Are computer-controlled pianos a reliable tool in music performance research? Recording and reproduction precision of a Yamaha Disklavier grand piano

Werner Goebl (1,2)
Roberto Bresin (1)
(1) Department of Speech, Music, and Hearing (TMH)
Royal Institute of Technology (KTH), Stockholm
(2) Austrian Research Institute for Artificial Intelligence (OeFAI), Vienna
werner.goebl@oefai.at; roberto@speech.kth.se

Abstract

In this study, a Yamaha Disklavier is tested on its measuring and reproducing capabilities, with the goal to examine its use in performance research. An experimental setup with accelerometers and a calibrated microphone is used to capture key and hammer movements, as well as the sound signal. Five selected keys are played by pianists with two types of touch (*'staccato – legato'*). Timing and dynamic differences between the original performance, the corresponding MIDI file recorded by a Disklavier, and its reproduction are analysed. Information of the MIDI file was more precise than the reproduction by the Disklavier. Timing errors are larger for soft tones and hammer velocities higher than 3.5 m/s could not be reproduced by the solenoids.

1 Introduction

Current research in expressive music performance mainly deals with piano interpretation since pianists are able to control only a few expressive parameters (note on/off, intensity, pedal information) on their instruments. Computer-controlled grand pianos are a practical source for the detection of these parameters and at the same time provide a natural setting for the players in a recording situation. Two systems are most commonly used in performance research: the Bösendorfer SE (Palmer & Brown, 1991; Repp, 1993; Palmer, 1996; Bresin & Widmer, 2000; Goebl, 2001; Widmer, 2001) and the Yamaha Disklavier (Behne & Wetekam, 1994; Repp, 1995a, 1995b, 1996a, 1996b, 1996c, 1997; Juslin & Madison, 1999; Bresin & Battel, 2000; Timmers, Ashley, Desain, & Heijink, 2000; Riley-Butler, 2001). These devices measure key and hammer movements in order to reproduce a human performance. These pianos are not designed for scientific purpose and their precise functionality is unknown or held back by the companies respectively. Therefore, exploratory studies on their recording and playback precision are necessary in order to examine the validity of the collected data.

Only a few studies provide some systematic information about the precise functionality of these devices. Coenen & Schäfer (1992) tested five different reproduction devices (among them a Bösendorfer and a Yamaha system) on various parameters, but their goal was to evaluate the reliability of these systems for compositional use, their main focus was therefore on the production unit. Bolzinger (1995) performed some introductory tests on a Yamaha upright Disklavier (MX-100 A), but his goal was to measure the interdependencies between the pianist's kinematics, performance, and the room acoustics. Maria (1999) developed a complex methodology to perform meticulous tests on a Disklavier (DS6 Pro), but no systematic results are reported so far.

The system of a Yamaha Disklavier consists of two main parts: (1) the measuring unit, and (2) the reproduction unit. The first stores information derived from two shutters per key: one below the key, and another at the hammer shank. The hammer shank shutter provides two trip points, whereas the one at the key only one. In this design, it is very similar to the Bösendorfer SE system (see Goebl, 2001 for more detailed information), but the authors cannot state anything about the interior processing of the measured data. The reproduction unit possesses solenoids below the back of each key in order to push them as the pianist did. Pedal measurement and reproduction is not discussed in the present study.

2 Aims

In this study we focus on the timing precision of the measuring and the reproduction unit of a Yamaha Disklavier grand piano. It is also important to determine the physical sound properties in relation to the MIDI velocity units, in order to understand their meaning for performance research (Palmer & Brown, 1991; Repp, 1993).

Another issue discussed in the following is the timing behaviour of the grand piano action in the context of different types of touch and their reproduction by a Disklavier. Selected keys distributed over the whole range of the keyboard are played by pianists in as many as possible different intensities with two kinds of touch: once accelerating from the keys (*'legato touch'*), once with an attack from a certain distance above the keys (*'staccato touch'*). These different kinds of touch are described in Askenfelt & Jansson (1991). In this context we try to address the still unanswered question whether there could be a perceivable difference between two differently played keys with the same hammer velocity.

3 Method

For the present study a Yamaha Disklavier grand piano of the Mark II series (DC2IIXG, 173 cm, serial number: 5516392)[1], situated at the Department of Psychology, Uppsala, Sweden, was used. Immediately before the experiment the instrument was tuned and the piano action and the reproduction unit adjusted by a specially trained Yamaha piano technician.

The tested keys were equipped with two accelerometers: one mounted on the key (Brüel & Kjær Accelerometer type 4393) and one on the hammer shank (Picomin Accelerometer Model 22). A microphone placed next to the strings of that key (approx. 10 cm distance) recorded the sound. The instantaneous speed of the key and the hammer as well as the sound were recorded on a multi-channel DAT recorder (TEAC RD-200 PCM data recorder) with a sampling rate of 16 kHz and a word length of 16 bit. The data stream was transferred to sound files and processed with the help of Matlab routines written by the authors. The recordings were preceded by calibration tests, to get meters per second for the hammer and key velocities, and dB SPL for the intensities.

Five keys were tested: C1 (MIDI note number 24), G2 (43), C4 (60), C5 (72), and G6 (91). Each key was hit in as many different dynamic levels (hammer velocities) as possible, in two different kinds of touch: once accelerating from the key ('legato touch'), once hitting the key from above ('staccato touch'). The two authors played these artificial tones. In an informal comparison between the two 'pianists' no difference was found. Parallel to the recordings on the multi-channel data recorder, the Disklavier recorded these test tones with its internal device on a floppy disk. For each of the five notes and the two types of touch, 65 to 113 attacks were performed, in total 972 attacks. Immediately after the recording of each key, the Disklavier reproduced the recorded MIDI file, the key and hammer movements were stored on the data recorder as before.

This method delivered (1) the precise timing and dynamics of the **original recording**, (2) the internally stored **MIDI file** of the Disklavier, and (3) the precise timing and dynamics of the **reproduction**.

To extract discrete data from the hammer and key velocity channels, several signal processing definitions had to be made:

1. The **hammer-string contact** was defined as the time of maximum deceleration of the hammer shank which corresponds well with the physical onset of the sound, and conceptually with the note onset in the MIDI file.
2. As **hammer velocity**, the maximum hammer velocity of the hammer shank before the hammer-string contact was taken.
3. An **intensity value** was derived by taking the maximum of the low-pass filtered (cut-off frequency 30 Hz) audio signal. The audio channel of each file was calibrated with a 1 kHz sinus tone at 94 dB (Sound Level Calibrator Type 4230).
4. The **MIDI onset**, and the **MIDI velocity** number were taken from the MIDI file.

The onset differences between the original recording and the MIDI file, and those between the original recording and its reproduction were calculated[2] and normalised (subtraction of the average difference), since the precise synchronisation between the three files (original – MIDI – reproduction) was not possible by that method.

---

[1] This Mark II XG series was issued by Yamaha in 1997 (information by Yamaha Germany, Rellingen, personal communication, 1999).

[2] $delay_{MIDI}$ = MIDI onset – original onset;
$delay_{repro}$ = reproduced onset – original onset

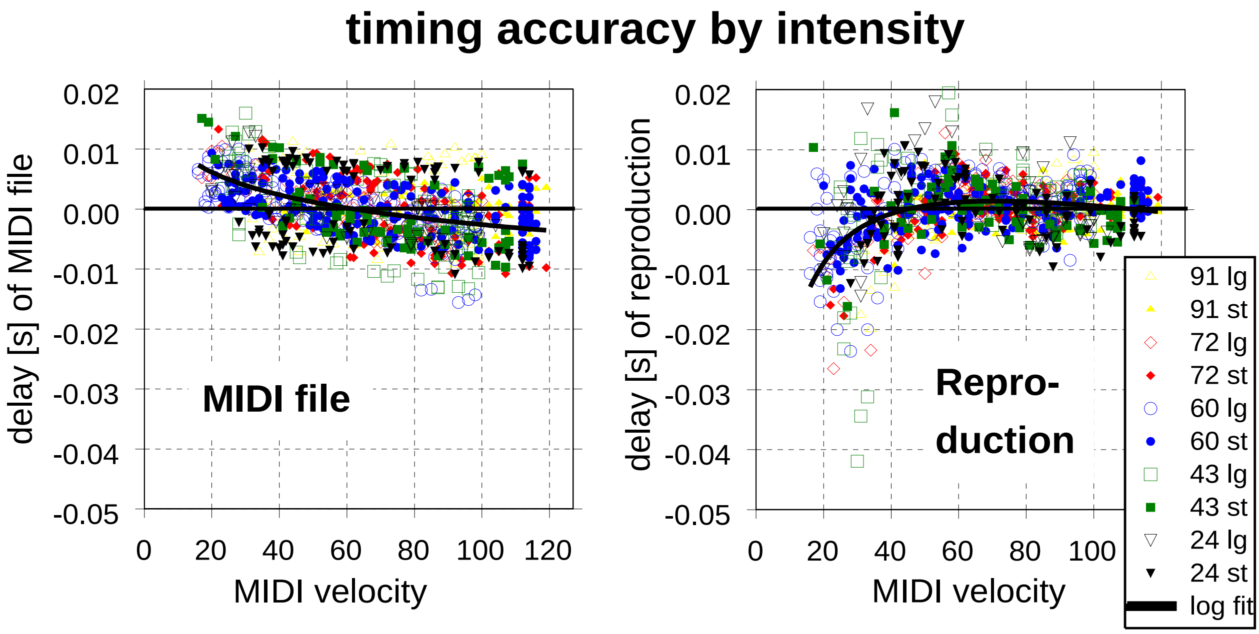

Figure 1 Delay times of the MIDI file (left panel) and the reproduction (right panel) in relation to the original recording. Negative values indicate too early, positive too late.

4 Results

4.1 Timing accuracy

The delay times, sorted by the intensity values, are plotted in . The onsets in the MIDI file are more precise than the onsets of the reproduction. They range between +/–10 ms, with a tendency for soft tones to be recorded later than louder tones. The reproduction shows the same error range for MIDI velocities higher than 60, for softer tones the error goes from +20 ms too late up to 42 ms too early.

The reason for the systematic error in the MIDI file is not clear, since according to information from the literature (Coenen & Schäfer, 1992; Repp, 1996c), the onset is taken from the hammer shutter, as done by the Bösendorfer SE system (Goebl, 2001). However, the too early reproduction of soft tones can be explained either by the faster travel times, especially for legato notes (see Figure 4), or by the too loudly reproduced very soft tones (see next section).

4.2 Dynamic accuracy

The reproduction accuracy of hammer velocity and sound pressure level is plotted in Figure 2. The limitation of the play-back device becomes clear in the left panel of Figure 2. The solenoids are not able to reproduce hammer velocities higher than 2.5 m/s in the bass range, and 3.5 m/s in the treble range. Also, the pianists could produce higher hammer velocities the higher the tones get. This is due to smaller hammer mass towards the treble.

The same picture can be seen at the sound pressure level (Figure 2, right panel), where higher notes achieve a lot higher dB values and their SPL reproduction is accurate over a greater range than for lower tones. Also, very soft notes are reproduced louder than they were originally played (for tones softer than 80 dB).

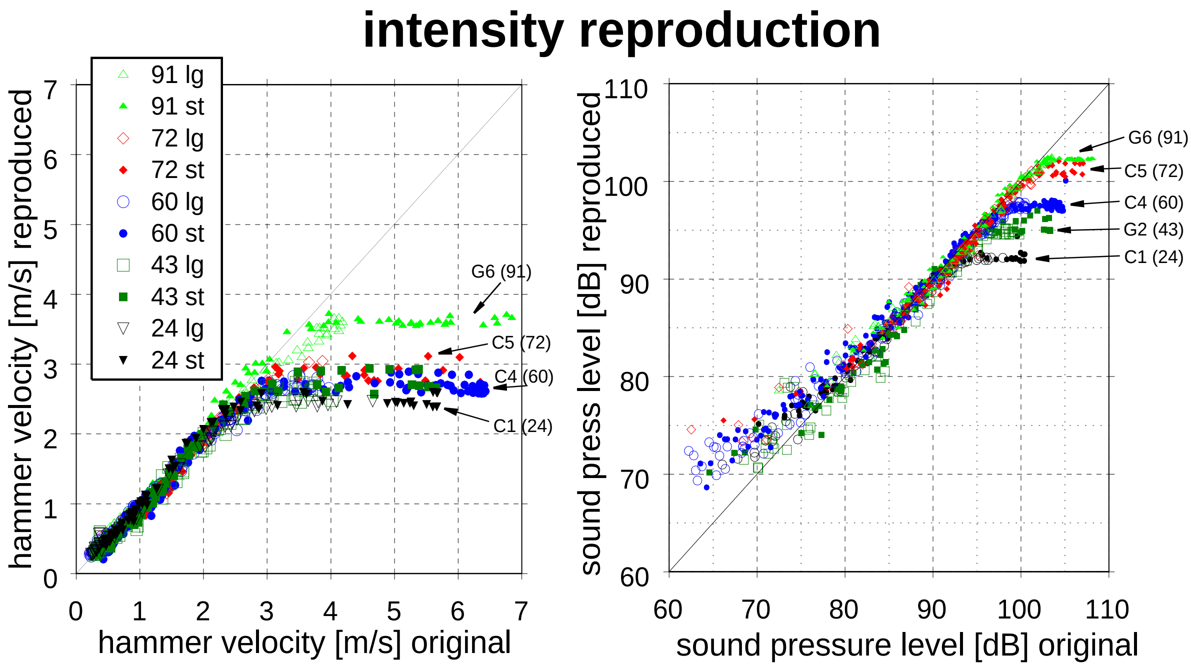

Figure 2 Reproduction accuracy of hammer velocity (left panel) and sound pressure level (right panel).

4.3 Relation hammer velocity to MIDI velocity

The relation between hammer velocity in meters per second and MIDI velocity is well matched by a logarithmic curve (see Figure 3). This figure shows also at what MIDI velocity the solenoids do not increase their drive force anymore. For the lower keys it is already at around 85, the middle keys (G2 – C5) saturate at around 90, only the G6 increases its hammer velocity up to a MIDI velocity value of 100. However, in an informal recording of expressive performances by the authors (over 26000 notes) the majority of MIDI velocity values were between 40 and 80, the highest value was 99, the lowest 6. That means, only in extreme cases (*ff* and *pp*) the piano clearly does not reproduce properly, the middle range seems to be quite accurate.

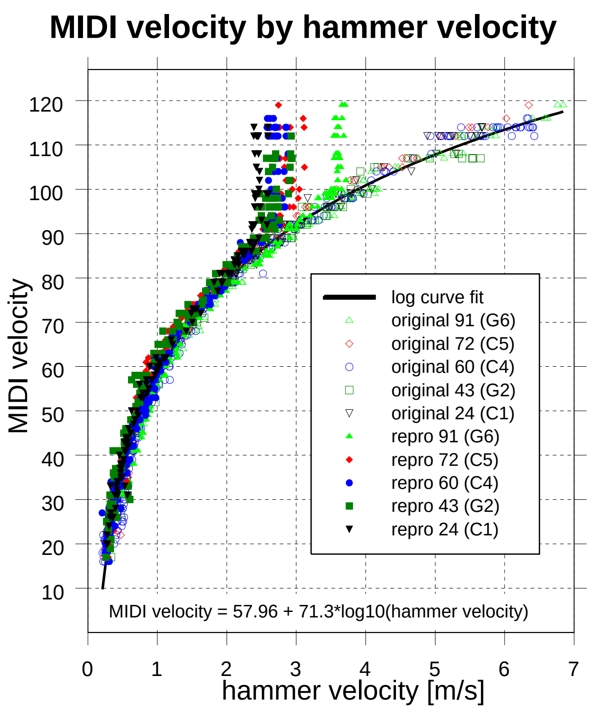

Figure 3 Relationship of hammer-velocity to MIDI velocity of the original performance measured by the Disklavier and its reproduction. A logarithmic curve is fit onto the original data (see formula).

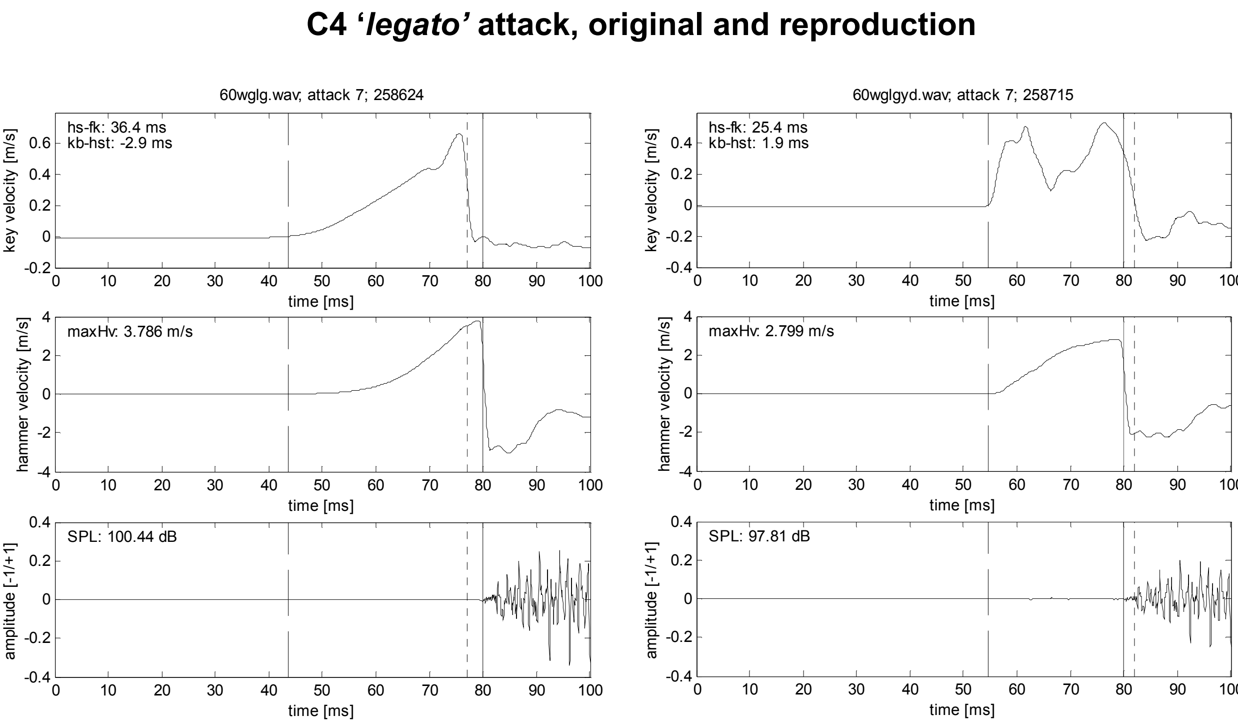

Figure 4 A *forte* attack (C4, MIDI note number 60) played by one pianist (left panel) from the key (*'legato touch'*), and its reproduction by the Yamaha Disklavier (right). The upper panel shows the key velocity, the middle the hammer velocity, the bottom the sound signal. The three lines indicate the finger-key contact (start of the key movement, left broken line), the key-bottom contact (dotted line), and the hammer-string contact (solid line).

4.4 Touch reproduction

In Figure 4 the instantaneous key and hammer velocity with the sound signal is plotted. On the left side the *'legato'* attack as played by one of the authors is shown with its smooth acceleration, on the right its reproduction by the Disklavier. The Disklavier hits the key always in a *'staccato'* manner, with an abrupt acceleration in the beginning of the attack. The whole piano action is compressed, the hammer starts to move slightly later. The solenoid's action results in a shorter travel time (the time between finger-key contact (fk) and hammer-string contact (hs) is 25 ms instead of 36 ms, see Figure 4, upper panel). The travel time difference between production and reproduction is even larger at very soft keystrokes. This could be one reason why soft notes appear much earlier in the play-back than louder notes (see Section 4.1).

The difference in hammer velocity is clearly perceivable in this particular attack; if the hammer velocities become similar, the two sounds, independent of how they were produced (*legato* – *staccato* – reproduced) become indistinguishable.

5 Conclusions

This study is a first attempt to investigate the measuring and reproducing capabilities of a Yamaha Disklavier systematically. With a very precise experimental setup (accelerometers, parallel with the sound signal), the information stored in the MIDI file and the reproduction was measured.

The MIDI file provided more accurate results than the reproduction. Its timing accuracy came up to +/-10 ms which is probably not perceivable by a human listener, but it is astonishingly imprecise, given that similar accuracy could be obtained by analysing a sound file. The MIDI velocity value in the MIDI file appear to be logarithmically matched to the hammer velocity. Sound pressure level is in the middle range approximately linearly related to MIDI velocity (between 40 and 100; this figure not shown), which makes MIDI velocity units a robust representation of intensity. Still its perceptual relevance has to be discussed.

The play-back part of the Disklavier system is the weakest, although it was regulated before the measurements: the timing accuracy decreases for soft notes with a spread of about 60 ms; soft tones are reproduced mostly louder than played originally; and the solenoids do not reproduce MIDI velocities higher than about 90.

The instrument used in this study is not the state of the art in reproducing pianos. Since 1997, when the series of our instrument came out, the "Pro" series (1998), and the "Mark III" series were developed. It is probable that these systems are more precise, thus it remains subject for future work to examine this.

6 Acknowledgments

This study was supported by the European Union (Marie Curie Fellowship HPMT-GH-00-00119-02, and SOb - The Sounding Object project – no. IST-2000-25287; http://www.soundobject.org) and the START program of the Austrian Federal Ministry for Education, Science, and Culture (Grant No. Y99-INF). Special thanks are due to Alf Gabrielsson (Department of Psychology, Uppsala), who provided a well maintained Disklavier for experimental use. Thanks to Anders Askenfelt, Simon Dixon, Alexander Galembo, Eric Jansson, Tore Persson, and Giampiero Salvi for stimulating discussions and helpful suggestions.

7 References

Askenfelt, A., & Jansson, E. V. (1991). From touch to string vibrations. II. The motion of the key and hammer. *Journal of the Acoustical Society of America, 90*(5), 2383-2393.

Behne, K.-E., & Wetekam, B. (1994). Musikpsychologische Interpretationsforschung: individualität und Intention. In K.-E. Behne & G. Kleinen & H. d. la Motte-Haber (Eds.), *Jahrbuch für Musikpsychologie* (Vol. 10, pp. 24-32). Wilhelmhaven: Noetzel.

Bolzinger, S. (1995). *Contribution a l'étude de la rétroaction dans la pratique musicale par l'analyse de l'influence des variations d'acoustique de la salle sur le jeu du pianiste.*, Unpublished doctoral thesis, Université Aix-Marseille II, Marseille.

Bresin, R., & Battel, G. U. (2000). Articulation strategies in expressive piano performance. *Journal of New Music Research, 29*(3), 211-224.

Bresin, R., & Widmer, G. (2000). Production of staccato articulation in Mozart sonatas played on a grand piano. Preliminary results. *Speech, Music, and Hearing. Quarterly Progress and Status Report, 2000*(4), 1-6.

Coenen, A., & Schäfer, S. (1992). Computer-controlled player pianos. *Computer Music Journal, 16*(4), 104-111.

Goebl, W. (2001). Melody lead in piano performance: Expressive device or artifact? *Journal of the Acoustical Society of America, 110*(1), 563-572.

Juslin, P. N., & Madison, G. (1999). The role of timing patterns in recognition of emotional expression from musical performance. *Music Perception, 17*(2), 197-221.

Maria, M. (1999). *Unschärfetests mit hybriden Tasteninstrumenten.* Paper presented at the Global Village - Global Brain - Global Music. Klangart Kongreß 1999, Osnabrück.

Palmer, C. (1996). On the assignment of structure in music performance. *Music Perception, 14*, 23-56.

Palmer, C., & Brown, J. C. (1991). Investigations in the amplitude of sounded piano tones. *Journal of the Acoustical Society of America, 90*(1), 60-66.

Repp, B. H. (1993). Some empirical observations on sound level properties of recorded piano tones. *Journal of the Acoustical Society of America, 93*(2), 1136-1144.

Repp, B. H. (1995a). Acoustics, perception, and production of legato articulation on a digital piano. *Journal of the Acoustical Society of America, 97*(6), 3862-3874.

Repp, B. H. (1995b). Expressive timing in Schumann's "Träumerei:" an analysis of performances by graduate student pianists. *Journal of the Acoustical Society of America, 98*(5), 2413-2427.

Repp, B. H. (1996a). The art of inaccuracy: Why pianists' errors are difficult to hear. *Music Perception, 14*(2), 161-184.

Repp, B. H. (1996b). The dynamics of expressive piano performance: Schumann's "Träumerei" revisited. *Journal of the Acoustical Society of America, 100*(1), 641-650.

Repp, B. H. (1996c). Patterns of note onset asynchronies in expressive piano performance. *Journal of the Acoustical Society of America, 100*(6), 3917-3932.

Repp, B. H. (1997). Some observations on pianists' timing of arpeggiated chords. *Psychology of Music, 25*, 133-148.

Riley-Butler, K. (2001). Comparative performance analysis through feedback technology, *Meeting of the Society for Music Perception and Cognition (SMPC2001), 9-11 August 2001* (pp. 27-28). Kingston, Ontario, Canada.

Timmers, R., Ashley, R., Desain, P., & Heijink, H. (2000). The influence of musical context on tempo rubato. *Journal of New Music Research, 29*(2), 131-158.

Widmer, G. (2001). Using AI and machine learning to study expressive music performance: Project survey and first report. *AI Communications, 14*, in press.