A SYSTEMIC APPROACH TO MUSIC PERFORMANCE LEARNING WITH MULTIMODAL TECHNOLOGY SUPPORT

Tillman Weyde\textsuperscript{1}, Kia Ng\textsuperscript{2}, Kerstin Neubarth\textsuperscript{1}, Oliver Larkin\textsuperscript{2}, Thijs Koerselman\textsuperscript{2}, Bee Ong\textsuperscript{2}

\textsuperscript{1}City University London, School of Informatics, Department of Computing, London, UK
\textsuperscript{2}ICSRiM - University of Leeds, School of Computing & School of Music, Leeds, UK

Abstract: The use of technology-enhanced learning in education is mostly seen as support for conveying and testing clearly defined knowledge and skills. In music performance education, the goal of the education is more personal and dynamically determined in the pedagogic and aesthetic interaction between the teacher and student. The approach taken in the i-Maestro project is therefore to support the teaching process by adding interaction forms and feedback loops that create new dynamics in the learning process. For music performance on string instruments, we use multimodal feedback with 3D motion tracking and sensor input to give students and teachers new forms of feedback and to make available aspects of performance that are hard to access by direct perception alone. We give examples how the integration these tools into the interaction and feedback loops of traditional music education can help to address problems like the gap between music theory and practice teaching.

1 Introduction

In this paper we present pedagogical considerations and technological developments within the context of the i-Maestro project which is co-supported by the EC IST under the 6\textsuperscript{th} framework programme. In i-Maestro we work on the development of multimedia technologies to support music practice and theory learning and teaching, focusing on string instruments. The pedagogical approach taken in i-Maestro is systemic, analysing tuition and designing multimodal interfaces in terms of feedback and control loops in music performance and music learning.

Based on studies in music psychology and education as well as input from a user group of music teachers and students, we have identified music-pedagogical challenges regarding motor programming and control, auditory anticipation and control, musical imagination, expressivity, and creativity, and effective and sustained practising. To address these challenges we have developed scenarios and defined educational and technological user requirements for multimedia technologies supporting music tuition. The technological framework of i-Maestro includes client tools, applications and services with gesture support and augmented instruments, audio processing and symbolic music processing for interactive performance and theory learning; a school server and portal for sharing resources and keeping profiles to support personalised education; and support for collaborative work enabling collaboration among teachers and students; and production and authoring tools for designing pedagogical resources.

Our systemic approach to string instrument tuition focuses on the interaction within and between systems, such as teacher and student or the musician and his instrument, and on control and feedback loops in musical activities (sensorimotor, visual, auditory, cognitive) which are essential in an educational domain like music. A systemic view has proven fruitful in general pedagogy, philosophy, and aesthetics (e.g. Bateson 1979, Maturana and Varela 1992). Schläbitz (2004) and Hametner (2006) have applied a systemic approach to the study of music education in general schools. To our knowledge, however, the systemic perspective has not yet been explored for specialised music practice training and technology-enhanced music learning.

In the following, we present the general pedagogical approach of the i-Maestro project (section 2). In
section 3 we give examples of i-Maestro tools for music learning with multimodal interfaces and discuss pedagogical implications. We conclude by reflecting on results and future work in section 4.

2 A systemic view on music instrument tuition

To introduce the systemic approach adopted in i-Maestro we analyse interaction between systems in music education (section 2.1) and control and feedback cycles in music performance (section 2.2). In section 2.3 we give reasons for the importance of feedback in music learning, before we describe in section 2.4 how technology can serve for feedback and control.

2.1 Real-time interaction in music education

Traditionally, human interaction in instrumental music lessons takes place between a teacher and one or more students. According to the findings West & Rostvall (2003) and information from the i-Maestro User Group, this interaction consists mainly of student performances and verbal comments by the teachers. According to Olsson (1997), language is used mainly to give positive or negative feedback on the student’s performance. West & Rostvall 2003 observed teacher-student interaction during instrument lessons in Swedish schools, with groups of students at the beginner level, and recorded a high number of only short utterances of instructional rather than analytical nature. Apart from human-human interaction, there is another important level of interaction between the musician and his or her musical instrument (e.g. Fricke 1993).

i-Maestro explores new possibilities that technology-supported multimodal interaction and feedback can provide for music education. To illustrate the potential of this idea, we can compare the routes of interaction in a traditional musical instrument lesson to the possibilities of interaction that can be created with technology (Figure 1).

![Figure 1: Diagram of real-time interaction in one-to-one instrumental performance teaching, traditional and technology-enhanced scenario.](image)

Introducing a computer with an audio and/or gestural input into the setting of one-to-one instrumental teaching introduces a multitude of additional interactions as shown on the right side of Figure 1. The space of pedagogical possibilities is enlarged by the added or modified feedback-loops between the student and a computer as well as between the student and her/his instrument and the teacher. Technology can be inserted on different levels of control and feedback, e.g. using real-time auditory feedback to motion data or supporting analysis after a performance. Most importantly, the additional interactions provided by technology are - unlike the traditional ones - optional and can be much more freely designed by the teacher or student, providing more options for tuition.

2.2 Control and feedback cycles in music performance

Music performance involves several control processes (Figure 2). Musical understanding and imagination, informed by music-theoretical knowledge (e.g. on musical structure or style), lead to an idea
of desired sound which can be realised by an appropriate, desired movement (see Harvey 1985 for a model on singing). The motor control system compares proprioceptive, haptic, and possibly visual feedback with the desired movement and adjusts the motor program; the auditory control system monitors the actual sound against the desired one and informs the motor control system.

Figure 2: Control and feedback levels in music performance.

For a student who learns to play an instrument, improving his or her performance skills involves enhancing auditory imagination and anticipation, enhancing perceptual skills like discrimination of pitch for auditory control of intonation, and acquiring and tuning motor patterns. Knowledge on music theory and history and appreciation of existing interpretations of a piece of music helps in creating an expressive performance (e.g. Sloboda 1985, Gabrielsson 1999), underpinning the necessity to connect music theory and practice.

2.3 Feedback in music learning

The role of feedback and more specifically knowledge of results (KR) has been emphasised in the context of motor learning, including musical activities (e.g. Pressing 1988, Thorpe 2002). Experiments in psychology have shown consistently that more KR leads to improvements in performance, although long-term learning and retention may benefit from a reduced frequency (Blackwell and Newell 1996, Lee and Carnahan 1990). Motion feedback can therefore be expected to contribute to learning when it provides the student with information about the performance quality that is additional to the regular response of the instrument. Verbal feedback can be ambiguous or misleading (Thorpe 2002); visual monitoring of playing movements is restricted by the student’s playing posture and gestures and visual field. Here technology can provide additional feedback by visualisation or sonification of movement or posture features traditionally not accessible to the student (see section 3.2). A particularly effective form of motor learning support is Bandwidth KR (Winstein 1991): KR is only given if the measured parameter is outside a certain range. In music instrument tuition, Bandwidth KR can be used to help avoiding unhealthy postures or other unwanted movements. This is particularly useful as intensive practising with strained posture or gestures can lead to painful and pathological conditions (Gabrielsson 1999, Williamson and Thompson 2006).

Feedback also affects motivation and consequently learning activities and effects. Motivation is based on an evaluation of the learning situation with respect to the learner's belief systems (e.g. self-
concepts and learning expectations) and is shaped by the learner's dispositions (i.e. personal characteristics and habitual emotional and motivational disposition). Positive feedback can create confident learning expectations and increase motivation. Negative feedback can increase learning motivation if it entails information how to improve or if the student expects increased effort to lead to improved learning results (Pekrun and Schiefele 1995).

However, controlling feedback, as opposed to informational feedback, has been shown to have detrimental effects on creativity if ego-involvement outweighs task-involvement (Silvia and Phillips 2004, Priest 2006): If a student experiences failure to achieve a certain learning outcome as a threat to his or her self-esteem, this results in lower levels of creativity. On the other hand, immediate feedback together with clear goals, reduction of fear of failure, and a balance between challenge and skills induces flow and fosters creativity (MacDonald et al. 2006, see also Sheridan and Byrne 2002).

2.4 Technology-enhanced control and feedback cycles

There are several benefits of inserting multimodal interfaces into existing feedback cycles. Firstly, multimodal interfaces can provide nonverbal information and thus help to sidestep possible problems of applying language to musical and gestural aspects of performance (Thorpe 2002). Secondly, technology can deliver feedback information in real-time as well as offline. Real-time feedback enables the student to relate the feedback directly to his/her posture and gestures, without relying on memory (Thorpe 2002). On the other hand, offline feedback allows teachers and students to analyse and discuss without having to attend to the performance at the same time. In addition, selected features of the performance can be highlighted, and temporally or spatially zoomed.

Alternatively, multimodal interfaces can be integrated into control cycles. Using for example sensors to conduct a sound playback lets the student develop a representation and musical interpretation of a piece of music even when not yet able to realise it on a musical instrument. Savage & Challis (2001) report positive cognitive-perceptual and motivational effects when students got the opportunity to create sounds without being constrained by technical difficulties of performance.

Multimodal interfaces offer a range of media and modalities. This poses potentials as well as critical issues as to the suitability of feedback channels (e.g. Levie 1985, Petridis et al., 2006). In i-Maestro we explore different modes of visualisation and sonification and their combination of movement parameters (for examples see section 3.2).

The use of technologically created feedback can have an impact on the role of the teacher and on teacher-student interaction. As mentioned earlier, the teacher is the main source of feedback and reinforcement. Instructional software that produces feedback creates another feedback loop, of which the teacher is a part. In Figure 3, the situation is sketched with the new feedback loop drawn with bold lines.

![Figure 3: Change of reinforcement loop.](image-url)

In the extreme case, the teacher is not interacting directly with the student, but by choosing and adjusting the exercise and feedback in the computer the teacher still has some control over the learning process. The teacher can in this case take a consulting role, rather than assessing the student. This is
probably an advantage, as the teacher-student relation is less burdened with the necessity of feedback on individual performances. The teacher is also in a favourable position to give advice, which may be understood to be helpful in achieving better results rather than just criticism. For this approach it is important that the student adopts the goal of achieving a positive computer feedback. Given that many youths spend hours playing computer games with the goal of collecting virtual feedback, it seems likely that a significant number of students would like to engage in interaction with and achieve positive feedback from the computer. In this context it is the teacher’s responsibility to guide the student and to make sure that the ultimate goal of playing music is not lost but still effectively pursued (Mills and Murray 2000).

3 Music learning with multimodal interfaces

The systemic approach in i-Maestro leads us to the development of hard and software that should integrate seamlessly with traditional music teaching and learning, while extending it with new opportunities. In the following, we give an overview on modalities addressed by i-Maestro interface elements (section 3.1) and present examples for the design of i-Maestro tools and their musical applications (section 3.2).

3.1 Multimodal technology support in i-Maestro

In the context of i-Maestro, music specific interaction technologies are developed by partners in the project for different modalities. They are described here with regards to the part they can play in music tuition.

Audio: Audio input is available on almost every modern computer, and has therefore the advantage of not requiring additional purchases by the user/owner, although built-in microphones are often not of sufficient quality to support musical audio processing. Audio input is used in a number of ways, from simple playback to analysing sound qualities or musical structures, as in score following (Schwarz et al., 2006).

Sensors: The sensors used by the i-Maestro project register acceleration in three axes and rotation in two axes to capture motion dynamics. The sensor is small and light and communicates with the computer wirelessly, which makes it easy to use the sensor with hand, arm, or whole body movements. The tracking of 3D positions over time is limited, but this technology is relatively inexpensive (it is comparable to that used in the remote control of the game console Wii by Nintendo) and can be used to detect motion characteristics and align motion patterns for gesture-following (Bevilacqua et al., 2007). One module using sensor signals is the classification of bow strokes based on the acceleration and deceleration of the bow and the beginning and ending of the note.

3D Motion Capture: For 3D motion capture, a multi-camera system by VICON is used. This system tracks the position of markers in space with high temporal and spatial resolution (approx. 5 ms, 1 mm). The markers are usually small balls, which can be attached to the player and the instrument (e.g. violin body and bow). The VICON system is used to provide 3D motion data that can be used to create appropriate visualisation, sonifications or other applications.

Symbolic Music Representation: Music notation is central to teaching western string instruments and the corresponding repertoire. Although there is commercially available software for music notation, these software systems are hard to integrate and lack interoperability. i-Maestro therefore uses and supports the standardisation of Symbolic Music Notation SMR as part of MPEG-4 by the International Standards Organisation (see Bellini et al., 2005). SMR brings music notation into the MPEG family of standards, allowing interoperability such as synchronisation, graphical and audio rendering, linking, and mark-up. These features and the open standard make MPEG-SMR particularly suitable for applications in music education.
3.2 Motor Learning

Motor learning is one of the main interests in i-Maestro, providing tools for teachers and students. i-Maestro support for motor learning aims to help teachers and students in working towards the learning objectives of acquiring, monitoring and improving technical performance skills for string instruments. It intends to enhance motion control by isolating, analysing, highlighting, and practicing individual aspects of posture and gestures in string playing, before gradually coordinating and automating such aspects. Visualisation and sonification provide the student with feedback to adapt the control of his/her movements.

Frequently encountered problems reported by teachers or addressed by string playing methods include bowing movements (such as keeping the bow parallel to the bridge, suitable distribution of the bow, bow velocity, and resulting sound quality). To support motor learning, a 3D Augmented Mirror is being developed that enables to monitor not only bowing movements, but also general posture and left-arm movements from different visual perspectives. Its feedback can continuously inform the student about posture or movements. This may range from a numeric/graphical display to verbal comments, depending on the pedagogic requirements.

For the 3D data, geometric analysis and processing are usually needed, including calculations of biomechanical models and other information based on the players body features, relative positions of body and instrument and for bowed strings the position of bow and instrument and their velocities and accelerations. These can either be used for visualisation or sonification to provide direct feedback. To gain more meaningful feedback, it is necessary to extract features that describe some relevant property of the data, such as the angle between the strings and the bow or the distance travelled vs. time for each bow stroke etc., together with meaningful visualisation to provide further understanding to the player. Further processing includes the analysis of movement patterns with interaction between the instrument and the bow. The main point of the 3D interface in this context is to provide views and allow analyses that are normally not directly available to the students and teachers, such as the exact angle between the bow and the strings or movement the trajectories.

Figure 4: Bowing shape and trajectory visualisation.

The feedback provided is illustrated in Figure 4 below with a screen snapshot of the i-Maestro 3D Augmented Mirror prototype which integrates 3D motion, audio and video within a 3D environment. Figure 4 also illustrates relative angles and position of a selected marker (in 3D) with 2D graphs which can be configure to visualise other parameters depending on the context of the usage or exercise. Short video clips and further details on the system are available online on the project website http://www.i-maestro.org.

The 3D Augmented Mirror provides new forms of interaction that are not available with a regular
mirror in terms of spatial perspectives, and allows the measurements of angles, e.g. between the bow and the strings or the trajectory of the bow movement, which are otherwise hard to determine for the player. The teacher or the student can adjust the system to give an auditory feedback when the angle departs from the parallel to the bridge by a certain amount, providing Bandwidth KR.

### 3.3 Auditory Sensitivity and Control

Auditory self-monitoring of one’s performance involves auditory anticipation of what a musical pattern should sound like and auditory control of the playing which compares the anticipated with the resulting sound. The control relies on perceptual discrimination and recognition skills. In addition, it is crucial for self-monitoring to know what to listen to (e.g. Sloboda 1985).

To enhance monitoring skills, interactive feedback can be provided by performance tracking. Using the above mentioned score-following module, the student’s performance can be mapped automatically onto the score being played: The score follower analyses the sound data of the performance in real-time and identifies the corresponding notes in the score. The notes currently played are highlighted in the score display by colour or a cursor. This feature is very valuable for teaching purposes, as it gives musically relevant information and allows musical localisation and contextualisation of pedagogical interaction.

A pedagogical application of score following is providing feedback on the tempo of a performance. Some teachers in the i-Maestro User Group have named tempo stability as a frequent problem with students. Having tempo information allows creating feedback that can draw the user’s attention to tempo deviations. The feedback can be designed in different ways in a lesson, depending on the individual learning and teaching objective. If the student’s focus should be on the tempo as such, a direct display of the numerical value may be sufficient. A probably more usable display would show the value graphically on a scale, or using different colours (e.g. blue: too slow, green: good, red: too fast), possibly a combination of these. It would also be possible to use auditory feedback, but that might be perceived as interfering with the music. It remains yet to be tested, which display leads to best results. An advantage of this approach is clearly that it is less rigid than using a metronome, which impedes even slight tempo deviations, which are often musically meaningful and intended.

On a subsequent level of training, tempo curves instead of instantaneous tempo values can be useful feedback. Sloboda (1985) compares novice and expert expressive performance and observes that less experienced performers sometimes lack appropriate timing of tempo changes, like ritardando or accelerando, with respect to the musical structure. Feedback on the position and extension of tempo curves thus can help students to analyse their expressive performance. In an advanced feedback mode, the feedback could be related to the tempo changes imagined during performance planning, i.e. students design their individual target tempo curve against which the tempo measured during the subsequent realisation is matched.

A common problem that leads to lack of auditory control is a cognitive and control overload. Students are often distracted by the many tasks they have to master simultaneously, and find it difficult to focus on a specific problem. i-Maestro gesture-tracking and audio playback can reduce the cognitive load for the student by allowing to control specific parameters of a recorded music performance by conduction or bowing gestures, realising a teaching strategy we call cognitive relief. The gesture tracking tool records movement data from a sensor integrated into a conductor’s baton or a violin bow and uses these data to regulate a recorded performance according to user-set options. With this tool, specific features of the performances can be isolated and studied. Also, such technology-supported control cycles enable students to explore music performance without being limited to their level of achievement in playing a difficult instrument like a string instrument.
3.4 Musical Understanding and Theoretical Knowledge

Musical performance relies on the performer’s mental representation of the piece to be performed. This representation includes the musical structure, musical expression, and the performer’s interpretative intentions. In addition, musicians often form a motor representation of the piece (Sloboda 1985, Gabrielsson 1999). Building a representation is based on an analysis of the piece, informed by theoretical knowledge. Expressive characteristics are related to the structure of the music and the historical and stylistic context of the composition, and teachers also discuss the composer’s background with their students.

Traditionally, the planning of a performance involves hand-written annotations in the score on paper. According to Oerter and Bruhn 2005, graphical mark-up of a score seems to be more effective than verbal accounts in analysing, understanding, and representing pieces of music. Using the SMR format and software, we can annotate scores with any kind of media (text, graphics, audio, video) and generate audible renderings of the music, which allows us to combine theoretical work on a score with performance. Based on a theoretical analysis, a student can create a performance plan. And even without being able to play the piece, the student can create a playback according to the plan. Although much of this can be done with commercially available music notation software, these are closed systems that lack the annotation capabilities and the integration of SMR and i-Maestro tools.

Vice versa, a student can record tempo, dynamics, and articulation using audio and sensors, which will be added as annotations to the score, generating a symbolic annotation plan from the performance. The student will then be able to edit score annotations and generate a new audio rendering. In this way, students may use their interactively created performance plan as the basis for their own practicing, e.g. using the earlier mentioned tempo bandwidth feedback or the bow stroke classification module. The possibilities of audio and sensor analysis and audio score rendering allow therefore changing freely between symbolic and practical work as sketched in Figure 5.

![Figure 5: Round-Trip Exercise.](image)

Using MPEG-SMR and score annotations with playback functions thus supports linking music theory and practice, a challenge which was noted as early as by Rimskij-Korsakov (1885) and was not solved in the late 20th century neither on an institutional level nor in instruction theory (Richter 1997). Bridging the gap between music theory and practice is still reported to be a persisting task by members of the i-Maestro User Group.

4 Conclusions

A systemic approach to music performance learning emphasises real-time and short-term interaction in music instrument playing and teaching. Such interaction involves feedback from different systems (like teacher feedback or student self-monitoring) and on multiple levels (cognitive, visual, auditory, proprioceptive), which is used to adapt auditory and motor control in playing the instrument. The European i-Maestro project explores multimodal interfaces to technologically extend feedback and control cycles in music performance.
3D motion tracking and the use of movement sensors with visualisation and sonification of movement data, providing either real time or offline feedback, can raise the student’s awareness of motor control and improve self-perception. Additionally, immediate informational feedback is expected to increase the student’s motivation. We also hope that the use of motion feedback contributes to students’ well-being, as it may help to avoid problematic posture and gesture and related health risks.

With different forms of sound production and manipulation, using audio processing techniques and possibly sensor technology, the student can be offered to control sound, sidestepping difficulties of instrument playing technique, and focus on auditory imagination and monitoring. Sound processing can be applied to synthesised sounds or sound recordings, such as the student’s own recorded performance. MPEG Symbolic Music Representation technologies, with score annotation and sound generation, provide a means to link music performance and symbolic knowledge. In combination with motion tracking and audio processing, score annotation can be used for feedback (automatic annotation from motion data) as well as for control (audio data generation from annotated score).

The possibility to create multi- and inter-modal feedback is one of the particularly interesting features of current technologies. Actions and the related perceptions in different senses are linked and the involvement of multiple perceptual modalities could therefore support holistic learning and experiences. However, the actual effect of using these tools depends on the particular design of the interaction and pedagogical setting. The main research question is therefore how to design interaction and integration it into pedagogical context, which will be validated in teaching practice.

Acknowledgement

The i-Maestro project is partially supported by the European Community under the Information Society Technologies (IST) priority of the 6th Framework Programme for R&D (IST-026883, www.i-maestro.org). We thank all i-Maestro project partners and participants for their interest, contributions and collaboration.

References


