IDENTIFICATION OF THE CONSEQUENCES OF A PIANIST’S MOTIONS ON HIS/HER SOUND BY MEANS OF PSYCHOACOUSTIC TESTS AND MOTION CAPTURE.

Jean-Christophe Valière, Camille Lefèvre and Floren Colloud
CNRS – UPR 3346, Université de Poitiers, ENSMA, Poitiers, France
e-mail: jean.christophe.valiere@univ-poitiers.fr

Alain Villard
Conservatoire à Rayonnement régional du Grand Poitiers, Poitiers, France
e-mail: alain.villard@poitiers.fr

For musicians, musical experts or critics, it appears obvious that the sound produced during musical performance is highly dependent on the performer’s movements. If this hypothesis is easily admitted in the scientific community when the body is in direct contact with the parts of the instrument that produce the sound (strings, pipes ...), it is less obvious when the contact is through a sophisticated mechanism system as for the case of the modern piano. In first analysis, the variability of the sound seems to depend exclusively on the acceleration of the hammer induced by the finger on the key. In fact, this is an apparent contradiction because the sound of the piano, as in most other instruments, cannot be reduced to a simple sequence of notes. The variability of playing the same piece is infinite even with the use of very precise music scores. For instance, how the notes are connected (slightly superimposed) or the use of pedals are some of these variables. But can these very small timing variations allow the perception of an overall sound difference between two pianists playing the same piece? To answer this question, an experiment has been conducted which includes acoustic pressure measurements inside and outside a concert piano, two artificial heads and motion capture of the whole body of the pianist with a focus on fingers. The aim of this experiment is to find which movements are responsible for the general impression of the sound of a pianist. This paper will focus on our first results obtained in psychoacoustics, that partially prove that the subjects identify the difference between pianists when the same musical phrase is played, but this identification decreases when pianists play different ones.

1. Introduction

For pianists, musical experts and critics it seems obvious that the particularity of pianist gesture induces an identifying sound associated to each one. By comparison with instruments for which the body directly interacts with the resonant part of the instrument (strings, pipes ...), it seems more speculative for the modern piano for which the sound is obtained by means of a very sophisticated mechanical system.

Due to its status in Western music, the piano has been the subject of a long-standing attention both on the mechanical aspects (vibratory and acoustic) and perception. The operating chain of piano sound production could be described as follows. The pianist produces an acceleration on the piano key, with a certain freedom in the position of contact, which is transmitted via a complex hammer mechanism that
strikes the string and relaxes. The action of removal of the key brings a damping, except if the foot pedal is actuated by the musician, in which case the damping buffers remains raised. According to the frequency domain, the hammer impacts a string (low frequencies), two strings (medium) or three strings (treble) to compensate for both the intensity of the sound but also the variation of inertia and therefore the damping of sound. The goal of the designers was to make the piano tessitura as homogeneous as possible. The vibrations of the strings are transmitted to the soundboard via the bridge, which plays the classic role of amplifier as for most string instruments. It is the essential source of radiation that continues through the geometry of the piano.

The complex mechanisms of sound production from the hammer [1] to the radiation have been studied, as well by experiments [2] than numerical simulation [3], avoiding as far as possible to consider the variability of the striking of the piano key. This could be easily understood because of the non-linearities, which appears at each step of this sound production. For a complete synthesis, see the recent book of Chaigne and Kergomard [4].

If the mechanical process of producing sound, from the key to the soundboard, has been well studied for a long time, the sound produced cannot be summed up by simply superimposing the successive states of the action on the keys. In other words, the variability of the possible attacks, the connection between the notes, and the use of the pedal over longer or shorter times leads to a specific sound.

Thus, a number of studies have focused on questions of global perception or timbres regardless of the issues of physical production. The most advanced work in the field of timbre characterization has identified 14 adjectives, which could be reduced in five major classes [5]. These timbres have been linked to the different actions performed on the piano (speed on the key, pedal action ...) [6,7]. These studies showed that the performances of the pianists shared common traits, characteristics of the structure of the piece interpreted, or, characteristics of the shades they want to express.

These gestural analyzes highlight common biomechanical and cognitive constraints that should condition the movements of all pianists. In order to try to highlight this fact, for the first time an experiment associating gesture capture gestures and sound recording was made in an auditorium dedicated to orchestral music with four pianists playing the same works on the same piano in exactly the same acoustic space [8,9]. They played several short works each their turn, their movements and the resulting sounds being recorded by a motion capture system and artificial heads, respectively.

In this paper, we will present first the experimental setup and the analyzed data. Then the psychoacoustic tests put in place on these data will be presented. Forty-five subjects did these tests with a panel of non-musicians, amateur musicians and expert musicians. The analysis of the results will be presented as well as future perspectives.

2. Description of the Experiment

2.1 Basic principles

The experiment which was carried out on 21 and 22 December 2017, has been fully described in the article of Villard et al. [8] and a video filmed by CNRS is available on line [9]. The experimental set-up is displayed on Figure 1.

Making a complete data acquisition of the entire operating chain seemed to us too tedious, it was thus privileged the direct relation between gesture and sonority without taking data on the mechanics of the system which has already been well studied. It was therefore necessary, as much as possible, to neutralize the variations of the mechanical aspects (vibratory and acoustic) in the operating chain. For this, a precise repertoire was defined as it will be described in the next section, the piano, the room and the measuring devices were not modified during the experiment.

The experiment has been achieved in conditions close to those of a concert. We benefit from a Steinway and Sons piano model D (Concert Grand) lent by the Theater Auditorium of Poitiers (TAP). It is ten
years old and was tuned for this occasion by his usual tuner. This piano was placed on the stage in the configuration "Recital", that is to say in the center of the scene space.

It seemed important to focus our experience around a musical work of indisputable artistic quality. We cannot dissociate our research from a requirement and a real artistic content. For this, we chose four pieces from Kinderszenen (Scenes from Childhood) op. 15 by Robert Schumann [10]: Von fremden Ländern und Menschen (Of Foreign Lands ans Peoples), Kuriose Geschichte (A curious Story), HaschecMann (Blind Man’s Bluff) and Wichtige Begebenheit (An Important Event). These four pieces, beyond an obvious artistic content, had the interest to be brief, to offer much differentiated modes of play on which we could consider obtaining sufficiently clear and distinct findings on the gestures used by the pianists and their sound renders, in order to be differentiated by auditory tests, signal processing and gestural analysis.

2.2 Acoustic recordings

The acoustic recordings were aimed at two sets of data. On the one hand, usable data in terms of signal processing with possible future interactions with vibroacoustic aspects and, on the other hand, usable data for listening tests of the type performed in psychoacoustics.

For the first data set, two sound recordings were made using microphone pairs, the first near the impact of the hammers to be able to accurately estimate the impacts of the notes (better transient detection) and the other at a distance of 2 m from the piano to capture the radiation of the piano as a whole and as it propagates in the room.

The second set of acoustic data is produced by two artificial heads (Head-Acoustics®), one positioned next to the pianist to reproduce the audition of the player and the other near the microphone pair reproducing the perception of a person near the piano.

It is clear that the sound depends on the room but it seemed preferable to avoid to achieve the experiment in an anechoic place for fear that the pianists compensate for the lack of naturalness by more provided or more resonant, as per example a too long use of the sustain pedal. In parallel with these tests, the main auditorium characteristics were measured, the results give a reverberation time of 1.8 s calculated on average on the bands [250 Hz - 2000 Hz], compliant for a dedicated auditorium orchestral music of this size (1000 people), and a great homogeneity of the reverberation time in octave bands from 125 Hz to 4000 Hz. The proximity of the acoustic sensors makes it possible to have the direct sound rather than the reverberant sound.

2.3 Pianist's movements recordings

The pianist's movements recordings are carried out simultaneously to the acoustic recording, and includes the accurate characterization of the attack of the key, in order to assess its relationship to the temporal and / or frequency variations of the sound. The aim is to establish the relationships between these different variables.

Ninety passive makers placed on the full body equip the pianist. These markers are positioned on a garment (upper and lower limbs, head and trunk) and directly on the skin for the hands and fingers (48 hemispherical markers of 3 mm in diameter). Markers are also positioned on a few keys of the piano to be able to reposition the pianist with respect to the instrument. Their spatial evolutions are recorded by a motion capture system composed by 15 high frequency cameras (Oqus 7, Qualisis. Sweden). The motion capture system is synchronized to the acoustic system and the sampling frequencies are in a multiple ratio to facilitate the tracking of time events, i.e. 48 kHz for acoustics and 240 Hz for motion capture. The compliance of the fingers of the pianists is also collected to be able to characterize the structures contributing to generate force and speed of the fingers.
3. Psychoacoustic tests

3.1 Methodology

The approach chosen for data analysis is to start with auditory evaluation and to end with gestural analysis as shown on Figure 2. First, it is necessary to find out perceptive invariants or to determine the variability into the excerpts or the pianists with regard to timbre. Second, from this first analysis, indicators must be found with signal processing analysis of the sounds taken close to strings. Third, the analysis of pianist’ movements have to be both correlated with these perceptive and signal processing indicators.

For the auditory tests, a panel of 45 people volunteering have participated. This panel includes 10 non-musicians, 23 non-professional musicians, 10 expert non-pianists (professional musicians or high musical level students) and 2 professional pianists. The tests were considered as difficult by all the listeners, naive subjects or not.

We call “musical phrase”, eight measures (for example) of a music score and “excerpt” one realization of this phrase. We have selected five different musical phrases (p = 1 to 5), lasting from 5 s to 10 s, in the 4 songs that have been interpreted. For each piece and each pianist, we selected the recording which was of the best musical quality possible. Finally, for each pianist (i = 1 to 4) we extracted in the selected recordings two excerpts (k = 1 or 2) of each of the selected musical phrases (p). These two excerpts correspond for all pianists to the same measures and the same resumption of these measures (either played for the first time, or for the second time in the case where there is a recovery). So, the subjects will listen sometimes two excerpts of the same pianist in order to check their reliability. We thus obtain a corpus of 5 sentences times 2 excerpts times 4 pianists, that is 40 excerpts.

Two tests have been conducted. The first on, called “Test #1” is a “pair differentiation” test for which two excerpts are listened, and the subject has to recognize if the player is the same. The second test called “Test #2” is a test of “verbalization of timbres”, by using adjectives previously defined in the literature.
Figure 2: Description of the operating chain of piano production, data captures. Below, the methodology of data analysis which is performed backwards from the sound production.

### 3.2 Pairwise Differentiation (Test #1)

In this test, we propose to the subject a series of pairs of excerpts to listen to. For each pair, the listener is asked to evaluate whether the two excerpts are played by the same pianist or by two different pianists. The test is divided into two parts:

- **Part #1**: The excerpts to be compared correspond to the same musical phrase,
- **Part #2**: The excerpts to be compared correspond to two different musical phrases.

In the Part#1, only the musical phrases $p = 1$ and $2$ are used. The pairs compare two excerpts from the same musical phrase: either phrase #1 or phrase #2. Therefore, we have in this part three types of different pairs:

- Pairs of the same excerpt (so necessarily same pianist)
- Different pairs of the same musical phrase but two excerpts of the same pianists
- Pairs of different pianists.

The aim of this test is to observe if the listeners recognize a player independently on the excerpt of the musical phrase.

In the Part#2, the excerpts to be compared do not correspond to the same musical phrase. The goal here is to try to see if, beyond the melodic differences, the pianists keep a specific musical signature that allows them to be distinguished from each other.

We have subdivided this part into two subparts: the first subpart (2.1) proposes to compare excerpts of the phrase #1 with excerpts of the phrase #2, while the second subpart (2.2) part proposes to compare these two sentences #1 and #2 with phrase #3, that the subjects have not heard yet. Therefore, in these two subparts we have two types of pairs:

- Pairs of the same pianist
- Different pairs of pianists.
All the pianists are compared the same number of times to each of the other pianists and the same number of times to themselves. However, to avoid having too many pairs, not all possible combinations are tested and the pairs are randomly selected for either the musical phrase, the pianist or the excerpt. This random selection is made for each subject, so each of them is offered different pairs. In the same way, the order of the pairs and the order of the excerpts within each pair is generated randomly for each subject. Thus, the statistics will not depend on the order or the composition of the corpus.

3.3 Verbalization (Test #2)

Previous studies achieved with seventeen expert pianists permit to obtain a list of 14 adjectives describing the characteristic timbres of piano [5]. With a calculation of semantic proximity by using Principal Component Analysis of the results, the authors reduce this selection of 14 adjectives to 5 classes for which the adjectives is considered to be close by the subjects [6,7].

Table 1: Timbre qualification among authors [5-7]

<table>
<thead>
<tr>
<th>Bright</th>
<th>Clear</th>
<th>Shimmering</th>
<th>Brassy</th>
<th>Harsh</th>
<th>Dry</th>
<th>Metallic</th>
<th>Soft</th>
<th>Velvety</th>
<th>Full</th>
<th>Round</th>
<th>Distant</th>
<th>Dark</th>
<th>Molded</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bright</td>
<td>Dry</td>
<td>Velvety</td>
<td>Round</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In the present test (Test#2), the subject listens the excerpts one by one and has to assign each of them 1 to 3 adjectives among the 14 proposed, in order to describe the sound. For this test, we choose to test the 5 musical phrases we have selected. On the other hand, to reduce the duration of the test, we are still forced to consider, for each subject, for each pianist only one excerpt of each phrase. The excerpts appear randomly and the results to be analysed are composed of a series of couple pianist/phrase characterized by 1 to 3 adjectives.

4. Results of auditory tests

Table 2 shows all the results for the Test#1. The results should be interpreted taking into account that the value of 0.5 corresponds to full random. When it is over 0.5 with a low standard-deviation (std) the test is successful. When the result is under 0.5 with low std, listeners find the opposite of the true value. Fortunately, this extreme case does not occur. When it is around 0.5 with high std the results are not conclusive.

Table 2: Rate of good results of identification and standard deviation.

<table>
<thead>
<tr>
<th>Kind of pair</th>
<th>Non-musicians (Eff. 10) Mean (std)</th>
<th>Musician Non-experts (Eff. 23) Mean (std)</th>
<th>Musician Experts (Eff. 10) Mean (std)</th>
<th>Pianist Expert (Eff. 2) Mean (std)</th>
<th>Global (Eff. 45) Mean (std)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Same Phrase (part #1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Same Excerpt Different Pianist</td>
<td>0.83 (0.16)</td>
<td>0.86 (0.19)</td>
<td>0.8 (0.18)</td>
<td>1 (0)</td>
<td>0.84 (0.18)</td>
</tr>
<tr>
<td>Different Excerpt/ Same Pianist</td>
<td>0.48 (0.18)</td>
<td>0.46 (0.17)</td>
<td>0.46 (0.16)</td>
<td>0.56 (0.44)</td>
<td>0.47 (0.18)</td>
</tr>
<tr>
<td>Different Excerpt/ Different Pianists</td>
<td>0.83 (0.1)</td>
<td>0.79 (0.15)</td>
<td>0.83 (0.12)</td>
<td>0.88 (0.06)</td>
<td>0.81 (0.13)</td>
</tr>
<tr>
<td>Two different Phrases (part #2.1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Same Pianist</td>
<td>0.58 (0.24)</td>
<td>0.54 (0.22)</td>
<td>0.55 (0.2)</td>
<td>0.5 (0)</td>
<td>0.55 (0.21)</td>
</tr>
<tr>
<td>Two different Pianist</td>
<td>0.49 (0.19)</td>
<td>0.52 (0.17)</td>
<td>0.54 (0.07)</td>
<td>0.5 (0)</td>
<td>0.52 (0.15)</td>
</tr>
<tr>
<td>Three different Phrases (part #2.2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Same Pianist</td>
<td>0.55 (0.15)</td>
<td>0.58 (0.16)</td>
<td>0.55 (0.22)</td>
<td>0.56 (0.09)</td>
<td>0.56 (0.17)</td>
</tr>
<tr>
<td>Two different Pianist</td>
<td>0.52 (0.16)</td>
<td>0.53 (0.16)</td>
<td>0.58 (0.14)</td>
<td>0.53 (0.12)</td>
<td>0.54 (0.15)</td>
</tr>
</tbody>
</table>
The first remark is that the tendencies are quite similar whatever the categories of listener. This fact is really singular and proves the difficulty of such test. The differentiation by pair usually gives better results for experts. In part #1, the listeners recognised quite well the pianists when only one excerpt was played or when different pianists (soft grey underline) played excerpts. On the other hand, the test is undermined (around 0.5) when two excerpts of the same phrase are played by the same pianist. A partial conclusion could be that the listeners focused on differences between pairs more than on difference of pianist playing.

In the part #2, two different phrases were played by two pianist or by the same pianist. Globally, the test is difficult and the results are over 0.5 but with a high std. Nevertheless, the results are a little better when the two musical phrases were played by the same player (grey underline). The listeners, on this part #2, focused more on the difference of timbre and had not to compare differences between excerpts of the same musical phrase. It is probably why the results of this part are better than those of the part #1 with the same pianist and different excerpts.

For the Test #2, no reference was given concerning the pianist. The listeners only have to choose one to three adjectives among a series of 14. Figure 3 left displays the occurrence of all the adjectives used whatever the pianists and the musical phrases, compared to the occurrence found out by [5].

One surprise is the good similarity between the two distributions at the exception of “dark” which occurs less. The occurrence of the adjective is obviously both dependent on the musical choice and on the pianist style. It confirms that our experiment is sufficiently various to obtain large amount of timbres. In the same time, it confirms results of the literature. When correlating the occurrence of the 14 adjectives found in literature with the occurrence found during the present experiment, the coefficient is 0.78. Comparing to the 5 families of adjectives yields and increase of the correlation to 0.87. It confirms the pertinence of the use of only 5 adjectives found out by authors.

Figure 3 right displays the timbre profile of the pianists whatever the musical phrase. It shows that each distribution differs notably from the mean (yellow bars). For instance, for the bright timbre, pianist 2 and 4 had a high value, 10% more from pianist 1 and 3. For dry timbre, pianist 3 and 4 differs from 15%.
5. Perspectives

The first auditory test does really permit to distinguish the pianists, because the listeners probably focused more their attention on the little difference of the interpretation than of the players. The second test gives more promising results regarding this first analysis. First, it confirms in several points the results of the literature which proves the pertinence and coherence of our experiment. Second, the differentiation of the pianists regarding their timbres appears clearly. This will permit to carry on with the finding of signal processing indicators as well as those related with the pianist’ movements.

6. Acknowledgment

The authors acknowledge Catherine Schneider, Pierre-Yves Jalicon, Alan Kenneth, Frédéric Palpacuer, Danae Le Guennic, Sébastien Noly, Jean-Pierre Manceau, Maëva Retailleau, and Joël Rakoto Ramiarantsoa, the « Théâtre Auditorium/Scène Nationale de Poitiers », the « Conservatoire de Région », the « CNRS » and the Institut Pprime (ACI Grant).

REFERENCES

10 Schumann, R. Kinderszenen (Scenes from Childhood), op. 15, Henle Verlag, HN 9924 (2009 [1839]).