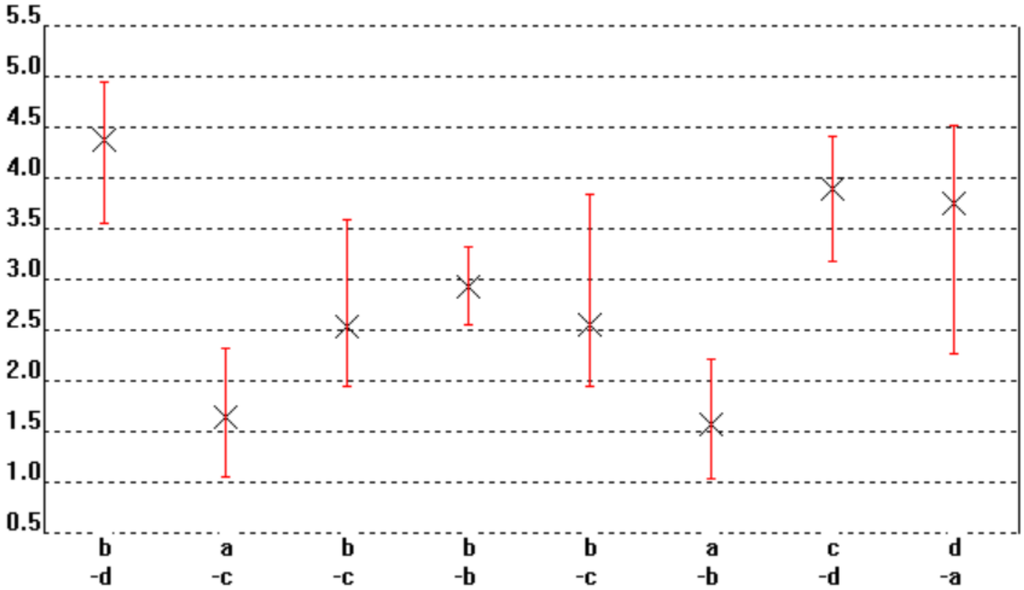

Figure B.17: Median and quartiles of block 4 (expression).

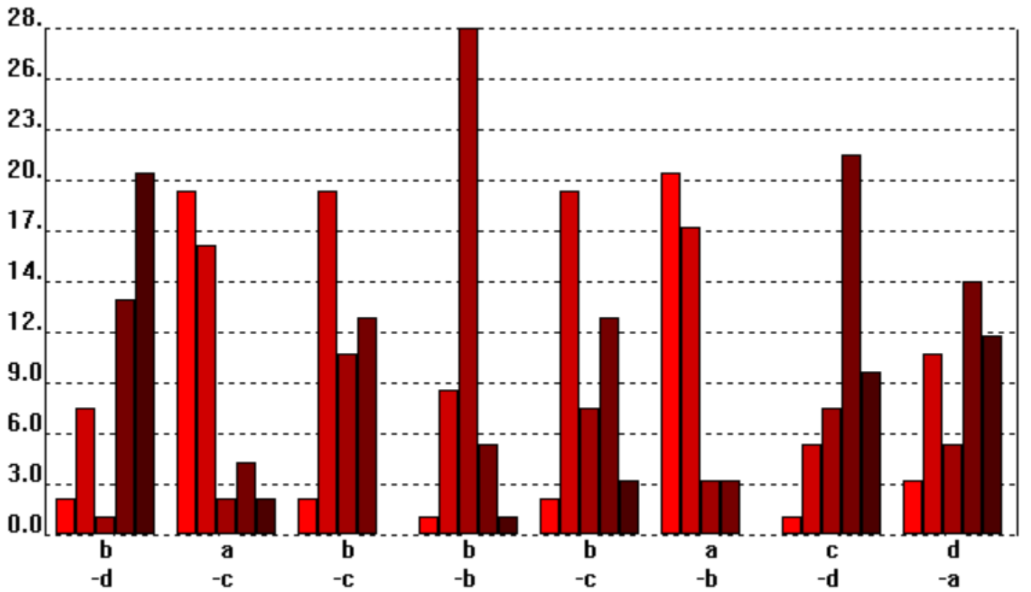

Figure B.18: Histogram of block 4 (expression).

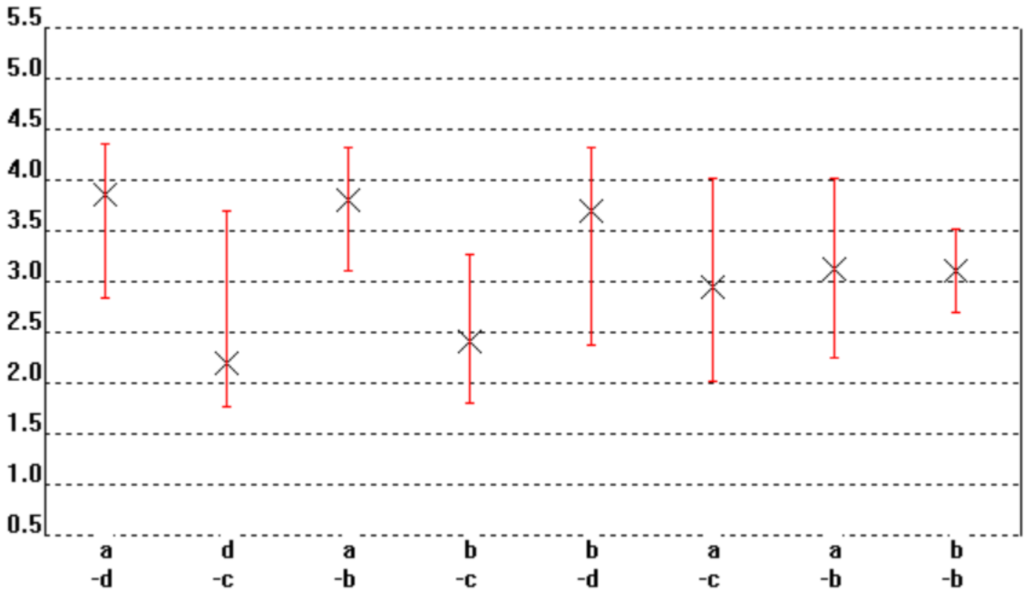

Figure B.19: Median and quartiles of block 5 (vibrato).

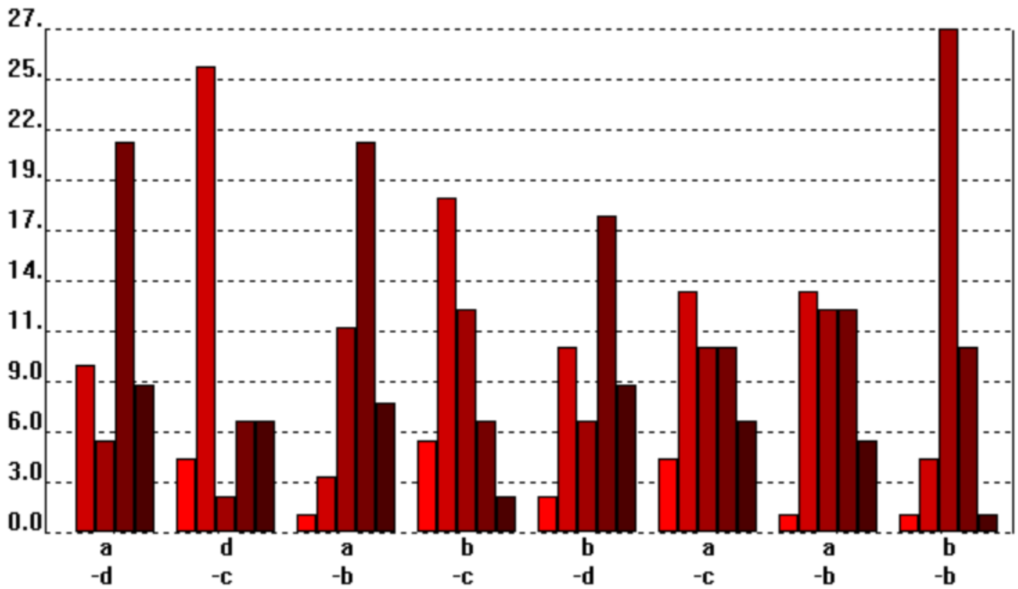

Figure B.20: Histogram of block 5 (vibrato).

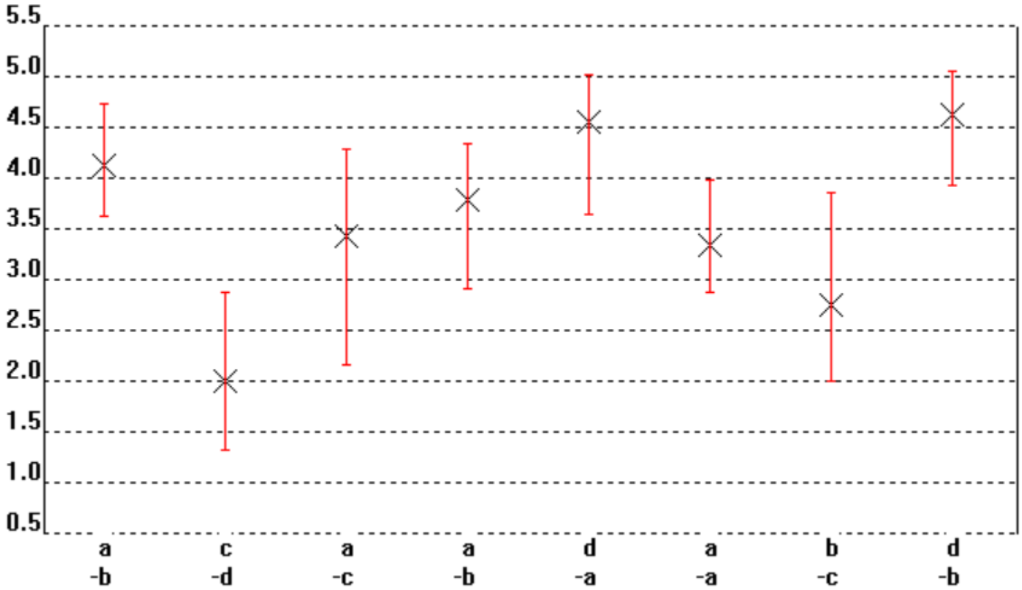

Figure B.21: Median and quartiles of block 6 (crescendo – decrescendo).

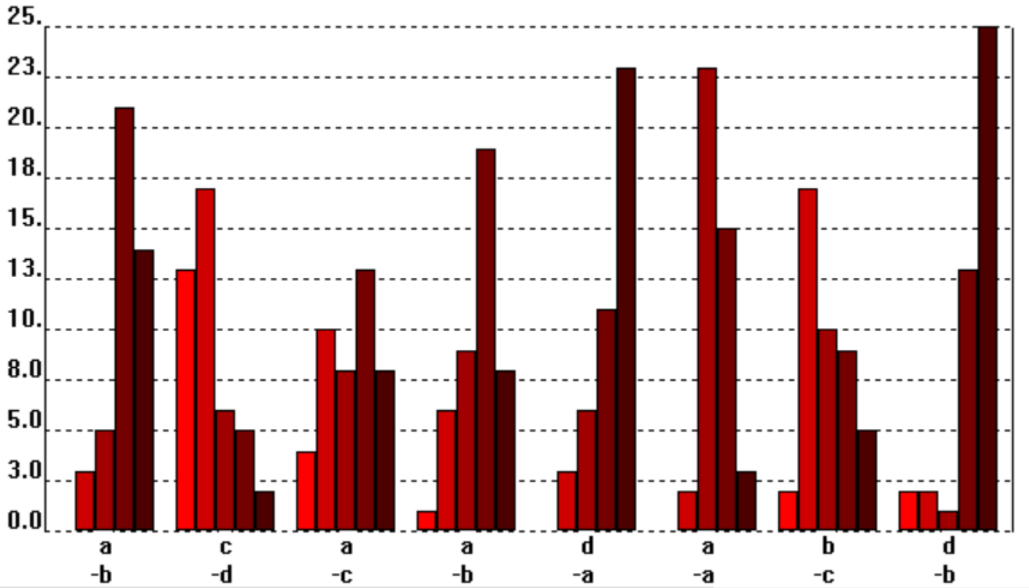

Figure B.22: Histogram of block 6 (crescendo – decrescendo).

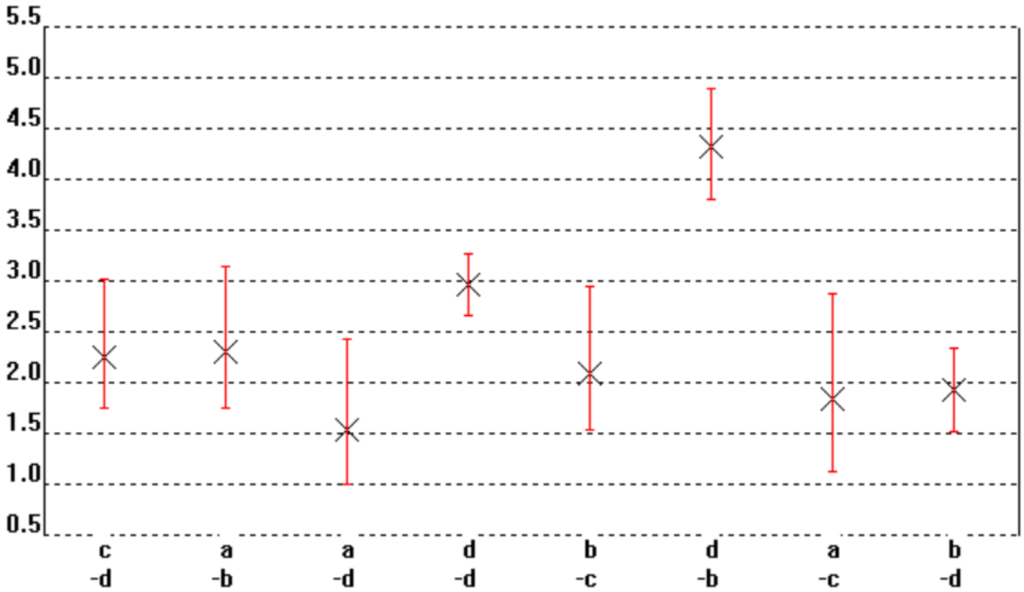

Figure B.23: Median and quartiles of block 7 (timbre change).

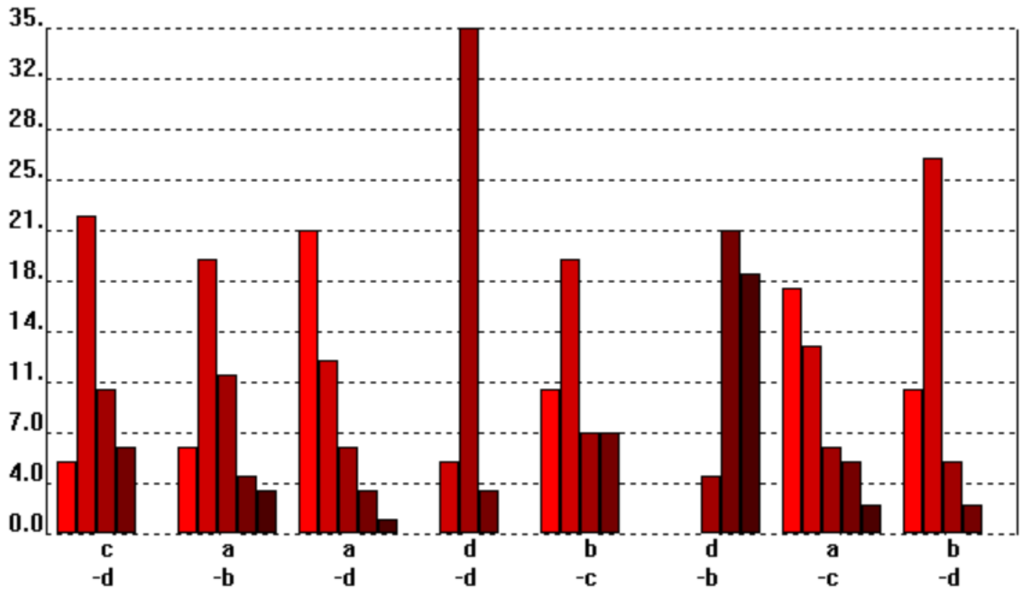

Figure B.24: Histogram of block 7 (timbre change).

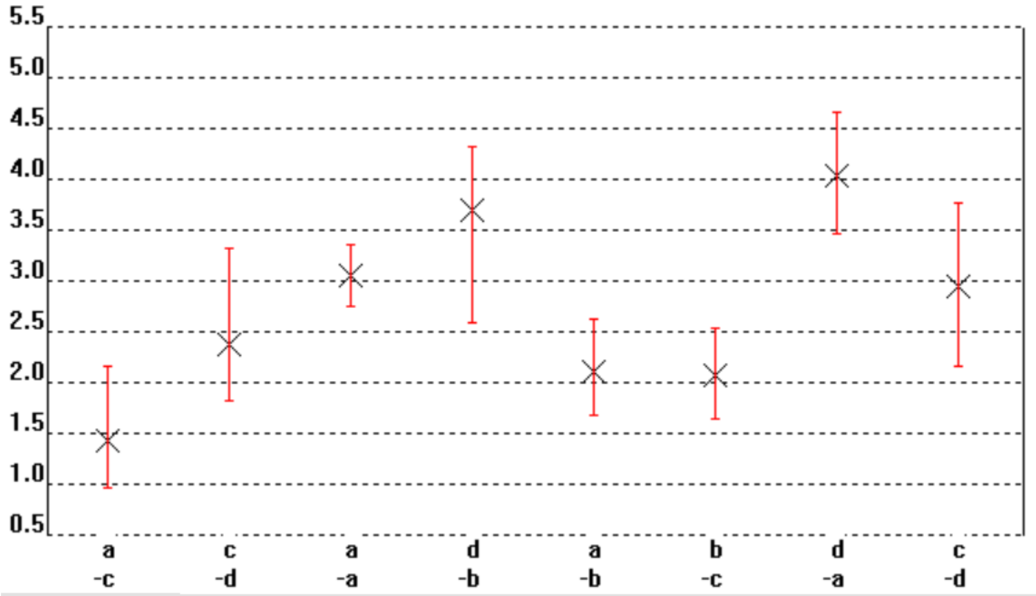

Figure B.25: Median and quartiles of block 8 (rhythmical precision).

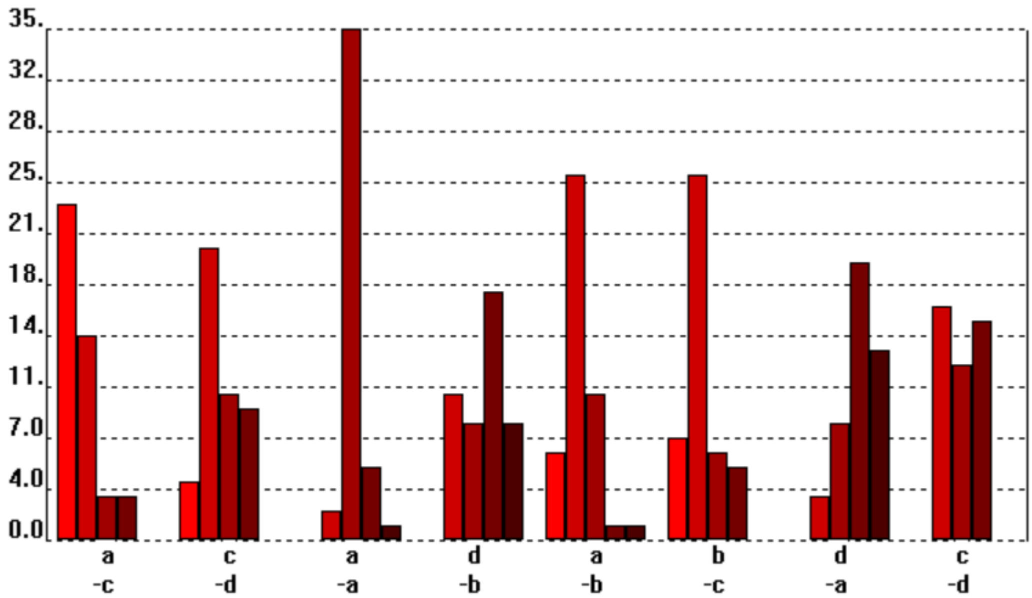

Figure B.26: Histogram of block 8 (rhythmical precision).

Appendix C

Code

Code examples implemented in MaxMSP 4.6[1] can be found on the enclosed CD-Rom. A pdf document of the thesis is also found on the CD-Rom. System requirements for using the MaxMSP patches and information on how to install these are provided in the "Read me MaxMSP-patches" file.

The folder "MaxMSP-patches" includes the following patches:

6.2.2 SD_SubtrSynth

6.2.3 SD_simplFMSynth

6.2.4 SD_MC_FMSynth

6.2.5 SD_MultModFM

6.2.6 SD_feedbFM

6.2.7 SD_feedbPM

6.2.7 SD_MultModPM

[1] Cycling74 (2010b)

6.2.7 SD_simplPMSynth

6.2.9 Granulation groove~

6.2.9 Granulation munger~

6.2.9 Granulation tapout~

6.6 CBVA

Additionally, the folder includes a folder called "Put in folder externals". This folder contains externals necessary to run the MaxMSP patches. It should be put into the "externals" folder of the current Max application in order to open the above-listed patches correctly. The MaxMSP patches make use of the third-party externals munger~, which is part of the PerColate collection (Trueman & DuBois, 2006), fiddle~ (Puckette, Apel, & Zicarelli, 2007), smoother~ (Jehan, Smith, & Zbyszynski, 2007), f0.wrap~ (Olofsson, 2009), and phasor.shift~ and window functions, which are part of the Granular Toolkit of Nathan Wolek (2002).

The folder "MaxMSP-standalone-patches" contains all the above mentioned MaxMSP patches as standalone applications. These applications can be run without the need to use an installed version of MaxMSP. System requirements for using the MaxMSP patches and information on how to use these are provided in the "Read me MaxMSP-standalone-patches" file.

30 sound examples are available and can be used to test the synthesis patches. Sound examples no. 1 to 10 and 21 to 35 were performed by Cornelius Poepel, and sound examples no. 11 to 20 were performed by Günter Marx. Staff notation of the musical patterns 1 to 20 can be found below for quick reference and selection.

Figure C.1: Musical score of 01.Pattern.

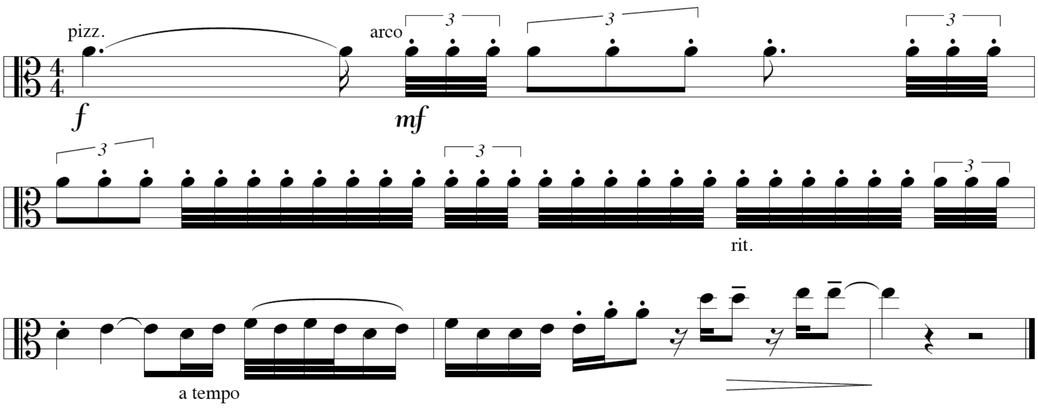

Figure C.2: Musical score of 02.Pattern.

Figure C.3: Musical score of 03.Pattern.

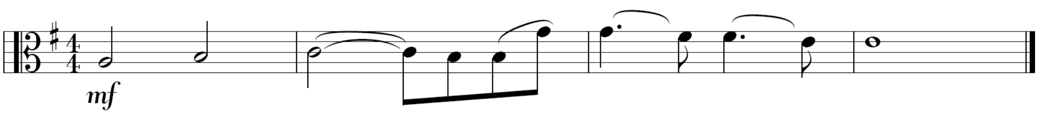

Figure C.4: Musical score of 04.Pattern.

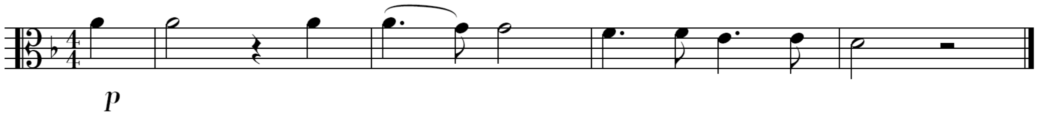

Figure C.5: Musical score of 05.Pattern.

Figure C.6: Musical score of 06.Pattern.

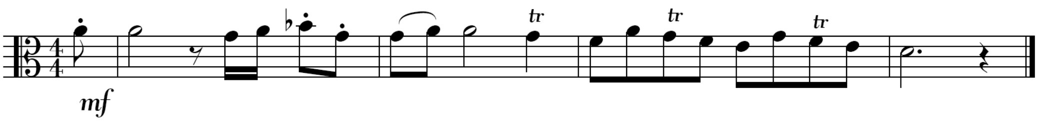

Figure C.7: Musical score of 07.Pattern.

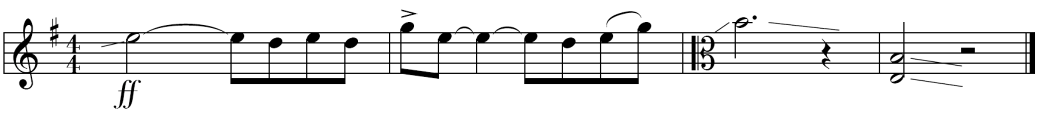

Figure C.8: Musical score of 08.Pattern.

Figure C.9: Musical score of 09.Pattern.

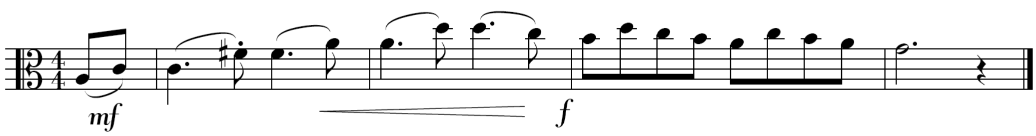

Figure C.10: Musical score of 10.Pattern

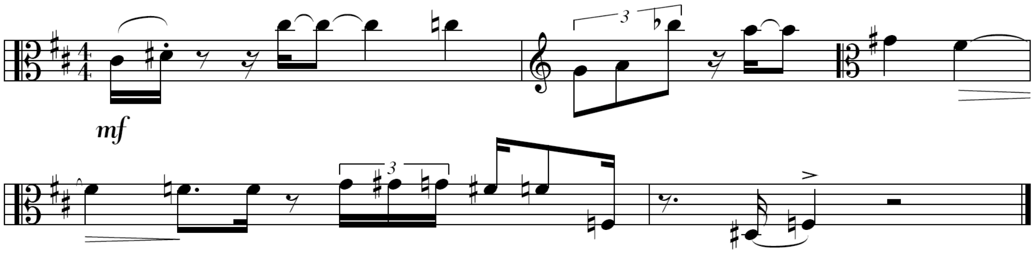

Figure C.11: Musical score of 11.Pattern.

Figure C.12: Musical score of 12.Pattern.

Figure C.13: Musical score of 13.Pattern.

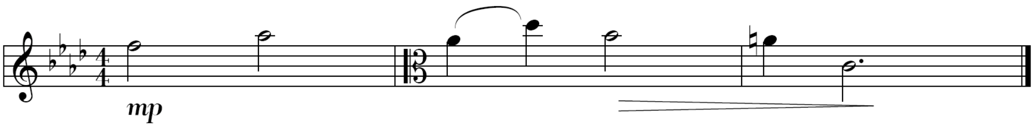

Figure C.14: Musical score of 14.Pattern.

Figure C.15: Musical score of 15.Pattern.

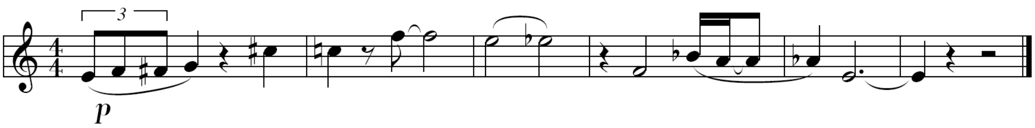

Figure C.16: Musical score of 16.Pattern.

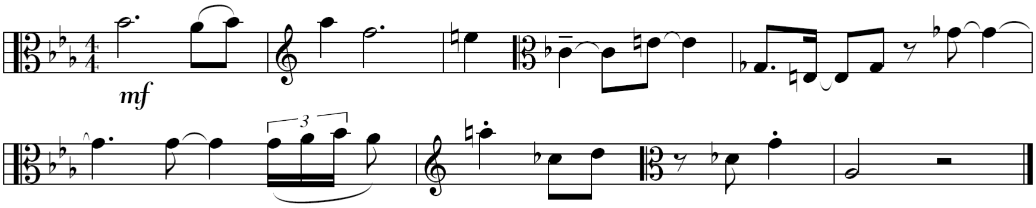

Figure C.17: Musical score of 17.Pattern.

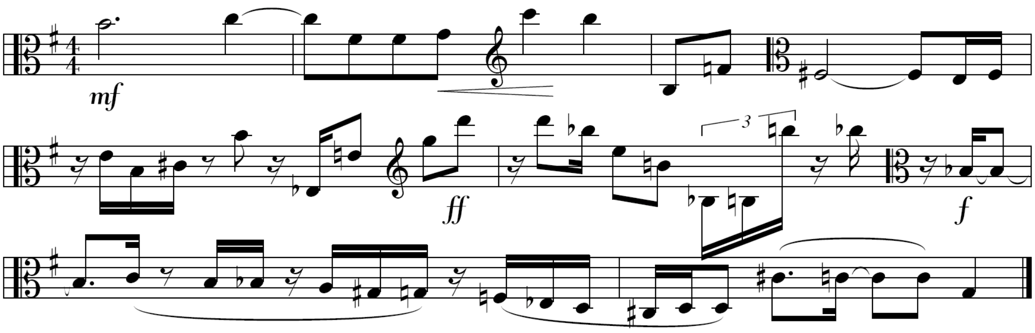

Figure C.18: Musical score of 18.Pattern.

Figure C.19: Musical score of 19.Pattern.

Figure C.20: Musical score of 20.Pattern.

Appendix D

Bow Tracking System

In addition to the extracted parameters of the audio signal, discrete bowing parameters can be of use to the ASDSS approach. Originally I wanted to build a Hyperbow (see section 4.3.5). However, I was not able to find any detailed documentation of technical specifications, circuit boards etc., and it was not possible to get any assistance from the developer for an implementation.[1] By contrast the digital Stradivarius project's hardware (Schoner, 2000b) was found to be well documented.[2] Bernd Schoner gave permission to build a version of this and offered to answer related questions.

The system tracks bow position, speed, and pressure. Bow pressure is measured by a pressure sensor under the index finger. In addition, left finger position on the finger board is measured. The bow tracking system uses the technology of electric field sensing (Paradiso & Gershenfeld, 1997).

[1] Sadly, Diana Young did not have the time to offer help (email communication, September 25, 2003).

[2] Schoner (2010)

Three antennae are mounted horizontally on the bow. The antennae send square-waves with different frequencies which are received by another antenna mounted behind the bridge of the instrument (Schoner, 2000b, p. 173).

An analogue square-wave oscillation with fixed frequency (44 kHz) feeds antenna no. 1. Measurement of this frequency's amplitude in the receiving antenna provides information on bow-bridge distance. Antenna no. 2 is a long resistor (in the present case, a chain of soldered resistors). At the tip of the bow, a square-wave oscillator (35 kHz) is mounted and connected to the resistor chain. At the frog of the bow, another square-wave oscillation (25 kHz) is fed into the antenna. Both frequencies are received by the bridge antenna. If the bow is closer to the tip, a greater amount of the tip frequency (35 kHz) is received by the bridge antenna. If the bow is closer to the frog a greater amount of frog frequency (25 kHz) is received. The pressure sensor modulates the frequency of the third oscillator connected to antenna no. 3. A pitch tracker decodes the pressure intensity received by means of a varying frequency in the bridge antenna.

A board was built to band-pass- ilter the oscillations coming from the bridge antenna. In addition, the amplitude of the filtered oscillations was measured on the board. Figures D.1 presents the board mounted on a test platform. More images of the board can be found on the DVD in chapter “Images” (images no. 3- 7) in Appendix E.2.

While testing this measurement device, I came to the conclusion that it would be easier and more adaptable if the filtering and amplitude measuring

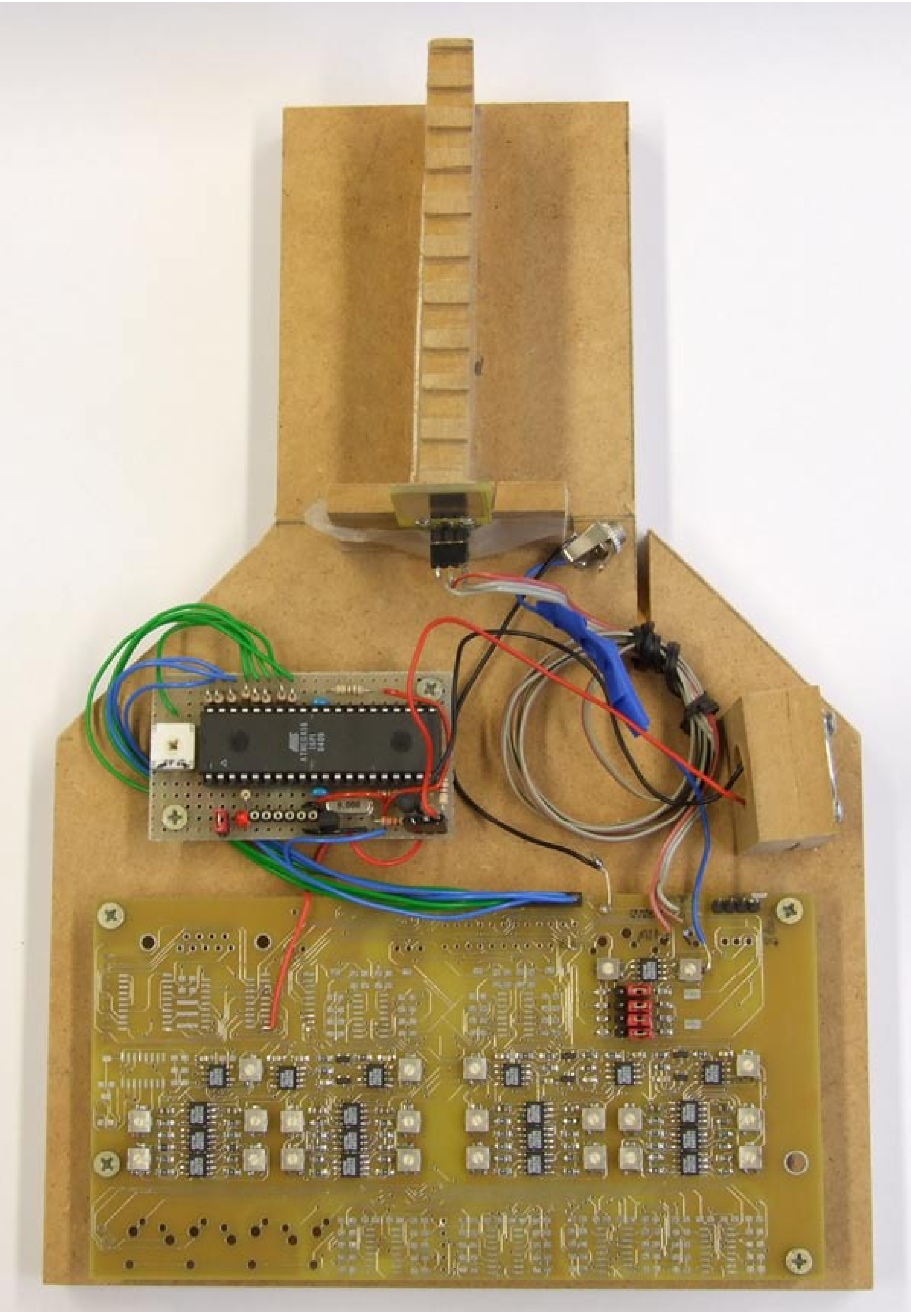

Figure D.1: Mainboard, MIDI translator board and receiving antenna mounted on dummy string instrument.

were to be carried out inside the computer. Therefore, the bow antenna data was analysed using FFT filters which were set to the frequencies of the bow oscillators and the amplitudes in the frequency bins were tracked.

The following problems occurred: the frequencies and amplitudes of the bow antennae were not precisely stable and were prone to interference. To achieve more stable frequencies and amplitudes in the bow electronics, the analogue oscillators (square-waves) were replaced by digital oscillators. The oscillations were generated with an ATMEL ATmega8 Microcontroller.[3] In addition, amplitude measurement of the received oscillations was carried out in the computer instead of on the board. This meant that only the antenna amplifier of the main board was used. Modifying the system in this way resulted in greater flexibility to deal with the signals sent by the bow.

In order to gain experience with the measuring device, I built a test set-up which made it possible to carry out repeated movements with a dummy string instrument and a dummy bow. Unfortunately the system did not work as reliably as expected. One of the reasons for this was that the electrical field was relatively easy to disturb, and the disturbance of the field was a recurrent problem. A significant amount of recalibration was necessary after short periods of time. The bowing hand could also disturb the field itself just by bowing, so that it was unclear whether the bowing had changed its direction (from up to down bow) or whether the bowing hand was disturbing

[3] The ATMEL-based oscillators were done by me and engineer Martin Nawrath. The frequencies of the oscillations were set as follows: frog 35 kHz, tip 30 kHz, bow-bridge distance 25 kHz.

the field. While it is possible to solve this problem with research in electrical engineering, I decided not to do so, because the development of electrical engineering electric field sensing technology would lie beyond the scope of this dissertation.

Another problem was that other factors to be measured were not registered exactly by the measurement device. Bow position (bow-bridge distance), for example, relates only to where the bow hair meets the string. The antenna mounted on the bow stick, however, can change its distance from the antenna on the bridge by turning the bow around its longitudinal axis, while the bow-bridge distance of the bow hair remains the same. Because of this and the above mentioned experiences, I decided to use only audio signal feature extraction in this research.

Appendix E

DVD and Audio CDs

E.1 Audio CDs

The two audio CDs present sounds created with ASDSS technology and compositions that make use of ASDSS. Tracks 1 to 30 of CD 1 are built with the MaxMSP patch *6.6 CBVA* (see Appendix C). The preset numbers written in the description of each track are related to the presets of this MaxMSP patch. Tracks 31 to 44 of CD 1 make use of the MaxMSP patch *6.2.9 Granulation munger~* (see Appendix C). The preset numbers written in the description of each track are related to the presets of this MaxMSP patch. Tracks 1 to 26, 31 to 36 and 39 to 42 were performed by Cornelius Poepel. Tracks 27 to 30 and 43 to 44 were performed by Günter Marx.

The following abreviations are used:
SD: Signal-Driven
PD: Parameter-Driven
SFM: Simple FM (frequency modulation)

MSSM: Modified Single Sideband Modulation
SS: Subtractive Synthesis

Tracks of audio CD 1:

Sounds examples using ASDSS methods:

Track 1: SD SFM, preset no. 2
Track 2: Pickup signal of previous example
Track 3: SD SFM, preset no. 3
Track 4: Pickup signal of previous example
Track 5: SD SFM, preset no. 4
Track 6: Pickup signal of previous example
Track 7: SD SFM, preset no. 16
Track 8: Pickup signal of previous example
Track 9: SD SS, preset no. 6
Track 10: Pickup signal of previous example
Track 11: SD SS, preset no. 10
Track 12: Pickup signal of previous example
Track 13: SD SFM & PD SFM & SD SS, preset no. 11
Track 14: Pickup signal of previous example
Track 15: SD SFM & PD SFM & SD SS, preset no. 12
Track 16: Pickup signal of previous example
Track 17: SD SFM & SD SS, preset no. 17
Track 18: Pickup signal of previous example
Track 19: SD SFM & PD SFM fed into SD SS, preset no. 20
Track 20: Pickup signal of previous example
Track 21: SD SFM & PD SFM fed into SD SS, preset no. 21
Track 22: Pickup signal of previous example
Track 23: MSSM & SD SS, preset no. 22
Track 24: Pickup signal of previous example
Track 25: MSSM fed into SD SS, preset no. 23
Track 26: Pickup signal of previous example
Track 27: SD SFM & PD SFM & MSSM & MSSM fed into SD SS, preset no. 24
Track 28: Pickup signal of previous example
Track 29: SD SFM & PD SFM & SD SS, preset no. 29
Track 30: Pickup signal of previous example
Track 31: Granulation, preset no. 1
Track 32: Pickup signal of previous example

Track 33: Granulation, preset no. 2
Track 34: Pickup signal of previous example
Track 35: Granulation, preset no. 3
Track 36: Pickup signal of previous example
Track 37: Granulation, preset no. 5
Track 38: Pickup signal of previous example
Track 39: Granulation, preset no. 6
Track 40: Pickup signal of previous example
Track 41: Granulation, preset no. 7
Track 42: Pickup signal of previous example
Track 43: Granulation, preset no. 10
Track 44: Pickup signal of previous example

Compositions using ASDSS methods:

Track 45: Recording of the composition *Torque* (2006) by Alexander J. Harker, for electric violin and MaxMSP. Written for Cornelius Poepel's electronic violin system. Violin: Alexander J. Harker.
Track 46: Recording of the composition *Chemical Sunset* (2007) by Carter Williams, for two electric violins using Audio Signal-Driven Sound Synthesis and video. First performance, June 26, 2007, Academy of Media Arts, Cologne. Violins: Günter Marx and Carter Williams.

Tracks of audio CD 2:

Sound examples presented to listeners in the study (see section 7.4).
Track 1: Block 1 (clearness)
Track 2: Block 2 (legato)
Track 3: Block 3 (precision in spicatto)
Track 4: Block 4 (expression)
Track 5: Block 5 (vibrato)
Track 6: Block 6 (crescendo – decrescendo)
Track 7: Block 7 (timbre change)
Track 8: Block 8 (rhythmical precision)
Track 9: Example of changing bow position to explain the meaning of bow position to the test subjects.

E.2 DVD

The DVD presents compositions and concerts involving ASDSS technology. It also presents string players testing ASDSS sounds.

Chapters of DVD:

Chapters 1-4 (ASDSS Examples): The author performing musical phrases with varying bow positions, speed, bowings, and finger pressure using ASDSS sounds.
Chapter 5 (Chemical Sunset): Composition *Chemical Sunset* for two electric violins using Audio Signal-Driven Sound Synthesis and video, composer: Carter Williams. First performance on June 26, 2007, in the 1st concert of the group hot_strings SIG at the Academy of Media Arts, Cologne. Violins: Günter Marx and Carter Williams.
Chapter 6 (ASDSS Instrument Testing 1): Gerardo Vitale (traditional viola player) testing an audio signal-driven sound.
Chapters 7-9 (ASDSS Instrument Testing 2): Participants of NIME04 testing audio signal-driven sounds. Keynote speaker Robert Moog was interested and listened to the test as well (chapter 9).
Chapter 10 (G. Marx at 5th SIG meeting): Excerpt of the 5th meeting of hot_strings SIG (October 23, 2006, IRCAM, Paris). Günter Marx performing a study etude including audio signal-driven sounds.
Chapter 11 (Discussion on ASDSS): Discussion on audio signal-driven sounds after Günter Marx had finished his improvisation (5th meeting of hot_strings

SIG).

Chapter 12 (digiStrings): Performance excerpts from the duo *digiStrings*. E-cello and cello: Torsten Harder; E-viola using ASDSS sounds: Cornelius Pöpel.

The chapter “Images” contains pictures documenting the Harms viola, the bow tracking system (see Appendix D), the 1st concert of hot_strings SIG, and the 6th meeting of hot_strings SIG which took place on the June 27, 2007, at the Academy of Media Arts, Cologne, Germany.

References

Ableton. (2010). *Website of the company Ableton.* Retrieved 16 December 2010, from http://www.ableton.com

Adrien, J.-M. (1991). *The missing link: Modal synthesis* (G. D. Poli, A. Picciali, & C. Roads, Eds.). Cambridge, MA, USA: MIT Press.

Aksnes, H., Godøy, R. I., Kvifte, T., & Ruud, E. (2007). *The musical gestures project 2004 - 2007* (Tech. Rep.). Oslo, Norway: Department of Music and Theatre, University of Oslo.

American Psychological Association (Ed.). (2009). *Publication manual of the american psycholgical association* (6th ed.). Washington, DC: American Psychological Association.

Arfib, D., Couturier, J. M., Kessous, L., & Verfaille, V. (2002). Strategies of mapping between gesture data and synthesis model parameters using perceptual spaces. *Organised Sound*, *7*(2), 127–144.

Ascension. (2010). *Space Pad.* Retrieved 8 December 2010, from http://www.ascension-tech.com/realtime/SpacePad.php

Askenfelt, A. (1986). Measurement of bow motion and bow force in violin playing. *Journal of the Acoustical Society of America*, *80*(4), 1007–

1015.

Bahn, C. (2010). *SBASS – sensor bass.* Retrieved 14 November 2010, from `http://www.arts.rpi.edu/crb/Activities/sbass.htm`

Banff Centre. (2009). *Department of Music & Sound, The Banff Centre.* Retrieved 10 October 2010, from `http://www.banffcentre.ca/music/`

Banken, V. (1993). *Die computerunterstützte Durchführung und Auswertung von Hörversuchen [The computer-aided conducting and analysis of listening studies (translation by the author)].* Diploma thesis, University of Applied Sciences Düsseldorf, Düsseldorf, Germany.

Bargar, R., Choi, I., Das, S., & Goudeseune, C. (1994). Model-based interactive sound for an immersive virtual environment. In *Proceedings of the 1994 international computer music conference* (pp. 471–474). Aarhus, Denmark.

Beller, S. (2004). *Empirisch forschen lernen [Learning empirical research (translation by the author)].* Bern, Switzerland: Verlag Huber.

Bevilacqua, F., Rasamimanana, N. H., Fléty, E., Lemouton, S., & Baschet, F. (2006). The augmented violin project: research, composition and performance report. In *Proceedings of the 2006 conference on new interfaces for musical expression* (pp. 402–406). Paris, France.

Biesenbender, V. (1992). *Von der unerträglichen Leichtigkeit des Instrumentalspiels [The unbearable lightness of instrumental performance (translation by the author)].* Aarau, Switzerland: Musikedition Nepomuk.

Brun, M. L. (1979). Digital waveshaping synthesis. *Journal of the Audio Engineering Society*, *27*(4), 250–266.

Buchla. (2010). *Website of the current Buchla Lightning system.* Retrieved

31 December 2010, from http://www.buchla.com/lightning3.html

Buxton, W., Buchla, D., Chafe, C., Machover, T., Mathews, M., Moog, B., et al. (2000). Round table [CD-Rom]. In M. M. Wanderley & M. Battier (Eds.), *Trends in gestural control of music* (pp. 415–437). Paris, France: IRCAM Centre Pompidou.

Caceres, J.-P. (2007). *CCRMA SoundWIRE.* Retrieved 28 October 2010, from https://ccrma.stanford.edu/groups/soundwire/

Cadoz, C. (1988). Instrumental gesture and musical composition. In *Proceedings of the 1988 international computer music conference* (pp. 1–12). Cologne, Germany.

Cadoz, C., Luciani, A., & Florens, J. L. (1984). Responsive input devices and sound synthesis by simulation of instrumental mechanisms: The cordis system. *Computer Music Journal*, *8*(3), 60–73.

Cadoz, C., & Wanderley, M. M. (2000). Gesture - music [CD-Rom]. In M. M. Wanderley & M. Battier (Eds.), *Trends in gestural control of music* (pp. 28–65). Paris, France: IRCAM Centre Pompidou.

CAML. (2010). *Website of the Computational Acoustic Modeling Laboratory.* Retrieved 15 November 2010, from http://www.music.mcgill.ca/caml/

Canazza, S., Roda, A., & Orio, N. (1999). A parametric model of expressiveness in musical performance based on perceptual and acoustic analyses. In *Proceedings of the 1999 international computer music conference* (pp. 379–382). Hong Kong, China.

Casals, P. (1983). *Pablos Casals. Licht und Schatten auf einem langen Weg. Erinnerungen aufgezeichnet von Albert E. Kahn [Pablos Casals.*

Light and shadow on a long road. Memories recorded by Albert E. Kahn (translation by the author)]. Frankfurt am Main, Germany: Fischer Taschenbuch Verlag.

CCRMA. (2010). *Center for Computer Research in Music and Acoustics.* Retrieved 31 December 2010, from https://ccrma.stanford.edu/

Chadabe, J. (1997). *Electric sound: The past and promise of electronic music.* Upper Saddle River, NJ, USA: Prentice Hall.

Chafe, C. (1993). Tactile audio feedback. In *Proceedings of the 1993 international computer music conference* (pp. 66–69). Vaseda, Japan.

Chafe, C., & O'Modhrain, M. S. (1996). Musical muscle memory and the haptic display of performance nuance. In *Proceedings of the 1996 international computer music conference* (pp. 428–431). Hong Kong, China.

CHI. (2010). *ACM Conference on Human Factors in Computing Systems.* Retrieved 31 December 2010, from http://www.chi2010.org/

Chowning, J. M. (1973). The synthesis of complex audio spectra by means of frequency modulation. *Journal of the Audio Engineering Society*, *21*(7), 526–534.

CNMAT. (2010). *Website of the Center for New Music and Audio Technologies.* Retrieved 28 November 2010, from http://cnmat.berkeley.edu/

ConGAS. (2007). *Gesture Controlled Audio Systems.* Retrieved 31 December 2010, from http://www.cost.esf.org/domains_actions/ict/Actions/Gesture_Controlled_Audio_Systems

Cook, P. R., & Scavone, G. P. (2009). *The Synthesis ToolKit in C++ (STK).* Retrieved 15 November 2009, from https://ccrma.stanford

.edu/software/stk/

Crab, S. (2005). *Computer Music: Music1-V & GROOVE.* Retrieved 31 December 2010, from http://120years.net/machines/software/index.html

Csíkszentmihályi, M. (1990). *Flow: The psychology of optimal experience.* New York: Harper and Row.

Csound. (2010). *Csound.* Retrieved 15 November 2010, from http://www.csounds.com/

Cycling74. (2010a). *FAQ: Max 4, History and Background.* Retrieved 8 December 2010, from http://cycling74.com/support/faq_max4/#1

Cycling74. (2010b). *Website of the company Cycling74.* Retrieved 8 December 2010, from http://cycling74.com/

Degele, N. (2000). *Informiertes Wissen. Eine Wissenssoziologie der computerisierten Gesellschaft [Informed knowledge. A sociology of knowledge of computerised society (translation by the author)].* Frankfurt am Main, Germany: Campus Fachbuch.

Farwell, N. (2001). *being what i am: doing what i do, manifesto of a composer.* Doctoral dissertation, University of East Anglia, East Anglia, UK.

Fels, S. (2004, April). Designing for intimacy: Creating new interfaces for musical expression. *Proceedings of the IEEE*, *92*(4), 672–685.

Fels, S., Gadd, A., & Mulder, A. (2002). Mapping transparency through metaphor: towards more expressive musical instruments. *Organised Sound*, *7*(2), 109–126.

Fischer, S. (1997). *Basics, 300 exercises and practice routines for the violin*

(No. 7440). London, UK: Edition Peters.

Fischer, S. (2010). *Biography of Simon Fischer.* Retrieved 31 December 2010, from http://www.simonfischeruk.com/page7.htm

Freed, A. (2008). *CNMAT resonators~ MSP object.* Retrieved 15 December 2010, from http://cnmat.berkeley.edu/downloads

Freed, A., Uitti, F. M., Zbyszynski, M., & Wessel, D. (2006, June). Augmenting the cello. In *Proceedings of the 2006 conference on new interfaces for musical expression* (pp. 409–413). Paris, France.

Galamian, I. (1962). *Principles of violin playing and teaching.* Englewood Cliffs, N. J., USA: Prentice-Hall, Inc.

Geiger, C., Reckter, H., Paschke, D., Schultz, F., & Poepel, C. (2008). Towards participatory design and evaluation of theremin-based musical interfaces. In *Proceedings of the 2008 conference on new interfaces for musical expression* (pp. 303–306). Genova, Italy.

Gill, C. (1997). The stomping ground: Musitronics, mu-tron, and the gizmotron. *Vintage Guitar Magazine*, *11*(12).

Goto, S. (1999). The aesthetics and technological aspects of virtual musical instruments: The case of the SuperPolm MIDI violin. *Leonardo Music Journal*, *9*, 115–120.

Goudeseune, C. (2001). *Composing with parameters for synthetic instruments.* Doctoral dissertation, University of Illinois at Urbana-Champaign, Urbana, IL, USA.

Goudeseune, C. (2002). Interpolated mappings for musical instruments. *Organised Sound*, *7*(2), 85–96.

Goudeseune, C. (2010). *E-Violin Study* [Video documentation of the composition]. Retrieved 31 December 2010, from http://zx81.isl.uiuc.edu/camilleg/evln1.qt

Goudeseune, C., Garnett, G., & Johnson, T. (2001). Resonant processing of instrumental sound controlled by spatial position. In *Publications of the 2001 workshop on new instruments for musical expression.* Seattle, USA. Retrieved 8 December 2010, from http://zx81.isl.uiuc.edu/camilleg/nime01.pdf

Graesser, H. (1996). *Electric violins.* Frankfurt am Main, Germany: Verlag Erwin Bochinsky.

Greenbaum, J., & Kyng, M. (1992). *Design at work, cooperative design of computer systems.* Mahwah, NJ, USA: Lawrence Erlbaum Associates, Inc.

Gwosdz, T. (2007). *Vergleichende Studie von Streichersynthesizern [A comparative study of string synthesisers (translation by the author)].* Diploma thesis, University of Applied Sciences Düsseldorf, Düsseldorf, Germany.

Harms, A. (2010). *Website of the luthier Arwed Harms.* Retrieved 31 December 2010, from http://www.arwedharms.com/

Harnoncourt, N. (1992). Interview by Jürgen Kesting, Track 14, [CD]. In *Nikolaus Harnoncourt: Das Portrait [Nikolaus Harnoncourt: The portrait (translation by the author)].* no statement of city: Teldec Classics International GmbH.

Hazzan, O., & Kramer, J. (2007, January). Abstraction in computer science & software engineering: A pedagogical perspective. *Frontier Journal*, *4*(1), 6–14.

Heidegger, M. (1967). *Sein und Zeit [Being and time (translation by the author)]* (11th ed.). Tübingen, Germany: Max Niemeyer Verlag.

Hermida, A. (2002). *E-strings for the future musician.* Retrieved 8 December 2010, from http://news.bbc.co.uk/2/hi/science/nature/1930950.stm

Hinton, G. (2010). *A guide to the EMS product range: 1969 to 1979.* Retrieved 15 December 2010, from http://www.ems-synthi.demon.co.uk/emsprods.html

Hiraga, R., Bresin, R., Hirata, K., & Katayose, H. (2004). Rencon 2004: Turing test for musical expression. In *Proceedings of the 2004 conference on new interfaces for musical expression* (pp. 120–123). Hamamatsu, Japan.

Hodgson, P. (1958). *Motion study and violin bowing.* Urbana, IL, USA: American String Teachers Association.

Holm, F. (1992). Understanding FM implementations: A call for common standards. *Computer Music Journal*, *16*(1), 34–42.

Hunt, A., & Kirk, R. (2000). Mapping strategies for musical performance [CD-Rom]. In M. M. Wanderley & M. Battier (Eds.), *Trends in gestural control of music* (pp. 382–427). Paris, France: IRCAM Centre Pompidou.

Hunt, A., & Wanderley, M. M. (2002). Mapping performer parameters to synthesis engines. *Organised Sound*, *7*(2), 97–108.

Hypercello. (2010). *Hypercello / Yo-Yo Ma.* Retrieved 14 October 2010, from http://park.org/Events/BrainOpera/Archive/Hyperinstruments/hypercello.html

ICMC. (2009). *International Computer Music Conference.* Retrieved 14 November 2009, from http://www.computermusic.org/page/23/

IEEE. (1998). *IEEE recommended practice for software requirements specifications.* New York, NY, USA: The Institute of Electrical and Electronics Engineers, Inc.

IRCAM. (2010). *Institute for Music/Acoustic Research and Coordination.* Retrieved 31 December 2010, from http://www.ircam.fr/

Isaacs, D. (2003). *Evaluating input devices for musical expression.* Thesis, University of Queensland, Brisbane, Australia.

Jaffe, D. A., & Smith, J. O. (1995). Performance expression in commuted waveguide synthesis of bowed strings. In *Proceedings of the 1995 international computer music conference* (pp. 343–346). Banff, Canada.

Janer, J. (2005a). Feature extraction for voice-driven synthesis. In *Proceedings of 118th audio engineering society convention.* Barcelona, Spain. Retrieved 8 December 2010, from http://mtg.upf.edu/files/publications/9cf4b9-AES2005-jjaner.pdf

Janer, J. (2005b). Voice-controlled plucked bass guitar through two synthesis techniques. In *Proceedings of the 2005 conference on new interfaces for musical expression* (pp. 132–135). Vancouver, Canada.

Janer, J. (2008). *Singing-driven interfaces for sound synthesizers.* Doctoral dissertation, Universitat Pompeu Fabra, Barcelona, Spain.

Japenga, R. (2010). *How to write a software requirements specification.* Retrieved 14 September 2010, from http://www.microtoolsinc.com/Howsrs.php

Jehan, T. (2001). *Perceptual synthesis engine: An audio-driven timbre gen-*

erator. Master's thesis, Massachusetts Institute of Technology, Cambridge, MA, USA.

Jehan, T. (2010). *Max/MPS objects of Tristan Jehan download page.* Retrieved 15 November 2010, from `http://web.media.mit.edu/~tristan/maxmsp.html`

Jehan, T., & Schoner, B. (2001). An audio-driven perceptually meaningful timbre synthesizer. In *Proceedings of the 2001 international computer music conference* (pp. 381–388). Havana, Cuba.

Jehan, T., Smith, W., & Zbyszynski, M. (2007). *smoother~* [MaxMSP third-party external]. Retrieved 31 December 2010, from `http://web.media.mit.edu/~tristan/maxmsp.html`

Jensen, E. (2010). *Website of the company Jensen Musical Instruments.* Retrieved 31 December 2010, from `http://www.halcyon.com/jensmus/`

Jordà, S. (2005). *Digital lutherie, crafting musical computers for new musics' performance and improvisation.* Doctoral dissertation, Departament de Tecnologia, Universitat Pompeu Fabra, Barcelona, Spain.

Jordà, S., Kaltenbrunner, M., Geiger, G., & Bencina, R. (2005). The reactable. In *Proceedings of the 2005 international computer music conference* (pp. 579–582). Barcelona, Spain.

Juslin, P. N. (2001). Communicating emotion in music performance: A review and theoretical framework. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and emotion* (pp. 309–337). Oxford, UK: Oxford Univ. Press.

Juslin, P. N. (2003). Five facets of musical expression: a psychologist's perspective on music performance. *Psychology of Music*, *31*(3), 273–

302.

Kiefer, C., Collins, N., & Fitzpatrick, G. (2008). HCI methodology for evaluating musical controllers: A case study. In *Proceedings of the 2008 conference on new interfaces for musical expression* (pp. 87–90). Genova, Italy.

Kimura, M. (2010). *Website of Mari Kimura.* Retrieved 16 December 2010, from `http://web.me.com/marikimura/Site_2/main.html`

Kirk, R., & Hunt, A. (1998). Computer assisted performance control for audio-visual instruments. In R. Kopiez & W. Auenhagen (Eds.), *Controlling creative processes in music* (Vol. 12). Frankfurt am Main, Germany: Peter Lang GmbH, Europäischer Verlag der Wissenschaften.

Klot, K. von. (2007). Verspielter Klangforscher [Playful sound explorer (translation by the author)]. *mobil, das Magazin der Bahn*(07), 103.

Krämer, S. (1988). *Symbolische Maschinen. Die Idee der Formalisierung in geschichtlichem Abriß [Symbolic machines. A historical survey of the idea of formalisation (translation by the author)].* Darmstadt, Germany: Wissenschaftliche Buchgesellschaft.

Lazzarini, V., Timoney, J., & Lysaght, T. (2007). Adaptive FM synthesis. In *Proceedings of the 10th international conference on digital audio effects 2007* (pp. 21–26). Bordeaux, France.

Levenson, T. (1994). Taming the hypercello. *Sciences*, *34*(4), 15–17.

Levitin, D. J., McAdams, S., & Adams, R. L. (2002). Control parameters for musical instruments: a foundation for new mappings of gesture to sound. *Organised Sound*, *7*(2), 171–189.

Lippe, C. (1994). Real-time granular sampling using the IRCAM signal processing workstation. *Contemporary Music Review*, *10*(2), 149–155.

Lount, P. W. (2004). *What is Smalltalk?* Retrieved 8 December 2010, from `http://www.smalltalk.org/smalltalk/whatissmalltalk.html`

Machover, T. (1991). *Begin Again Again ...* [Musical score]. Milano, Paris: Ricordi.

Machover, T. (1992). *Hyperinstruments: A progress report* (Tech. Rep.). Cambridge, MA, USA: Massachusetts Institute of Technology.

Machover, T. (2002). Instruments, interactivity, and inevitability. In *Proceedings of the 2002 conference on new instruments for musical expression* (p. 1). Dublin, Ireland.

Machover, T. (2003). CD-booklet [CD]. In *Hyperstring trilogy.* Shutesbury, MA, USA: Oxingale Records.

Maramax. (2010). *The Father of Computer Music.* Retrieved 31 December 2010, from `http://www.csounds.com/mathews/`

Marrin, T. (2000). *Inside the conductor's jacket.* Doctoral dissertation, Massachusetts Institute of Technology, Cambridge, MA, USA. Retrieved 8 December 2010, from `http://web.media.mit.edu/~tristan/Classes/MAS.945/Papers/Main/Marrin_Thesis.pdf`

Marrin Nakra, T. (2001). Translating conductor's gestures to sound. In *Proceedings of the 2001 workshop on human supervision and control in engineering and music.* Retrieved 8 December 2010, from `http://www.engineeringandmusic.de/individu/nakrmarr/Nakra-Paper.html`

Mastin, L. (2010). Glossary. In *The basics of philosophy.* Retrieved 23 October 2010, from `http://www.philosophybasics.com/general` _glossary.html

Mateescu, A., & Salomaa, A. (1997). Formal languages: An introduction and a synopsis. In G. Rozenberg & A. Salomaa (Eds.), *Handbook of formal language theory* (Vol. 1, pp. 1–39). New York, NY, USA: Springer Verlag.

Mathews, M. V., & Kohut, J. (1973). Electronic simulation of violin resonances. *Journal of the Acoustical Society of America*, *53*(6), 1620–1626.

McEnery, T., & Wilson, A. (2001). *Corpus linguistics* (2nd ed.). Edinburgh: Edinburgh University Press.

McIntyre, M. E., Schumacher, R., & Woodhouse, J. (1983). On the oscillation of musical instruments. *Journal of the Acoustical Society of America*, *74*(5), 1325–1345.

McMillen, K. A. (2008). Stage-worthy sensor bows for stringed instruments. In *Proceedings of the 2008 conference on new interfaces for musical expression* (pp. 347–348). Genova, Italy.

Menuhin, Y. (1985). *Ich bin fasziniert von allem Menschlichen [I am fascinated by everything that is human, (translation by the author)]* (3rd ed.). München, Germany: R. Piper.

MIDI. (2010). *MIDI Manufacturers Association.* Retrieved 31 December 2010, from http://www.midi.org/

Miranda, E. R., & Wanderley, M. M. (2006). *New digital musical instruments: Control and interaction beyond the keyboard.* Middleton, WI, USA: A-R Editions, Inc.

MIT. (2010). *Massachusetts Institute of Technology.* Retrieved 31 December

2010, from http://www.mit.edu/

Moore, F. R. (1988). The dysfunctions of MIDI. *Computer Music Journal*, *12*(1), 19–28.

Müller, S., & Mazzola, G. (2003). The extraction of expressive shaping in performance. *Computer Music Journal*, *27*(1), 47–58.

Native Instruments. (2010). *Website of the company Native Instruments.* Retrieved 31 December 2010, from http://www.native-instruments.com/

Neuhaus, H. (1967). *Die Kunst des Klavierspiels [The art of piano playing (translation by the author)].* no statement of city: Musikverlage Hans Gerig.

Neumann, J. von. (1955). *Mathematical foundations of quantum mechanics.* Princeton, NJ, USA: Princeton University Press.

Neumann, J. von. (1966). *The theory of selfreproducing automata.* Urbana, IL, USA: University of Illinois Press.

Ng, K., & Nesi, P. (Eds.). (2008). *Interactive multimedia music technologies.* Hershey, PA, USA: Information Science Reference.

Nichols, C. (2002). The vBow: A virtual violin bow controller for mapping gesture to synthesis with haptic feedback. *Organised Sound*, *7*(2), 215–220.

Nichols, C. (2003). *The vBow: An expressive musical controller haptic human-computer interface.* Doctoral dissertation, Stanford University, Department of Music, Stanford, CA.

NIME. (2010). *International Conference on New Interfaces for Expression.* Retrieved 31 December 2010, from http://www.nime.org/

Olofsson, F. (2009). *f0.ext* [Collection of MaxMSP third-party externals]. Retrieved 31 December 2010, from http://www.fredrikolofsson.com/pages/code-max.html

O'Modhrain, M. S. (2000). *Playing by feel: Incorporating haptic feedback into computer-based musical instruments.* Doctoral dissertation, Stanford University, Stanford, CA, USA.

O'Modhrain, M. S., & Chafe, C. (2000). The performer-instrument interaction: A sensory motor perspective. In *Proceedings of the 2000 international computer music conference* (pp. 145–148). Berlin, Germany.

OSC. (2010). *Open Sound Control.* Retrieved 8 December 2010, from http://opensoundcontrol.org/

Overholt, D. (2005a). The overtone violin. In *Proceedings of the 2005 conference on new interfaces for musical expression* (pp. 34–37). Vancouver, Canada.

Overholt, D. (2005b). The overtone violin: A new computer music instrument. In *Proceedings of the 2005 international computer music conference* (pp. 604–607). Barcelona, Spain.

Overholt, D. (2009). *Image of the Overtone Violin.* Retrieved 10 October 2009, from http://www.mat.ucsb.edu/~dano/violin/

Paine, G. (2007). The thummer mapping project (ThuMP). In *Proceedings of the 2007 conference on new interfaces for musical expression* (pp. 70–77). New York, NY, USA.

Paradiso, J. A. (1997). Electronic music: New ways to play. *IEEE Spectrum*, *34*(12), 18–30.

Paradiso, J. A., & Gershenfeld, N. (1997). Musical applications of electric field sensing. *Computer Music Journal*, *21*(2), 69–89.

Pierce, J. R. (1989). *Klang mit den Ohren der Physik [The science of musical sounds(title of the English edition)]* (2nd ed.). Heidelberg, Germany: Spektrum-der-Wissenschaft-Verlagsgesellschaft.

Pitteroff, R., & Woodhouse, J. (1998). Mechanics of the contact area between a violin bow and a string. *Acustica*, *84*, 543–562.

Poepel, C. (1999). *Meta strings.* Diploma thesis, part 2, Musikhochschule der Musikakademie der Stadt Basel, Basel, Switzerland.

Poepel, C. (2004). Synthesized strings for string-players. In *Proceedings of the 2004 conference on new interfaces for musical expression* (pp. 150–153). Hamamatsu, Japan.

Poepel, C. (2005). On interface expressivity: A player-based study. In *Proceedings of the 2005 conference on new interfaces for musical expression* (pp. 228–231). Vancouver, Canada.

Poepel, C. (2008a). Driving sound synthesis with a live audio signal. In K. Ng & P. Nesi (Eds.), *Interactive multimedia music technologies* (pp. 167–194). Hershey, PA, USA: Information Science Reference.

Poepel, C. (2008b). Formalized and non-formalized expression in musical interfaces. In R. Adamns, S. Gibson, & S. M. Arisona (Eds.), *Transdisciplinary digital art: Sound, vision and the new screen* (pp. 358–373). Heidelberg, Germany: Springer.

Poepel, C., & Dannenberg, R. B. (2005). Audio signal driven sound synthesis. In *Proceedings of the 2005 international computer music conference* (pp. 391–394). Barcelona, Spain.

Poepel, C., & Marx, G. (2007). hot_strings SIG. In *Proceedings of the 2007* *conference on new interfaces for musical expression* (pp. 281–284). New York, NY, USA.

Poepel, C., & Overholt, D. (2006). Recent developments in violin-related digital musical instruments: Where are we and where are we going? In *Proceedings of the 2006 conference on new interfaces for musical expression* (pp. 390–395). Paris, France.

Poli, G. D., & Rochesso, D. (1998). Physically based sound modelling. *Organised Sound*, *3*(1), 61–76.

Pont, O. (2010). *Website of the luthier Olivier Pont.* Retrieved 15 November 2010, from `http://www.olivierpont-luthier.com/`

Poupyrev, I., Lyons, M. J., Fels, S., & Blaine, T. (2001). Workshop proposal: New interfaces for musical expression. In *Proceedings of the 2001 workshop on new interfaces for musical expression.* Seattle, WA, USA. Retrieved 8 December 2010, from `http://www.nime.org/2001/docs/proposal.pdf`

Prem, E. (1997). Epistemic autonomy in models of living systems. In P. Husbands & I. Harvey (Eds.), *Proceedings fourth european conference on artificial life* (pp. 2–9). Brighton, UK.

Puckette, M. S. (1991). Combining event and signal processing in the MAX graphical programming environment. *Computer Music Journal*, *15*(3), 68–77.

Puckette, M. S., Apel, T., & Zicarelli, D. D. (1998). Real-time audio analysis tools for Pd and MSP. In *Proceedings of the 1998 international computer music conference* (pp. 109–112). Ann Arbor, MI, USA.

Puckette, M. S., Apel, T., & Zicarelli, D. D. (2007). *fiddle~* [MaxMSP third-party external]. Retrieved 31 December 2010, from `http://crca.ucsd.edu/~tapel/software.html`

Pyka, A. (1997). *Objektorientierte Entwicklung eines Statistiktools in C++ [Object-orientated development of a statistics tool using C++ (translation by the author)].* Diploma thesis, University of Applied Sciences Düsseldorf, Düsseldorf, Germany.

Rasamimanana, N., Fléty, E., & Bevilacqua, F. (2006). Gesture analysis of violin bow strokes. In S. Gibet, N. Courty, & J.-F. Kamp (Eds.), *Gesture in human-computer interaction and simulation: 6th international gesture workshop* (Vol. 3881, pp. 145–155). Berlin, Heidelberg, Germany: Springer.

Risset, J.-C. (1984). Musik und Computer: Der Computer als Kompositionshilfe [Music and computers: The computer as a composition aid (translation by the author)]. *Österreichische Musikzeitschrift*, *39*(9), 442–446.

Roads, C. (1980). Interview with Max Mathews. *Computer Music Journal*, *4*(4), 15–22.

Roads, C. (1996). *The computer music tutorial.* Cambridge, MA, USA: MIT Press.

Rose, J. (2005). *Improvisation & Computers.* Retrieved 31 December 2010, from `http://www.jonroseweb.com/c_articles_improvisation.html`

Rovan, J. B., Wanderley, M. M., Dubnov, S., & Depalle, P. (1997). Instrumental gestural mapping strategies as expressivity determinants in computer music performance. In *Proceedings "Kansei - the technology* *of emotion" workshop* (pp. 68–73). Genova, Italy.

Schelleng, J. C. (1973). The bowed string and the player. *Journal of the Acoustical Society of America*, *53*(1), 26–41.

Schleske, M. (2004, May). *Zeitgeist und Geigenbau – Meilensteine aus Kunst und Wissenschaft. Eine kleine Zeitreise [Spirit of the time and violin-making - milestones in art and science. A short journey through time (translation by the author)].* Retrieved 31 December 2010, from `http://www.schleske.de/fileadmin/user_upload/doc/Zeitgeist.pdf`

Schleske, M. (2010). *Website of the luthier Martin Schleske.* Retrieved 16 October 2010, from `http://www.schleske.de/en/master-luthier.html`

Schoner, B. (2000a). Cluster-weighted sampling for synthesis and cross-synthesis of violin family instruments. In *Proceedings of the 2000 international computer music conference* (pp. 376–379). Berlin, Germany. Available from `http://alumni.media.mit.edu/~schoner/`

Schoner, B. (2000b). *Probabilistic characterization and synthesis of complex driven systems.* Doctoral dissertation, Massachusetts Institute of Technology, Cambridge, MA, USA.

Schoner, B. (2010). *Website of Bernd Schoner.* Retrieved 16 November 2010, from `http://alumni.media.mit.edu/~schoner/`

Schoonderwaldt, E., Rasamimanana, N., & Bevilacqua, F. (2006). Gesture controlled audio systems: Combining accelerometer and video camera: reconstruction of bow velocity profiles. In *Proceedings of the 2006 conference on new interfaces for musical expression* (pp. 200–203). Paris, France.

Schriewer, K. (1995). *Objektorientierte Windowsprogrammierung eines statistischen Analysetools zur Auswertung von Hörversuchen [Object-orientated Windows programming of a statistics tool for the analysis of listening tests (translation by the author)].* Diploma thesis, University of Applied Sciences Düsseldorf, Düsseldorf, Germany.

Schulze, H. H. (1989). *Computer Enzyklopädie: Lexikon und Fachwörterbuch für Datenverarbeitung und Telekommunikation [Computer encyclopaedia: Lexicon and specialised dictionary for data processing and telecommunication (translation by the author)].* Reinbek bei Hamburg, Germany: Rohwolt.

Serafin, S. (2004). *The sound of friction: Real-time models, playability and musical applications.* Doctoral dissertation, Stanford University, Stanford, CA, USA.

Serafin, S., & Young, D. (2003). Bowed string physical model validation through use of a bow controller and examination of bow strokes. In *Proceedings of the Stockholm music acoustics conference (SMAC 03)* (pp. 99–102). Stockholm, Sweden.

Sethares, W. A. (1998). *Tuning, timbre, spectrum, scale.* Berlin, Heidelberg, New York: Springer.

Settel, Z., & Lippe, C. (2003, May 22-24). Convolution brother's instrument design. In *Proceedings of the 2003 conference on new instruments for musical expression* (pp. 197–200). Montréal, Canada.

Shannon, C. E. (1948, July and October). A mathematical theory of communication. *The Bell System Technical Journal*, *27*(3 and 4), 379–423 and 623–66.

Smith, D. C., Irby, C., Kimball, R., & Verplank, B. (1982). Designing the star user interface. *Byte*, *7*(4), 242–282. Retrieved 8 December 2010, from http://www.guidebookgallery.org/articles/designingthestaruserinterface

Smith, J. O. (1991). *Historical View of Synthesizer Development.* Retrieved 15 October 2010, from https://ccrma.stanford.edu/~jos/kna/Historical_View_Synthesizer_Development.html

Sommerville, I. (2004). *Software engineering* (7th ed.). Amsterdam, Netherlands: Addison Wesley.

Stachowiak, H. (1973). *Allgemeine Modelltheorie [A general theory of models (translation by the author)].* Vienna, Austria: Springer.

Steim. (2010). *Centre for research and development of instruments and tools for performers in the electronic performance arts.* Retrieved 31 December 2010, from http://www.steim.org/steim/

Stowell, D., Plumbley, M. D., & Bryan-Kinns, N. (2008). Discourse analysis evaluation method for expressive musical interfaces. In *Proceedings of the 2008 conference on new interfaces for musical expression* (pp. 81–86). Genova, Italy.

Taylor, C. (1999). *Design Document for a Computer Game.* Retrieved 8 December 2010, from http://www.runawaystudios.com/articles/ctaylordesigntemplate.doc

Trogemann, G., & Viehoff, J. (2005). *Code@Art, eine elementare Einführung in die Programmierung als künstlerische Praktik [Code@art, a basic introduction to programming as an artistic practice (translation by the* *author)]*. Vienna, Austria: Springer.

Truax, B. (1987). Real-time granulation of sampled sound with the DMX-1000. In *Proceedings of the 1987 international computer music conference* (pp. 138–145). Urbana, IL, USA.

Trueman, D. (1999). *Reinventing the violin.* Doctoral dissertation, Princeton University, Princeton, NJ, USA.

Trueman, D., & DuBois, R. L. (2006). *PerColate* [Collection of MaxMSP third-party externals]. Retrieved 31 December 2010, from `http://music.columbia.edu/percolate/`

UIST. (2010). *ACM symposium on user interface software and technology.* Retrieved 31 December 2010, from `http://www.acm.org/uist/uist2010/`

Verfaille, V. (2003). *Effets audionumériques adaptatifs: théorie, mise en œuvre et usage en création musicale numérique [Adaptive digital audio effects: theory, implementation and usage in music applications (translation by the author)].* Doctoral dissertation, Université de la Méditerranée Aix-Marseille II, Aix en Marseille, France.

Verfaille, V., & Arfib, D. (2001). A-DAFX: Adaptive digital audio effects. In *Proceedings of the 4th international conference on digital audio effects 2001* (pp. 10–13). Limerick, Ireland.

Verfaille, V., & Depalle, P. (2004). Adaptive effects based on STFT, using a source-filter model. In *Proceedings of the 7th international conference on digital audio effects 2004* (pp. 296–301). Naples, Italy.

Wanderley, M. M. (1999). Non-obvious performer gestures in instrumental music. In *Gesture-based communication in human-computer interac-*

tion: International gesture workshop, GW'99, proceedings (Vol. 1739, pp. 37–48). Berlin, Heidelberg, Germany: Springer.

Wanderley, M. M. (2007). *Clarinet Gesture Analysis.* Retrieved 8 December 2010, from `http://www.music.mcgill.ca/musictech/clarinetWiki/doku.php`

Wanderley, M. M., & Battier, M. (Eds.). (2000). *Trends in gestural control of music.* Paris, France: IRCAM Centre Pompidou.

Wanderley, M. M., & Orio, N. (2002). Evaluation of input devices for musical expression, borrowing tools from HCI. *Computer Music Journal*, *26*(3), 62–76.

Wanderley, M. M., Vines, B., Middleton, N., McKay, C., & Hatch, W. (2005). The musical significance of clarinetists' ancillary gestures: An exploration of the field. *Journal of New Music Research*, *34*(1), 97–113.

Wessel, D., & Wright, M. (2002). Problems and prospects for intimate musical control of computers. *Computer Music Journal*, *26*(3), 11–22.

Winograd, T., & Flores, F. (1986). *Understanding computers and cognition.* Norwood, NJ, USA: Ablex Publishing Company.

Wolek, N. (2002). *Granular Toolkit* [MaxMSP third-party externals, abstractions and window functions]. Retrieved 31 December 2010, from `http://www.lowkeydigitalstudio.com/2007/03/granular-toolkit-v1-49/`

Woodhouse, J. (1992). Physical modeling of bowed strings. *Computer Music Journal*, *16*(4), 43–56.

Woodhouse, J. (1993). On the playability of violins. Part I: Reflection functions. *Acustica*, *78*(3), 125–136.

Yoo, L., & Fujinaga, I. (1998). Zeta violin techniques: Limitations and applications. *Journal SEAMUS*, *13*(2), 12–6.

Young, D. (2001). *New frontiers of expression through real-time dynamics measurement of violin bows.* Master's thesis, Massachusetts Institute of Technology, Cambridge, MA, USA.

Young, D. (2002). The hyperbow controller: Real-time dynamics measurement of violin performance. In *Proceedings of the 2002 conference on new instruments for musical expression* (pp. 1–6). Dublin, Ireland.

Young, D. (2007). *A methodology for investigation of bowed string performance through measurement of violin bowing technique.* Doctoral dissertation, Massachusetts Institute of Technology, Cambridge, MA, USA.

Young, D., Nunn, P., & Vassiliev, A. (2006). Composing for hyperbow: A collaboration between MIT and the Royal Academy of Music. In *Proceedings of the 2006 conference on new interfaces for musical expression* (pp. 396–401). Paris, France.

Young, D., & Serafin, S. (2003). Playability evaluation of a virtual bowed string instrument. In *Proceedings of the 2003 conference on new instruments for musical expression* (pp. 104–108). Montréal, Canada.

Zender, H. (1991). *Happy new ears.* Freiburg im Breisgau, Germany: Verlag Herder.

Zeta. (2010). *Website of the company Zeta Music.* Retrieved 13 January 2010, from `http://www.zetamusic.com/`

Zielinski, S. (2010). *Siegfried Zielinski.* Retrieved 16 October 2010, from `http://entropie.digital.udk-berlin.de/wiki/`

Siegfried_Zielinski