An Exploratory Data Analysis of the Ability of Western Musicians to Cope with Latency

Jorge Medina Victoria[1,2]

[1] Darmstadt University of Applied Sciences, Darmstadt, Germany
[2] Cork Institute of Technology, Cork, Ireland
jorge.medina.victoria@h-da.de

Abstract. Technology based on networks has long ago become state of the art for music recording, concert venues and even as a medium for collaborative performances over the Internet. However, there is a drawback attached to this technology. Latency is a permanent issue associated to any network technology. On the other hand, musicianship relies on time perception. The performance of music and playing an instrument are activities highly affected by latency. The present paper describes a methodology in order to measure how latency and the performed western musical instruments are related. In addition, the latency tolerance range (LTR) is presented as a comparison measurement between the different musical instrument groups. The results are based on the exploratory data analysis of the gathered data.

Keywords: Latency · Latency Tolerance Range · Tempo · Musical Instrument Groups.

1 Introduction

Research on the issue of latency and musicianship is focused on timing and the improvement of network technical conditions, in order to enable collaborative performances [7, 3, 4]. This paper presents the exploratory data analysis of empirical observations of solo-playing musicians performances known as non-collaborative performances. To gather the data, a methodology was developed using a listening test setup and a questionnaire. The exploratory analysis and the numerical results may enhance the knowledge framework regarding the latency issue and the performance of music, enabling comparisons between different musical instrument groups. Beyond a further understanding of musicianship and the relation between performer and musical instrument, the outcome delivers important information for the design of virtual musical instruments where issues, such as haptics, virtual acoustics and immersion play a primary role.

2 Methodology

A total of 31 test subjects playing 17 different western musical instruments took part in a listening test and answered a questionnaire. Every musician played only his or her main musical instrument. The different musical instruments were divided into the groups of aerophones, chordophones, membranophones and idiophones based on the Hornbostel-Sachs taxonomy [8].

The setup of the listening test is described in [6] and consisted of an audio interface, a cardioid pattern microphone, headphones, a laptop with a digital audio workstation (DAW) and a 7-inch monitor as shown in Figure 1. The musician played his/her instrument, the audio signal was recorded and sent back delayed (latency was introduced progressively) to the musician's headphones, until the test subject was not able to play any more and disrupted the performance. Three different metronomes (aural, visual and aural-visual) were used as a control mechanism. In addition, a predefined score was the same for every test subject. The notes of the score defined a specific pitch and it was used to enable equal conditions to all participants while playing the musical instrument, comparisons between results are possible. Some membranophones and idiophones have no definite pitch. However, variations on the timbre of all different instruments are expected due to the spectrum and envelope characteristics. Having more than one frequency component traduces in a vibration response in different places along the inner ear [9]. The impact of the musical instrument frequencies with relation to the latency issue is irrelevant at this level.

Fig. 1. Listening test setup

Every one of the 31 musicians played the music score five different times according to the tempi 90 BPM, 120 BPM, 150 BPM, 180 BPM and 210 BPM, using a different metronome each time. The three metronomes were regulated using the MIDI protocol and have not been delayed. At the end of the listening test, 15 different latency values were gathered (5 tempi and 3 metronomes).

It is possible to define the total latency of the system used in the listening test by means of the following mathematical expression:

$$Lt = La + Lc + Ld \quad (1)$$

L_t is the total latency, which is usually measured in milliseconds. The total latency is the sum of the latency due to the sound transmission L_a, the latency produced by the analog (sound waves) to digital (information in the DAW) to analog (sound waves to the headphones) conversion L_c and L_d which is the network latency simulated in the listening test.

The latencies L_a and L_c can be considered constants. The value L_c was measured and was the same for every listening test (12.208ms). On the other hand, the value of L_a is directly related to the distance of the microphone to the instrument as presented in Table 1. Prior to the beginning of each listening test, the distance microphone to instrument and the sound pressure level (SPL) of the musical instrument were measured, in order to guarantee equality of conditions. In addition to those measurements, the level gain in dB of the microphone and the headphones, as well as the reverberation time (RT60 in seconds) and noise level of the different rooms were measured. The information presented in Table 1 summarizes the range, median, mean, variance and standard deviation for the 31 test subjects (nbr.val).

Table 1. Listening test related measurments.

	Mic. dist. (m)	Inst.dBSPL (A)	Mic. gain (dB)	HP gain (dB)	RT60 (sec)	Room dBSPL(A)
nbr.val	31	31	31	31	31	31
min	0.23	75.10	0.00	-10.00	0.30	31.20
max	0.71	96.40	12.00	10.00	1.64	38.70
range	0.48	21.30	12.00	20.00	1.34	7.50
median	0.39	83.40	10.00	-5.00	0.50	33.50
mean	0.41	85.40	6.52	-3.23	0.51	33.62
var	0.01	29.48	24.26	57.58	0.06	5.82
std.dev	0.11	5.43	4.93	7.59	0.25	2.41

Based on the information presented in Table 1 it is clear that all of the instruments were recorded under similar conditions. Furthermore, the results of the first column have a very low variance, which means that the distance between microphone and instrument was almost the same. In other words, L_a, which is the latency produced due to sound transmission, can be assumed as a constant.

The value L_d, which is the simulated network latency, is the numerical outcome of the listening test. As stated before, musicians played the score while listening to their own delayed[3] signal (latency was introduced). The score consisted of two bars and was played in a loop. Every second repetition (4 bars), the delay was increased by 10ms. The latency range L_d variated from 0 to 300ms. At 0ms only L_a and L_c were present and imperceptible for the majority of musicians. The latency value L_d was obtained by a performance disruption, when the musician interrupted the performance the value L_d was notated.

3 Exploratory Data Analysis

After having collected the different latency values L_d, in addition to further numerical information of the questionnaire such as age, hours of instrument practice with and without metronome and years of experience on the instrument, it is possible to visualize the information in a scatter plot as presented in Figure 2.

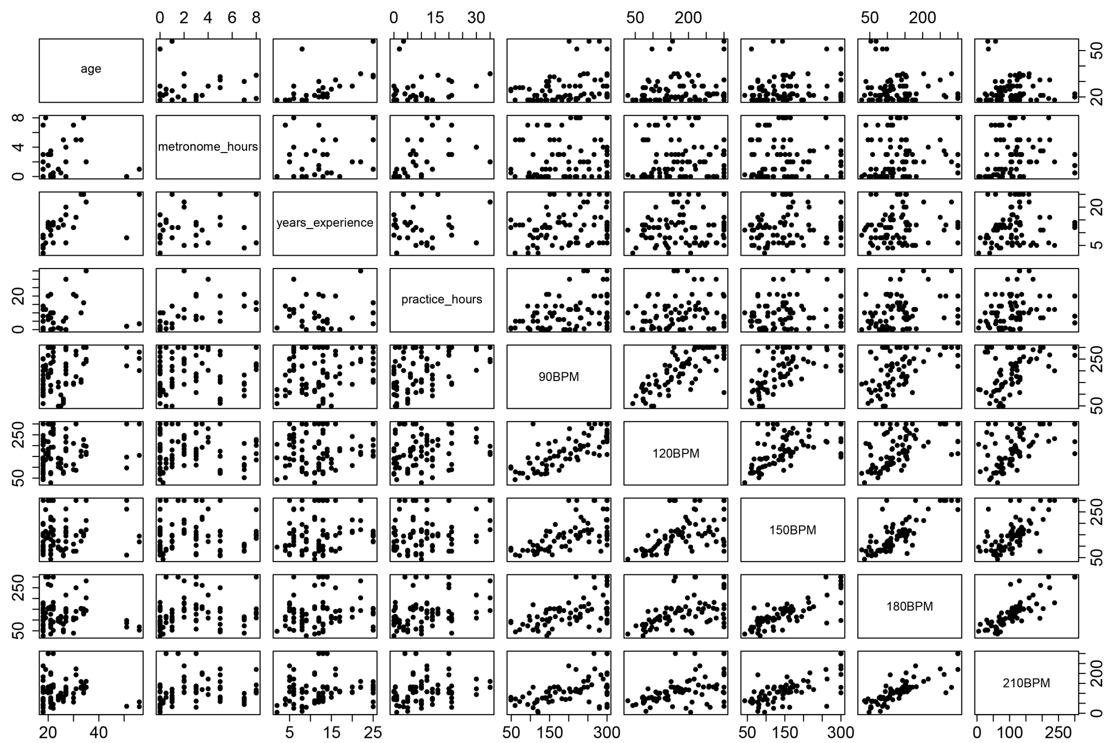

Fig. 2. Scatter plot matrix for all numerical variables

[3] The concepts of latency and delay are the same for this work.

In Figure 2, every numerical variable is plotted against each other. The variables are written on the diagonal and all relationships above and under the diagonal are mirrored. The numerical scale for age ranges from 0 to 50 years, however mainly young European musicians took part in the research. The scale for the years of experience ranges from 0 to 25 years, practice hours with metronome from 0 to 8 hours and without (practice hours) from 0 to 30 hours. Finally all latency values (L_d) distributed according to the five different tempi (90 BPM to 210 BPM) have the range 0 to 300 milliseconds.

The linear trends in Figure 2 are easy to identify for columns five to nine, which are related to the latency values L_d. Additionally, another linear relationship is present between age and years of experience (column 3, row 1) which is obvious. The older a musician is, the more years of experience he or she has. On the other hand, it is interesting that no visual relationships exists between the years of experience, musical practice with and without metronome and the latency values L_d for the different tempi. It may indicate that those variables have no meaningful role at all, regarding the ability to cope with latency.

A further procedure is to visually analyse the shape of the distribution. Figure 3 presents the explorative data analysis (EDA) for the distribution of the gathered data.

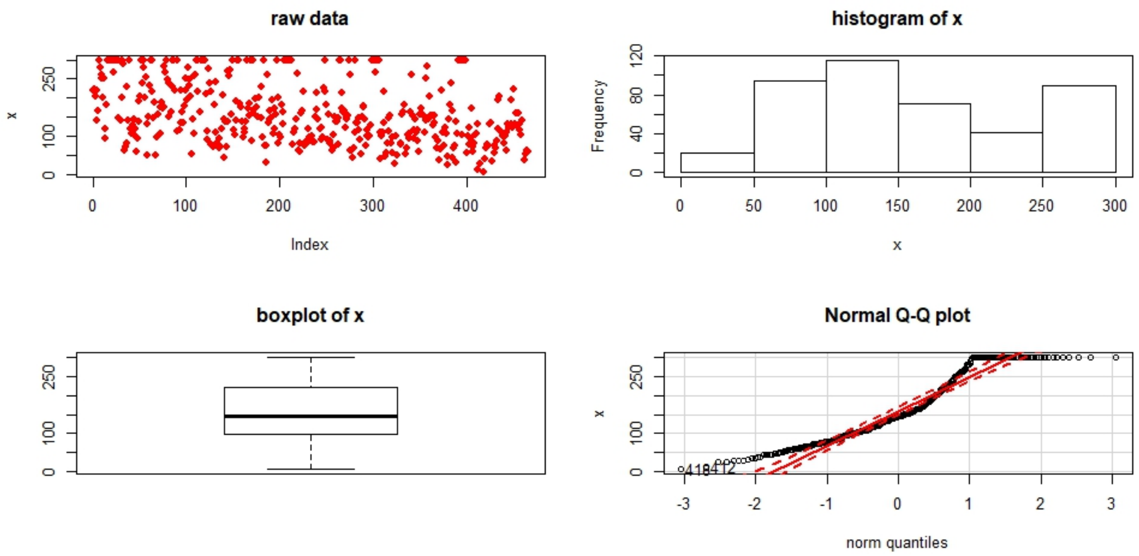

Fig. 3. EDA for the obtained data distribution

From the raw data plot in the upper left of Figure 3, two details are clear. Firstly, the data has not a random distribution and secondly, the data shows a decaying behaviour. The next plot on the right is a histogram, additional to the observation regarding a non-normal distribution. It is also possible to see the outliers (data in the 300 bar). Finally, the lower right plot (Q-Q plot) confirms that the data is non-normally distributed. The confidence intervals (red dashed lines) are clearly overpassed.

Table 2 lists all musical instruments that have been analysed. The selection of test subjects was not completely random. This purposive sample represents a virtual population and constitute a random sample [2].

Table 2. List of musical instruments

Subject	Instrument	Group	Subject	Instrument	Group
1	Piano	Chordophones	17	Trombone	Aerophones
2	Piano	Chordophones	18	Violin	Chordophones
3	Cello	Chordophones	19	Triangle	Idiophones
4	Cello	Chordophones	20	Tenor saxophone	Aerophones
5	Classical guitar	Chordophones	21	Classical guitar	Chordophones
6	French horn	Aerophones	22	Violin	Chordophones
7	Alto saxophone	Aerophones	23	Snare drum	Membranophones
8	Violin	Chordophones	24	Alto saxophone	Aerophones
9	Trumpet in B	Aerophones	25	Triangle	Idiophones
10	Snare drum	Membranophones	26	Marimba	Idiophones
11	Piano (upright)	Chordophones	27	French horn	Aerophones
12	Snare drum	Membranophones	28	Double bass	Chordophones
13	Violin	Chordophones	29	Harp	Chordophones
14	Transverse flute	Aerophones	30	Bassoon	Aerophones
15	Trombone	Aerophones	31	Tenor saxophone	Aerophones
16	Timpani	Membranophones			

In Figure 4, data from the three different metronomes (aural, visual and aural-visual) is presented and data from similar instruments, e.g. piano (1), piano (2) and piano upright (11) have been averaged. Figure 4 represents the latency value (L_d) vs. tempo in BPM. This relationship is known from Barbosa [1] as the latency adaptive tempo (LAT).

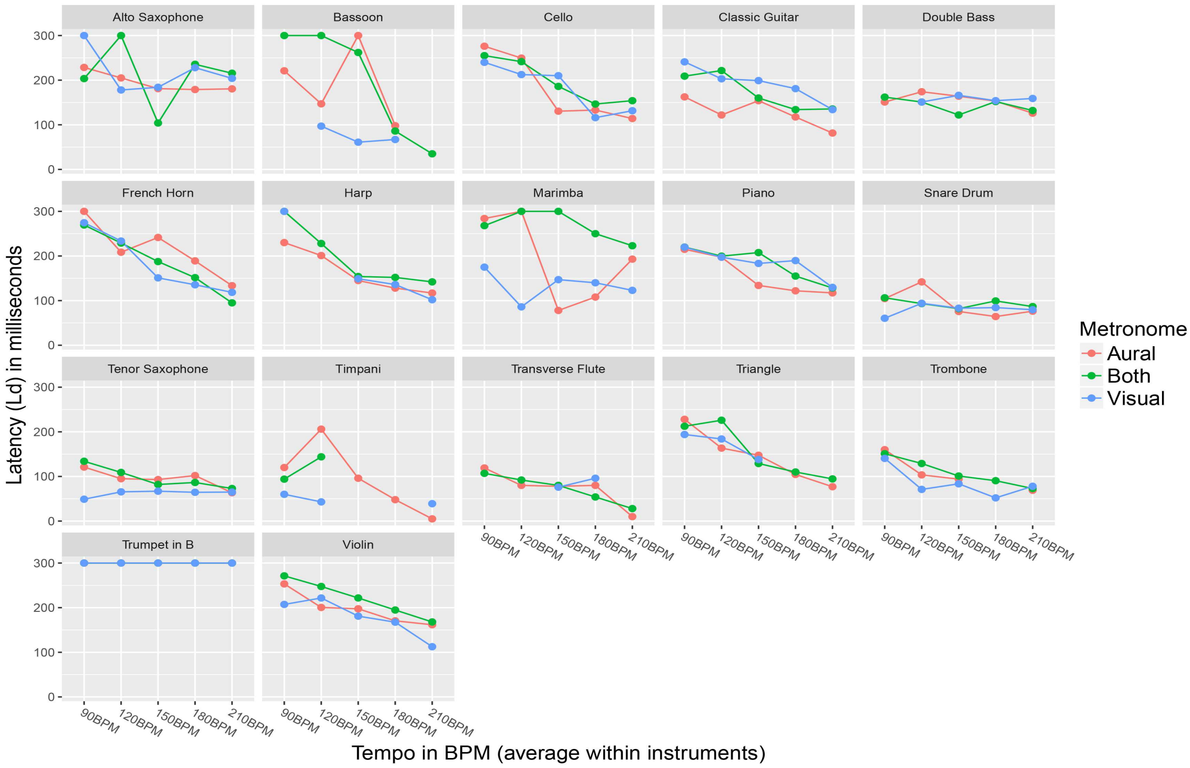

Fig. 4. Latency values (L_d) for the three metronomes (aural, visual and both) according to tempo

It can be noticed that there are very few differences between the results of the different metronomes (except for instruments such as marimba, bassoon and alto saxophone), therefore, it may be possible to average the latency results L_d of the three different metronomes. Observing Figure 4, it is clear that choosing a western musical instrument has an impact regarding the ability to cope with latency. For example, a musician playing the snare drum tends to interrupt the performance earlier (lower latency value L_d) compared to a cello player.

The data of all musical instruments, according to the different five tempi of the experiment could be presented using the kernel density estimation (KDE). The KDE is a smoothed version of the histogram [5] and determines the shape of distribution. In Figure 5 the kernel density estimation for a continuous value L_d is presented.

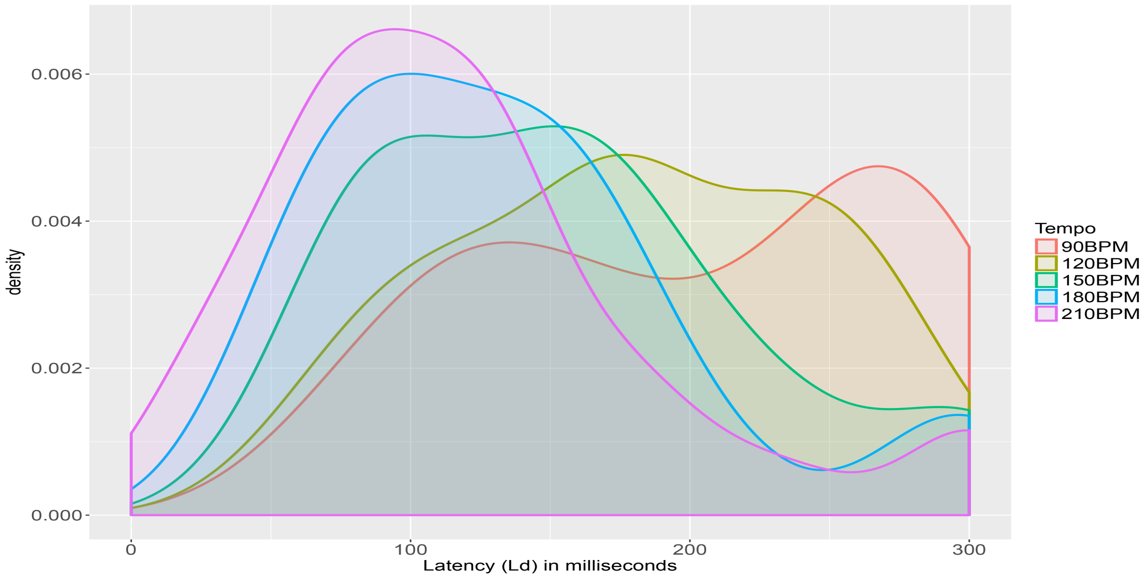

Fig. 5. Density plot for all instruments according to the different tempi (average of all 3 metronomes)

It is important to state that the listening test delivered only five discrete latency values L_d, according to every tempi and to every metronome, as seen in Figure 4. For the Kernel density estimation (KDE) in Figure 5, the values of every metronome (aural, visual and aural visual) were averaged.

Figure 5 clearly shows the peaks for the latency values L_d. The higher the tempo (210 BPM and 180 BPM) the narrower the peak. It indicates that by playing at higher tempi, all musicians stop the performance at approximately the same latency value L_d. For the tempi 90 BPM, 120 BPM and 150 BPM there is no observable narrow peak. On the contrary, for some tempi (90 BPM) there are more than one peak and different values of L_d present. It might suggest that different musical instruments have dissimilar latency values L_d, where the performance is interrupted. The lower peaks at 300ms (right side of Figure 5) are produced by those musicians who were able to play even at the higher latency values (L_d) up to 300ms (e.g. trumpet in B). Those values of L_d can be assumed as outliers.

3.1 Latency Tolerance Range

The measure of the latency tolerance range (LTR) is the estimation of a range in milliseconds where latency is tolerable according to the different musical instrument groups. The (LTR) can be defined as the difference between the third and first quartile of the gathered data and is better suited than the mean or median to describe non-normal distributions.

The latency tolerance range (LTR) is measured in milliseconds and its mathematical expression is:

$$LTR = Q3Ld - Q1Ld \tag{2}$$

Figure 6 displays the LTR for the different musical instrument groups based on the gathered data. As expected and based on the information of Figure 4, the LTR for membranophones is very narrow compared to the aerophones group. At this point, it is important to clarify that only few instruments (snare drum and timpani) are part of the membranophone group. On the contrary, the group of the aerophones involved more than six different instruments. The groups of chordophones and idiophones are similar in range but as the number of instruments differ, comparisons may be difficult.

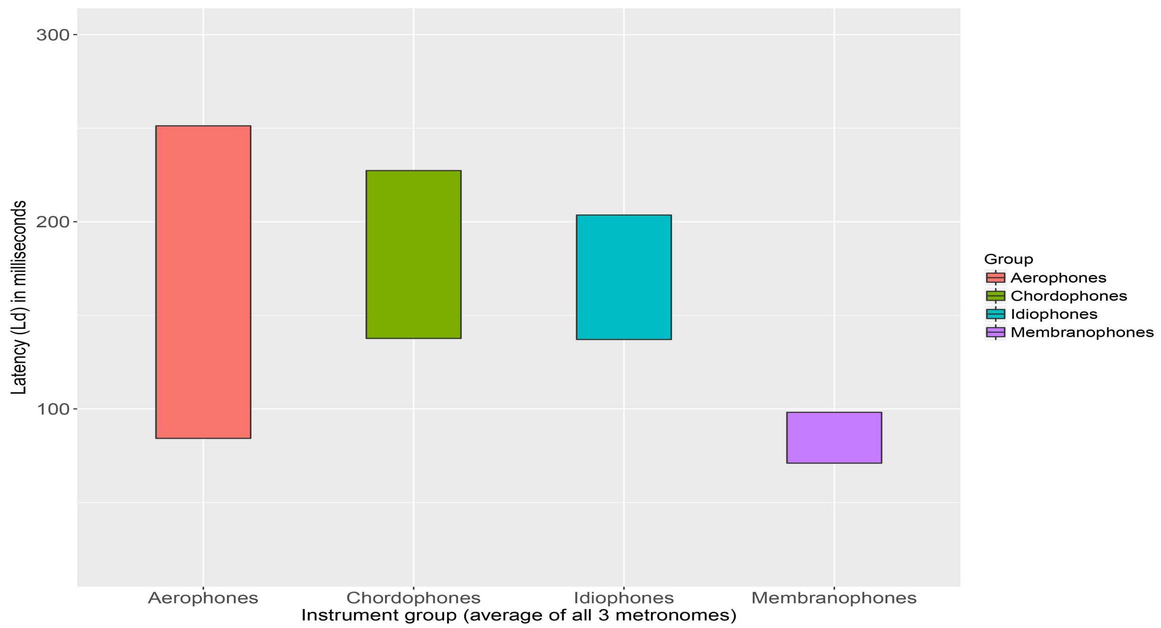

Fig. 6. Latency tolerance range (LTR) for musical instrument groups

The latency tolerance range is dependant on the amount of data gathered. However, the tendency is clear. Playing a membranophone may be a difficult task regarding latency. Aerophones, chordophones and idiophones may perform music as an ensemble within similar latency ranges, without any inconvenience.

4 Discussion and Conclusions

The present paper presented the exploratory data analysis of a controlled listening test. The research design and test characteristics such as the use of a metronome and a score enabled comparisons of the results. Based only on the exploratory data analysis of the gathered data, it is possible to claim that the ability to cope with latency is affected by the musical instrument played. Performing on different musical instruments may have an effect on the ability to cope with latency. Furthermore, the years of experience playing an instrument, the hours of practice and even the use of a metronome while practicing may not affect this ability at all.

The unbalanced design regarding the number of musical instruments tested per instruments type and the lower number of test subjects or musicians might diminish the external validity of the outcomes. However, there are clear tendencies and patterns and the effect is far from random.

The influence of the musical instrument, regarding the latency issue, decreases while performing at faster tempi above 180 BPM. On the contrary, this influence increases by playing a musical instrument at lower tempi under 150 BPM. In addition, the latency tolerance range (LTR) enable the comparison of results regarding the issue of latency between the different musical instrument groups.

The results and the methodology presented, may constitute a further step in the modelling of musical instruments for technologies such as virtual reality. Moreover, based on the results presented, software requirements for artistic and educational projects, such as network musical live performances and distance learning programs can be further developed to fit better the requirements regarding musical instrument groups.

References

1. Barbosa, Alvaro.: Displaced Soundscapes: Computer-Supported Cooperative Work for Musik Applications. Phd thesis, Pompeu Fabra University, Barcelona and Spain, April 2006
2. Bortz, Jürgen., Lienert Gustav A., Boehnke, Klaus.: Verteilungsfreie Methoden in der Biostatistik. Springer-Lehrbuch Bachelor, Master. Springer, Berlin and Heidelberg, 3rd edition, 2008
3. Carôt, Alexander., Werner Christian.: Network music performance - problems, approaches and perspectives. In Music in the Global Village Conference, Budapest, Hungary (2007)
4. Chafe, Chris., Gurevich, Michael.: Network time delay and ensemble accuracy: Effects of latency, asymmetry. AES Audio Engineering Society Convention Paper, In 117th San Francisco California, October 28-31, pages 1-7, 2004
5. Field, P. Andy., Miles, Jeremy., Field, Zoë.: Discovering statistics using R. SAGE, London (2012)
6. Medina Victoria, Jorge.: An approach to quantifying the latency tolerance range in non-collaborative musical performances. AES Audio Engineering Society Convention Paper, In 136th, Berlin, Germany, pages 1-6, 2014

7. Rottondi, C., Chafe, C., Allochio, C., Sarti, A.: An overview on networked music performance technologies. IEEE Access(4), 8823-8843 (2016)
8. Sachs, Kurt.: The history of musical instruments. Dover Publications, Mineola, N.Y. (2006)
9. Schnupp, J., Nelken, I., King, A.: Auditory Neuroscience - Making Sense of Sound. The MIT Press paperback edition, 2012.