INTRODUCING THE INTERPRETATION SWITCHER INTERFACE TO MUSIC EDUCATION

Verena Konz and Meinard Müller
Saarland University and MPI Informatik, Campus E1 4, 66123 Saarbrücken, Germany

Keywords: Music information retrieval, Music education, Music synchronization, Alignment, User interface.

Abstract: In the field of music information retrieval (MIR), great efforts have been directed towards the development of technologies and interfaces that allow users to access and explore music on an unprecedented scale. On the other hand, musicians and music teachers are often still skeptical about the benefits of computer-based methods in music education. In this paper, we report on an experiment conducted at the University of Music Saarbrücken with the goal to introduce a novel MIR user interface, referred to as Interpretation Switcher, to music education and to get feedback from music experts. To this end, we asked nine music students to analyze different performances of the same piece of music according to a well designed questionnaire, using the novel switching functionality of our interface. Doing so, we not only tested and evaluated our interface in a setting of practical relevance, but also indicated the potential of MIR methods in music education.

1 INTRODUCTION

Computers have become an indispensable tool for storing, processing, and generating music. Even though computer-based methods and interfaces are ubiquitously used for music synthesis, there is still a reluctance in using computers for music analysis and music education. Research in computer-assisted music education already started at the end of the 1960s mainly in the USA and the United Kingdom (Brown, 2007; Kiraly, 2003; Smith, 2009; Stevens, 1991; Guetl and Parncutt, 2008). For example, computers have served as a tool for creative music-making. Furthermore, the method of Computer-Assisted-Instruction (CAI), where students are taught a particular skill by a computer, has been applied in areas like music theory or aural training. Various studies have been conducted to investigate the effect of CAI-based methods within music education (Smith, 2009; Kiraly, 2003), and the usefulness of such methods seems to be a controversial issue.

In the field of music information retrieval (MIR), the development of technologies and interfaces for music exploration and analysis has been an active research area (Damm et al., 2008; Dixon and Widmer, 2005; Goto, 2003; Müller, 2007; Sonic Visualiser, 2009). However, these technologies and interfaces are often evaluated in the own lab environment, where people are familiar with computers. Though, for building up MIR systems of practical relevance one needs broader feedback, in particular from music experts. Hence, for a user-centered analysis, it is necessary to conduct "real" user studies that ensure a natural setting (Lesaffre et al., 2008).

There are many applications in the context of music education that may benefit from the above mentioned MIR-based technologies and interfaces. However, musicians and music teachers are often still reluctant in using novel computer-assisted methods and novel MIR interfaces in their lessons. Furthermore, many of the available MIR interfaces are still too complicated lacking the necessary user-friendliness and robustness to be operable by non-experts. Under such circumstances, it remains a challenge to raise the interest of music educators for using, testing, and participating in the development of novel MIR interfaces and for discussing possible application scenarios.

In this paper, we report on an investigation with the objective of introducing a novel MIR interface to music education. In collaboration with the University of Music Saarbrücken we conducted an experiment consisting of several steps. First, nine piano students were recorded playing the same piece of music, the first movement of Beethoven's Pathétique Sonata Op. 13, on the same piano and under the same recording conditions. In the next step, the nine audio recordings were temporally aligned and integrated in a user interface referred to as Interpretation Switcher (Fre-

Konz V. and Müller M. (2010).
INTRODUCING THE INTERPRETATION SWITCHER INTERFACE TO MUSIC EDUCATION.
In *Proceedings of the 2nd International Conference on Computer Supported Education*, pages 135-140
DOI: 10.5220/0002765201350140

merey et al., 2007; Müller, 2007), which allows for synchronous playback of the different performances. Upon using this interface, the music students were then asked to analyze the anonymised performances according to a well-designed questionnaire.

There are a number of achievements of our experiment. Firstly, we tested and evaluated our interface in a setting of practical relevance, thus indicating the potential of MIR methods in music education. Secondly, we generated royalty free music recordings without any copyright restrictions, which can be used freely for research purposes. Thirdly, using a Yamaha Disklavier for our experiments, we also obtained MIDI data (which was actually not used in the investigation described in this paper) along with audio recordings. Such MIDI-audio pairs can be used as ground truth material for various MIR tasks (Müller, 2007). Finally, we generated many different interpretations of the same piece, which yields valuable data for tasks such as automated performance analysis (Widmer et al., 2003).

The paper is organized as follows. In Sect. 2, we present in detail the setup of the conducted experiment describing the piece of music, the recording conditions, the interface, and the questionnaire. Then, in Sect. 3, we discuss the results of our experiment responding to the performance evaluation and to the interface evaluation. Finally, in Sect. 4, we sketch further application scenarios and indicate future work.

2 EXPERIMENTAL SETUP

2.1 Piece of Music

For our experiment, we chose the first movement of Beethoven's Pathétique Sonata Op. 13. This piece of music appeared to be a good choice for various reasons. Firstly, the Pathétique Sonata is a musicologically outstanding work, for which numerous detailed descriptions and scientific literature exists. Secondly, being a very popular and famous work, the Pathétique belongs to the standard repertoire of many pianists. Hence, there are numerous audio recordings for this piece. The third and most important reason for choosing the Pathétique is that it is very rich in contrast concerning tempo as well as dynamics.

To make the latter point clear, we describe the Pathétique's exposition in more detail. Beginning with the slow introductory theme marked *Grave* (measures (abbreviated mm.) 1-10, see Fig. 1 (a)), the work starts very dramatically. The introduction is characterized by its contrasts in dynamics—fortissimo passages are followed by subito piano and

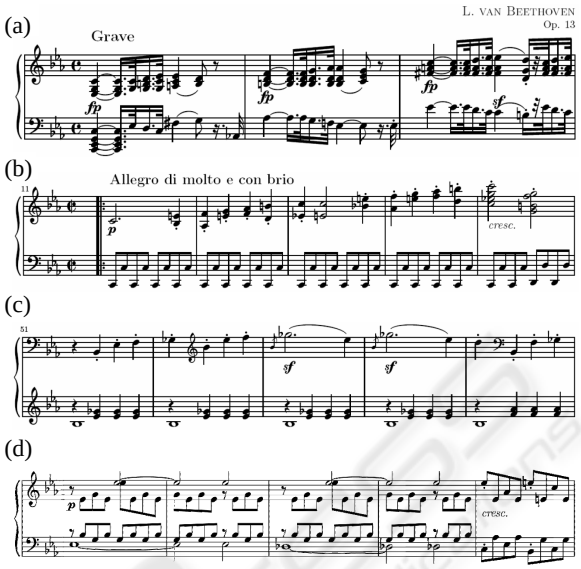

Figure 1: First movement of Beethoven's Pathétique Sonata Op. 13 (score obtained from (Mutopia Project, 2009)). **(a)** Beginning of the introduction (Section A, mm. 1 ff.) **(b)** First theme (mm. 11 ff.) **(c)** Second theme (mm. 51 ff.) **(d)** Section B (mm. 89 ff.)

vice versa. This contrast in dynamics is underlined by contrasts in rhythmics, articulation, and atmosphere. Ending with the chromatic run, the introduction leads in the first theme (mm. 11-27, see Fig. 1 (b)) of the sonata which is characterized by the tremolo in octavos in the left hand giving it a dramatic touch. In contrast to the dramatic first theme, the second theme (mm. 51-88, see Fig. 1 (c)) sounds more playful. It is based on the call and response principle and is characterized by a play with articulation.

As described above, one can find many contrasting elements in the exposition of the first movement of the Pathétique: Contrasts in the shaping of the two themes, contrasts in dynamics, and contrasts in atmospheres. In addition, there is an abrupt change in tempo at the beginning of the first theme (mm. 11), where the introductory *Grave* leads in the actual exposition marked *Allegro di molto e con brio*. Because of its musical richness, the Pathétique offers the pianist a wide range of possibilities for shaping the piece with respect to dynamics, tempo, and agogics. Therefore, this piece is very well suited in view of the musical evaluation within our experiment.

2.2 Performance and Recording Setup

The recordings of our experiment were performed at the University of Music Saarbrücken. Nine students of different study paths from the piano class of Prof.

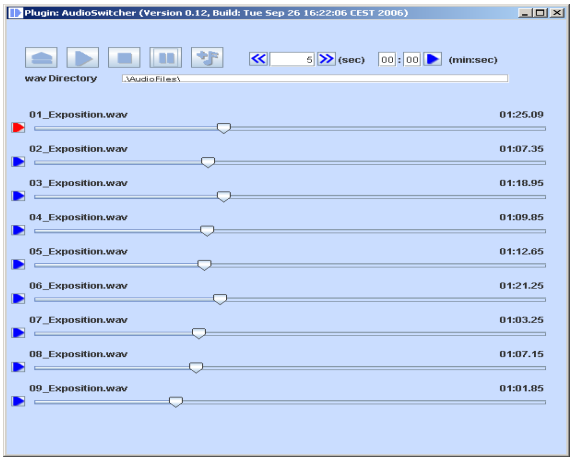

Figure 2: Instance of the Interpretation Switcher plugin of the SyncPlayer for synchronous playback of different audio recordings of the same piece of music. In this example, nine different recordings of the exposition of Beethoven's Pathétique Sonata are opened.

Thomas Duis were asked to play the first movement of the Pathétique. Being in several training states, they were on different performance levels. All students played on the same instrument under the same recording conditions on two different days (Friday, 06.02.2009 and Monday, 09.02.2009). In the recording sessions, only the performer, the technical staff, and the scientific investigators were present in the room—the other performers were not allowed to listen to their fellow students. Using two microphones, we did not achieve the quality of a recording studio. However, we obtained audio recordings of a sufficient quality in view of our experiments. Furthermore, using a Yamaha Disklavier, we also generated MIDI data along with the audio recordings. Actually, the MIDI files were not used in our experiments, but as mentioned before they are useful for later projects.

2.3 MIR User Interface

The *SyncPlayer* system is an advanced audio player for multimodal presentation, browsing, and retrieval of music data (Fremerey et al., 2007). One of the available plugins for the SyncPlayer, referred to as *Interpretation Switcher*, is the MIR interface used in our experiment, see Fig. 2. It allows the user to select several recordings of the same piece, which have previously been synchronized (Müller, 2007). Each of the selected recordings is represented by a slider bar indicating the current playback position with respect to the recording's particular time scale. The audio recording that is currently used for playback, in the following referred to as reference recording, is represented by a red marker. The slider of the reference recording moves at constant speed while the sliders of the other recordings move according to the relative tempo variations with respect to the reference. The reference recording may be changed at any time simply by clicking on the respective marker located on the left of each slider. The playback of the new reference recording then starts at the time position that musically corresponds to the last playback position of the former reference. One can also jump to any position within any of the recordings by directly selecting a position of the respective slider, which automatically triggers a switch of the reference to the respective recording. A similar functionality is provided by the Sonic Visualiser, a system for viewing and analyzing the contents of music audio files (Sonic Visualiser, 2009). Here, the MATCH plugin allows for temporally aligning two recordings and then for switching from one to the other (Dixon and Widmer, 2005).

Note that the SyncPlayer and the Sonic Visualiser provide many more functionalities comprising plugins for inter- and intra-document browsing and retrieval as well as data visualization and analysis. In our experiment, we restrict ourself basically to the switching functionality with the motivation to keep the interface as simple and intuitive as possible to avoid any rejections from the users. As Fig. 2 shows, the Interpretation Switcher looks like a standard audio player with the only difference that more than one slider control bar is available. After a short explanation of the main switching functionality, none of the students reported on difficulties in using our interface.

2.4 Survey and Questionnaire

Subsequent to the last recording session, the nine different performances were aligned and integrated in our Interpretation Switcher. Then, we conducted our survey in the evening of the second recording day (Monday, 09.02.2009). Eight music students participated in the survey, seven of whom were also among the nine performers. The different interpretations were anonymised within the interface and the participants listened to the recordings for the first time.

Each participant was provided with a computer running the Interpretation Switcher and with earphones. After a short introduction of the interface's switching functionality, the participants received a questionnaire having one hour for answering the questions. This questionnaire consisted of two main parts. In the first part, the students had to listen, to compare, and to rate the nine different interpretations with respect to various performance aspects. Here, the questions were designed in such a way that the

students naturally started to use the switching functionality of the interface, thus getting familiar with the Interpretation Switcher in a concrete application of musical relevance. In the second part, they were then asked to give feedback on the usefulness and operability of the interface itself.

The questions of the first part of the questionnaire referred to different sections of the first movement of the Pathétique. As a kind of warming up, we started with a short section (Section A), which only consisted of the first three measures, see Fig. 1 (a). This section was cut out from the nine aligned performances and presented to the students by the Interpretation Switcher interface. Even being rather short, Section A already offers the pianists a wide range of interpretation so that the comparison of the different performances constitutes a musically interesting task. In the first question (A1), the participants had to rate the nine different interpretations of Section A with respect to the three musical aspects dynamics, articulation, and agogics. Here, the rating scale ranged between 1 and 10, where 1 means poor and 10 excellent. In addition, they had to rate their total impression of this section's performances using the same scale. Afterwards, in question A2, they had to identify their own interpretation (if applicable) only by means of Section A. Then, the performances of Section A were closed and a different section (Section B) was presented by the Interpretation Switcher to them. Here, Section B consisted of the technically more involved mm. 89-100, see Fig. 1 (d). The students then had to answer corresponding questions (B1, B2).

At the beginning of the questionnaire, the students were confronted with different performances of relatively short sections. Here, only switching between the performances was required to properly answer the questions—jumping back and forth within a performance was not necessary. In this way, the students became familiar with the basic switching functionality of the interface. In the next stage, they were presented with the nine performances of the entire exposition. They now had to rate their total impression of the first theme (mm. 11 ff., see Fig. 1 (b)), of the second theme (mm. 51 ff., see Fig. 1 (c)), and of the entire exposition (questions E1, E2 and E5). Here the new challenge concerning the use of the Interpretation Switcher was not only to switch between the different performances but also to find the corresponding entry points of the two themes within the recordings. Another task (E3), was to order the nine different interpretations with respect to the tempo (beginning with the slowest, ending with the fastest) in the second theme. With this task the students had to constantly switch between and jump within the performances, being forced to use the functionality of the interface extensively. In question E4, again, they had to identify their own performance (if applicable) now having the entire exposition at their disposal.

After finishing the questions on music aspects, in the second part of the questionnaire the participants were asked to evaluate the Interpretation Switcher interface. Here, the idea was to let the participants first use the interface in an application scenario to gather practical experience without knowing about the final interface evaluation. In the first question (S1), they should rate the user-friendliness and the degree of usability of the Interpretation Switcher on the above described scale from 1 to 10. We then wanted to know if there were any problems while using the interface (S2). Furthermore, the students were asked to comment on possible improvements and to propose additional functionalities they would have liked when working on the first part of the questionnaire (S3). In a last question (S4), they should sketch possible application scenarios where they could imagine to use MIR user interfaces such as the Interpretation Switcher.

3 EVALUATION

3.1 Performance Evaluation

In the first part of the questionnaire, the participants had to analyze and compare the different performances against each other. Table 1 presents the results of question A1, where they had to rate the nine different performances of Section A with regard to dynamics, articulation, agogics, and in total. The first row of Table 1 shows the number of the respective performance; the values of each column correspond to the respective performance. The second row shows the ratings with regard to dynamics averaged over the eight participants. For example, the first performance was rated with a score of $\mu = 6.63$ on average. The third row shows the standard deviation, which is $\sigma = 1.19$ for the first performance. The following rows of Table 1 are to be read in the same fashion. For example, the participants rated the sixth performance on average with $\mu = 6.88$ ($\sigma = 1.36$) with respect to articulation, whereas the overall impression of this performance amounts to $\mu = 6.38$ (σ=1.51). As we can see, the eighth performance was ranked highest with respect to dynamics ($\mu = 6.75$), whereas the second one with respect to articulation ($\mu = 7.00$). The overall rankings for Section A are relatively close together, which may show that the section was too short for giving a well-founded evaluation or that it was played similarly by all students.

Table 1: Evaluation results of question A1 (Fig. 1 (a)). The average ratings μ along with the standard deviations σ are shown for the nine performances of Section A with regard to various musical aspects.

		1	2	3	4	5	6	7	8	9
Dynamics	μ	6.63	6.25	6.38	6.13	5.75	6.38	5.38	6.75	6.63
	σ	1.19	1.49	1.69	1.36	1.67	1.69	1.77	1.49	1.60
Articulation	μ	6.13	7.00	6.25	6.38	6.13	6.88	5.38	5.88	6.00
	σ	1.81	1.69	1.83	1.30	1.64	1.36	2.20	1.36	1.41
Agogics	μ	6.13	6.50	5.13	6.38	5.63	6.50	4.75	5.75	6.13
	σ	1.73	1.69	2.53	1.92	1.41	1.41	2.12	1.28	0.99
Overall	μ	6.25	6.25	6.00	6.25	5.88	6.38	5.25	6.13	6.00
	σ	1.58	1.67	2.14	1.28	1.25	1.51	1.58	1.55	1.20

Table 2: Evaluation results of question B1 (Fig. 1 (d)).

		1	2	3	4	5	6	7	8	9
Dynamics	μ	5.38	6.38	7.13	6.38	5.63	7.13	5.88	4.75	4.13
	σ	1.41	1.69	1.55	1.69	1.41	1.13	1.36	2.12	1.36
Articulation	μ	5.38	6.00	6.25	6.38	4.88	6.25	5.13	4.63	4.75
	σ	1.85	1.20	1.04	1.60	1.89	1.58	1.89	1.69	2.19
Agogics	μ	5.13	6.00	6.88	6.75	5.13	6.75	5.38	4.88	4.00
	σ	2.42	0.76	0.99	1.39	1.81	1.28	1.51	1.64	1.51
Overall	μ	5.25	5.88	6.75	6.63	5.00	6.63	5.63	5.00	4.25
	σ	1.75	0.99	1.49	1.41	1.51	1.41	1.69	1.51	1.39

Analogously, Table 2 shows the results of question B1. Here, the best performance concerning the overall impression is the third one ($\mu = 6.75$), whereas the worst performance is the ninth one ($\mu = 4.25$). Actually, the ninth performance was ranked worst with respect to all musical aspects. The reason for the poor rating is that the performing student struggled significantly with the technically more involved Section B, thus neglecting the musical shaping. This may also explain, why the given scores between the performances differ to a much larger degree for Section B than for Section A. Finally, Table 3 presents the results of questions E1, E2 and E5, where only the overall impression had to be rated. Here, the first, third, and second performances were rated best with regard to the first theme ($\mu = 6.88$), the second theme ($\mu = 6.75$), and the entire exposition ($\mu = 7.00$), respectively. Again, the ninth performance was rated worst with regard to all three categories. Interestingly, there does not exist a clear winner performance concerning all different musical aspects and themes.

3.2 Interface Evaluation

In the second part of the questionnaire, the students were asked about the operability and usefulness of the Interpretation Switcher. As mentioned before, none of them had serious problems in using the interface, which is also reflected by a high average rating of $\mu = 7.63$ given for the user-friendliness of the interface. Only one of the participants gave a low score of 4 explaining a relatively large standard deviation of $\sigma = 2.07$. As it turned out, the reason for this was that the student was pressured for time and not really in the mood of participating in our experiment. Actually, this student also admitted that she has had no time for properly practicing the piece, resulting in performance number nine with the lowest score, see Table 3. Most of the other participants emphasized that they found the handling and functioning of the Interpretation Switcher very intuitive, even music students who have had only little experience with computers. Furthermore, most students found the Interpretation Switcher very useful for tasks such as performance analysis, music comparison, and other analysis tasks. Here, the average rating amounted to $\mu = 7.13$ with standard deviation $\sigma = 1.89$.

Table 3: Evaluation results of question E1 (first theme, Fig. 1 (b)), question E2 (second theme, Fig. 1 (c)), and question E5 (entire exposition).

		1	2	3	4	5	6	7	8	9
1. Theme	μ	6.88	6.25	5.88	5.63	4.88	5.25	5.50	5.25	4.75
	σ	1.13	1.39	1.25	2.13	2.30	2.19	2.45	2.12	1.83
2. Theme	μ	6.25	6.38	6.75	5.63	5.00	5.25	6.38	5.13	4.50
	σ	1.83	1.06	1.39	1.30	1.60	1.04	1.85	1.96	1.60
Exposition	μ	6.50	7.00	6.88	5.63	4.75	5.13	5.50	5.63	4.38
	σ	1.41	1.51	1.46	1.85	1.58	1.81	2.00	1.77	1.60

After the general rating, the students were also asked to freely comment on problems, possible improvements, additional functionalities, and possible application scenarios (S2, S3, S4). At this point they all confirmed that they have had no problems while using the Interpretation Switcher interface. However, two students noted that the interface could have reacted faster while switching between the respective performances. One student would have appreciated to have an additional functionality for displaying the musical score during playback. Also user-defined auxiliary markers that can be freely fixed, adjusted, and removed along the various slider control bars should be introduced for additional orientation and navigation purposes. All but one of them affirmed that they could imagine to use the Interpretation Switcher within their studies or even for private use. In particular, they said that the interface may be useful in the context of special seminars, where the comparison of different performances play an important role. One student was enthusiastic about the features offered by the Interpretation Switcher. He usually records his piano lessons in order to listen to and to study his own playing afterwards. Here, he would significantly benefit from novel switching and navigation functionalities for comparing and analyzing the recorded audio material. Also, the Interpretation Switcher could be

very useful for compactly documenting the learning progress over a longer period in time. For example, it could synchronously present the various performances of a specific musical section recorded in different piano lessons over the semester.

4 CONCLUSIONS

In this paper, we presented a first experiment conducted at the University of Music Saarbrücken with the main objective of introducing MIR user interfaces with novel switching and navigation functionalities to music teachers and students. Even though this group tends to be skeptical about using computer-based methods in music education, most participants affirmed the usefulness of our interface for comparing and analyzing performances or simply for music listening and enjoyment. Testing and evaluating our interface within a concrete application of practical relevance, we not only made a new group of prospective users acquainted with MIR methods but also obtained valuable feedback from music experts.

The presented experiment only constitutes the beginning of a planned collaboration with music educators and students, who are usually not aware of the developments in music information retrieval. For the future, we plan to conduct similar experiments on a larger scale. One further idea is to participate regularly in the lessons of piano students to record their playing. We then plan to process (segment, classify, synchronize) the audio material automatically and to suitably integrate it in our Interpretation Switcher to document and analyze the students' learning process.

Finally, we plan to develop and combine various additional functionalities. For example, as mentioned by one of the participants, an additional sheet music interface for presenting the musical score while playing back associated audio material would be helpful. Actually, such functionalities have been presented in (Damm et al., 2008). Furthermore, we will integrate additional functionalities for inter- and intra-document music browsing including the possibility of setting user-defined auxiliary markers as well as precomputed markers that reflect the musical form of the piece (Fremerey et al., 2007; Goto, 2003). In introducing novel functionalities, one main challenge will be to keep the operability of the interface as intuitive as possible to avoid rejections from the users' side.

ACKNOWLEDGEMENTS

The research was funded by the Cluster of Excellence on *Multimodal Computing and Interaction* at Saarland University. We thank Wolfgang Bogler and the piano class of Prof. Thomas Duis at the University of Music Saarbrücken for their support.

REFERENCES

Brown, A. (2007). *Computers in Music Education*. Routledge.

Damm, D., Fremerey, C., Kurth, F., Müller, M., and Clausen, M. (2008). Multimodal presentation and browsing of music. In *Proc. ICMI, Chania, Crete, Greece*, pages 205–208.

Dixon, S. and Widmer, G. (2005). Match: A music alignment tool chest. In *Proc. ISMIR, London, United Kingdom*, pages 492–497.

Fremerey, C., Kurth, F., Müller, M., and Clausen, M. (2007). A demonstration of the SyncPlayer system. In *Proc. ISMIR, Vienna, Austria*, pages 131–132.

Goto, M. (2003). SmartMusicKIOSK: Music listening station with chorus-search function. In *Proc. ACM Symposium on User Interface Software and Technology*, pages 31–40.

Guetl, C. and Parncutt, R. (2008). An interactive tool for training and testing musical auditory skills. In *Proc. ED-MEDIA, Vienna, Austria*, pages 5229–5237.

Kiraly, Z. (2003). Solfeggio 1: a vertical ear training instruction assisted by the computer. *International Journal of Music Education*, 40(1):41–58.

Lesaffre, M., Leman, M., De Baets, B., De Meyer, H., De Voogdt, L., and Martens, J.-P. (2008). How potential users of music search and retrieval systems describe the semantic quality of music. *Journal of the American Society For Information Science and Technology*, 59(5):1–13.

Müller, M. (2007). *Information Retrieval for Music and Motion*. Springer.

Mutopia Project. Music free to download, print out, perform and distribute. `http://www.mutopiaproject.org`, Retrieved 12.05.2009.

Smith, K. H. (2009). The effect of computer-assisted instruction and field independence on the development of rhythm sight-reading skills of middle school instrumental students. *International Journal of Music Education*, 27(1):59–68.

Sonic Visualiser (2009). `http://www.sonicvisualiser.org/`, Retrieved 12.05.2009.

Stevens, R. S. (1991). The best of both worlds: An eclectic approach to the use of computer technology in music education. *International Journal of Music Education*, 17(1):24–36.

Widmer, G., Dixon, S., Goebl, W., Pampalk, E., and Tobudic, A. (2003). In search of the Horowitz factor. *AI Magazine*, 24(3):111–130.