Analysis of trumpet performance adjustments due to room acoustics

Sebastià V. AMENGUAL GARI[1]; Malte KOB[1]; Tapio LOKKI[2]

[1]Detmold University of Music, Germany

[2]Aalto University, Finland

ABSTRACT

In a musical performance, musician, instrument and room form a closed feedback loop that continuously shapes the generated sound. A virtual acoustic environment was developed to study performance adjustments systematically. In formal experiments, 11 trumpet players were recorded while performing several pieces with multiple auralized versions of real spaces. A set of 44 low level audio features was computed for each of the 364 final recordings. A Dual Multiple Factor Analysis (DMFA) was performed, reducing the data dimensionality to four dimensions related to musical aspects: Level/Timbre, Dynamics, Overall Tempo and Tempo Variations – accounting for 58% of the explained variance. Correlation analysis revealed a moderate effect of temporal (EDT, RT60) and energy (G) room acoustic parameters on performance level and timbre. Additionally, in an environment with considerably stronger energy, players tended to decrease the overall tempo of the performance. However, most of these effects were highly individual, and a further analysis was performed to assess the behavior of each player. To investigate the perceptual impact of performance adjustments on listeners, an online listening test was conducted. The results revealed that listeners are able to consistently perceive level and timbre variations induced by stage acoustic conditions. Results further suggest that the perception of dynamics or tempo variations cannot be fully described by single values and a multi-dimensional characterization is necessary.

Keywords: Musical performance, stage acoustics, real-time auralization, musician adaptation, room acoustics.

1. INTRODUCTION

In a musical performance, musician, instrument, and room can be conceptually regarded as a closed feedback loop that continuously shapes the generated sound. The mechanical actions of a musician on their instrument result in the generation of sound, which is then modified by the room and perceived again by the musician. As a result of a cognitive process and in response to the perceived sound, the musician then adjusts their actions. In addition to that, other environmental factors, such as lighting, temperature or the interaction of the musician and the audience, can potentially impact the performance as well. A diagram depicting this conceptual model is shown in Fig. 1. In this model, elements directly related to acoustics are represented in black, while environmental factors are represented in grey. Other conceptual performance models were previously introduced in (1, 2).

Past research suggests that room acoustics systematically influences certain musical aspects, while other adjustments seem to be highly individual and variable. Bolzinger, Kawai *et al.* (3–5) conducted experiments in a room with variable acoustics, effectively isolating acoustics from environmental factors. Pianists were recorded using a MIDI interface and the results suggest that they tend to reduce the overall level of the performance and the use of sustain pedal in more reverberant conditions. However, the same studies found conflicting results in terms of overall performance tempo, suggesting that adjustments are highly individual and might depend on context or musical piece.

Auralization has been used to study the influence of room acoustics to the performance of musicians. Ueno, Kato *et al.* (2, 6) recorded 5 musicians (oboe, two flutes, baritone singer and violin) playing the same two pieces in a virtual acoustic environment, focusing on a detailed analysis of individual performers. Schärer *et al.* (1) conducted experiments with 12 musicians playing in a binaurally rendered virtual acoustic environment. Both these studies concluded that musicians tended to

[1]amengual@hfm-detmold.de, kob@hfm-detmold.de
[2]tapio.lokki@aalto.fi

Peer reviewed

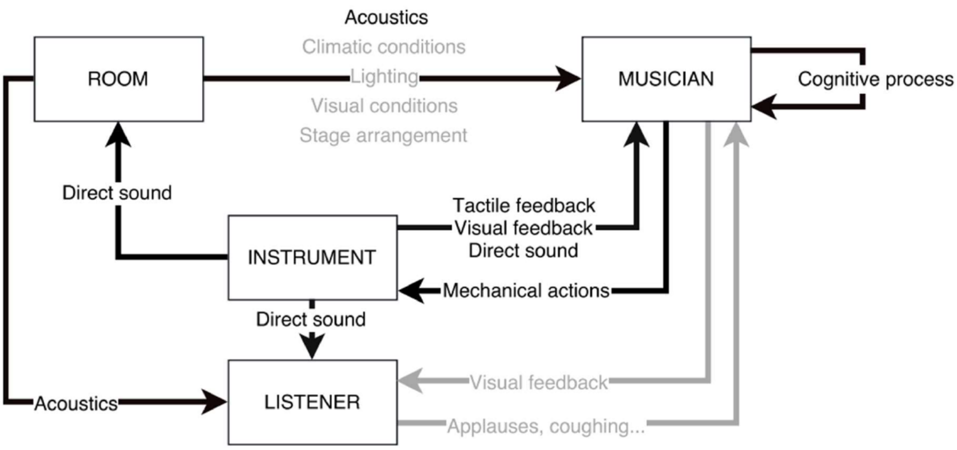

Figure 1 – Conceptual representation of the factors involved in a musical performance [extracted from (7)]

decrease the overall performance tempo in both very short and very long reverberation times. However, these adjustments did only generalize for certain pieces, and it was found that other adjustments related to other aspects such as timbre aspects, vibrato or dynamics (among others), were found to be highly individual.

The discussed studies investigating solo musicians were conducted with a limited number of instruments and a set of constraints in the recorded pieces – either explicitly selecting the pieces or stating a specific musical character. This paper presents an experimental investigation aiming at characterizing individual and generalized performance adjustments that solo trumpet players implement in order to accommodate room acoustics. In a laboratory setup, eleven trumpet players were recorded when performing under 4 different stage acoustic conditions auralized in real-time using loudspeaker-based spatial sound reproduction. In addition, personal interviews were conducted with each musician after the recording sessions. Finally, listener judgments were completed to evaluate the perceived extent of the performance adjustments.

2. Simulated Stage Acoustics

The first electroacoustic systems aiming at reproducing room or stage acoustics in real-time appeared towards the end of the 1970s. They were based on the synthesis of a very limited number of early reflections using delay lines, with the addition of late reverberation using either plate reverbs (8) or the sound reproduced and captured in a reverberation chamber (9). The limited fidelity of these early systems was overcome by the increase in computing power and availability of digital signal processing (DSP), allowing the realization of systems based on real-time convolution of measured or simulated room impulse responses and reproduced either using loudspeaker arrays (10, 11) or binaural technology (1).

2.1 Auralization

In the present study, the auralized stage acoustics were based on measured multichannel room impulse responses (MRIR) captured on different stages. The measurement set-up was composed of a directional loudspeaker (Neumann K120A) with radiation properties similar to the trumpet (12) and an open microphone array arranged on a stage, imitating a trumpet musician standing on stage 13. The loudspeaker was placed at the position of the trumpet bell, oriented towards the audience area, and the microphone was placed at the position corresponding to the musician's head (approximately 60 cm behind the loudspeaker). The measured MRIR were then analyzed using the Spatial Decomposition Method (SDM) (14) and re-synthesized for reproduction in the Detmold Surround Sound Sphere (D3S), a 13 loudspeaker set-up (see Fig. 2). Note that the direct sound must be removed from the resulting RIRs, as only the room reflections must be reproduced.

For real-time reproduction, a directional microphone was attached to the trumpet bell and the sound produced by the trumpet was convolved in real-time with the rendered RIRs, providing the musician with the acoustic feedback of a selected room. The convolution was performed using a custom Max/MSP patch and the total system latency was 5.9 ms, allowing for the reproduction of

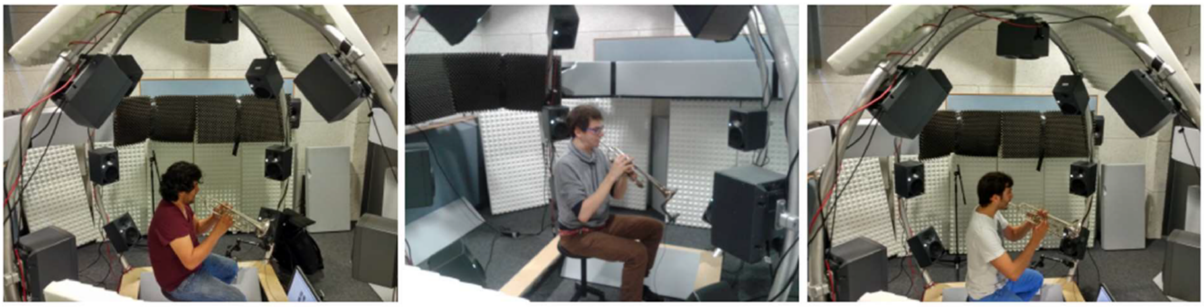

Figure 2 – Musicians performing in the D3S during the experiment [extracted from (7)].

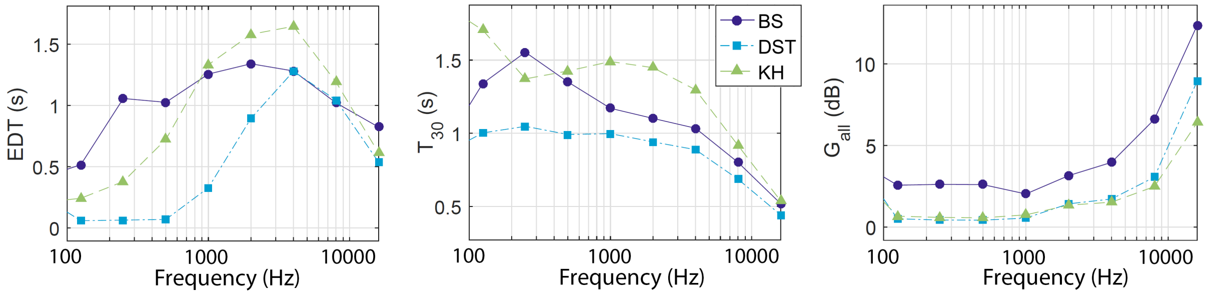

Figure 3 – Room acoustic parameters of the measured spaces.

the floor reflection. An extensive description of the system, including measurement details, implementation of the convolution engine, equalization, calibration and validation is presented in (7, 13).

2.2 Acoustic parameters

Three rooms from the Detmold University of Music were measured and used for auralization – the Brahmssaal (BS), a small shoebox music chamber room, the Detmold Sommertheater (DST), a shoebox shaped room with side balconies with approximately 300 seats; and the Detmold Konzerthaus (KH), a 600 seat concert hall. In addition, the quasi-anechoic acoustics of the laboratory were used as an experimental condition too. The Reverberation Time (T_{30}) of the laboratory is less than 100 ms at mid and high frequencies, and given that special care was taken to remove any prominent reflection by adding extra absorption, it is considered anechoic in the analysis.

Several acoustic parameters (EDT, T_{20}, T_{30}, C_{80}, G_{all}, G_{early}, G_{late}, ST_{early}, ST_{late}, T_S) were computed from the measured RIRs. Due to strong correlation between the computed parameters only three of them are considered for further analysis: Reverberation Time (T_{30}), Early Decay Time (EDT) and Room Strength relative to the direct sound (G_{all}). The parameters, expressed in octave bands, are shown in Fig. 3. These parameters were selected due to the extended use of T_{30} in concert hall research, the relation of EDT with perceived reverberance, and the fact that G_{all} can be computed with less uncertainty than ST parameters (15).

In terms of EDT, DST has the shorter time, with a shape similar to KH, with considerably longer times at higher frequencies, due to the strong source directionality. The fact that the stage dimensions of BS are much smaller, results in a flatter EDT, and the highest at low frequencies. In terms of T_{30}, all the rooms follow a similar trend, with DST having the lowest reverberation time, followed by BS and KH. Regarding G_{all}, all three rooms follow a very similar trend, with a much higher gain at high frequencies, due to source directivity. Both DST and KH present similar G_{all} values, while BS is considerably more energetic.

It is known that the SDM-based auralizations suffer from an increase of reverberation at high frequencies. Strategies to perform time-frequency equalization are available, although they were not yet developed at the time of the experiments (16, 17). Thus, the T_{30} of the auralized environments is approximately 30% and 50% higher than the original T_{30} at 8 and 16 kHz octave bands, respectively.

3. Experiment

Eleven participants (10 male, 1 female) were recruited to complete the study. All of them were either bachelor or master students of the Detmold University of Music at the time of completion of the study. One of the players specialized on jazz performance and the rest were classical musicians.

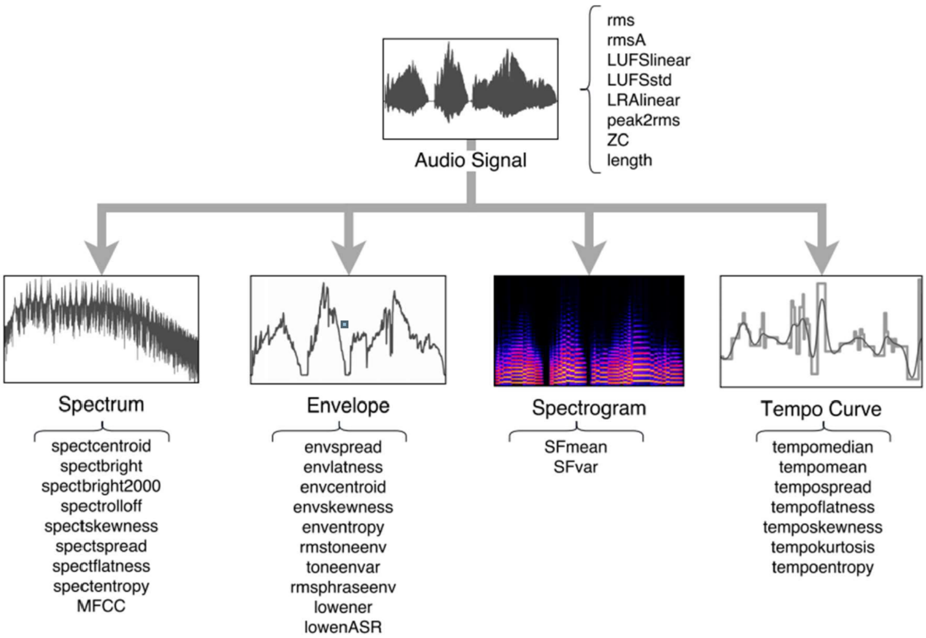

Figure 4 – Low level features extracted from the recordings [extracted from (7)].

Before the study, musicians were instructed to select between two and four musical excerpts of their choice, with the only limiting factor related to the duration – between 20 seconds and 1 minute – and their ability to perform the piece without difficulties – in order to reduce learning effects. Upon arrival, participants were introduced to the concept of the experiment, explaining that the acoustics of other spaces were being rendered in real-time as a response to their generated sound. The directional microphone was fitted to the trumpet and a calibration test was completed. Finally, the played pieces were discussed and recordings started.

The number of trials per piece was not determined at the beginning of the session, and musicians were allowed to switch to another piece to avoid fatigue. Each trial was done under a randomly selected acoustic condition, and prior to each recording, unlimited time was provided for the musicians to become familiar with the acoustics. This procedure was implemented to reduce the experimental constraints on the musicians and increase the ecological conditions of the study.

After the recording session was over, a survey was provided to collect demographic data and feedback, and an interview was conducted. The interview was based on a set of closed questions regarding the adjustment of their performance and cooperative conversation. Musicians were encouraged to attend to several recording sessions on different days.

4. Objective Analysis of Recorded Signals

4.1 Dataset

Approximately 400 recordings were collected in four different recording sessions. After removing recordings with audible errors, 364 recordings were left for the consequent analysis. The data have a nested structure with 3 levels: player, session, and piece, resulting in 46 sub-datasets (unique player, session and piece combinations). Some players participated in multiple sessions and some pieces were performed by more than one musician, allowing the exploration of the effect of external factors on the same musicians, or the comparison of multiple players performing the same piece.

4.2 Low level features

A set of 44 low level features was extracted from the recordings, either directly from the audio signal, or from transformations of it. A description of each of the features and its extraction method is included in (7, Chapter 3.2.2) and is out of the scope of this paper. Most of the features are shown in the diagram depicted in Fig. 4, which also shows the high level representation of the extraction process.

Table 1 – Contributions of each low level feature to the DMFA dimensions.

Dimension	Contributions (in descending order)
1 – Level & Timbre	Spectral centroid (0.93), energy ratio above 2000 Hz (0.93), Mel-Cepstral Coefficient 1 (-0.93), LUFS (0.92), A-weighted RMS (0.91), brightness (0.91), spectral rolloff (0.90), spectral skewness (-0.90), RMS (0.89), RMS of the tone envelopment (0.89), standard deviation (0.88) of the spectral flux, mean of the spectral flux (0.85), spectral entropy (0.82), RMS of the phrase envelopment (0.80), Mel-Cepstral Coefficient 2 (-0.67), spectral flatness (0.60).
2 – Dynamics	average to silence ratio (0.83), envelope entropy (-0.82), LUFS standard deviation (0.81), loudness range (0.74), low energy rate (0.69).
3 – Overall tempo	tempo mean (-0.87), length (0.72), temporal spread (-0.64), envelope centroid (0.64), envelope spread (0.62), median tempo (-0.62)
4 – Tempo variations	temporal entropy (-0.78), temporal flatness (-0.78), temporal spread (0.71).

4.3 Dimensionality reduction

In order to reduce the dimensionality of the data and discard features that are strongly correlated, a Dual Multiple Factor Analysis (DMFA) was performed on the dataset (18). This technique was chosen due to the nature of the dataset, by several observations (audio recordings) measured on the same set of variables (low level audio features) and organized in groups (player, session, piece).

The first 10 dimensions account for approximately 80% of the variance in the data, and the first 4 dimensions (58% of explained variance) were kept for further analysis, given their strong correlation with a subset of parameters. It is important to note that the values of the low level features are centered within each of the groups during the DMFA, thus the values of each dimension represent the relative differences between performances of the same group (player, session, piece). This is done to ensure that the dimensions capture the low level differences caused by performance variations rather than piece differences – given that each piece yields vastly different values for each feature.

Given the contributions of each low level feature, the dimensions are labelled as follows:

- Dimension 1 – Level/Timbre – 30.92% explained variance
- Dimension 2 – Dynamics – 11.57% explained variance
- Dimension 3 – Overall tempo – 8.61% explained variance
- Dimension 4 – Tempo variations. – 6.75% explained variance

Table 1 collects the individual contributions of each low level feature. Note that the highest contributor of Dimension 3 – mean tempo extracted from the tempo curve – has a negative value, thus this dimension correlates negatively with the overall tempo - higher values indicate a slower performance.

5. Results

From an analysis perspective, each of the performance groups (unique combinations of session, player and piece) are considered as independent sub-dataset with multiple observations (recordings under different acoustic conditions). Thus, the values of each observation are averaged in each of the groups for each player, and a correlation analysis is performed between these and the room acoustic parameters. The single values of the room acoustic parameters are computed by averaging between 250 and 4000 Hz. The results are presented in Fig. 5 for each of the players individually.

5.1 Level & Timbre – Dimension 1

In more energetic and reverberant environments, trumpet players tend to decrease the overall level and brightness of their performance. Six out of eleven players exhibit significant correlations with at least one of the analyzed parameters, while none of the players significantly increases the level and brightness. Only two players seem to be unaffected by acoustics, with regards to level and timbre adjustments.

5.2 Dynamics – Dimension 2

In terms of adjustments of dynamics, the correlation analysis suggests that there are no clear generalized trends. However, some players do present significant correlations, suggesting that musicians actively modify the dynamic variations of their performance as a function of the acoustic properties

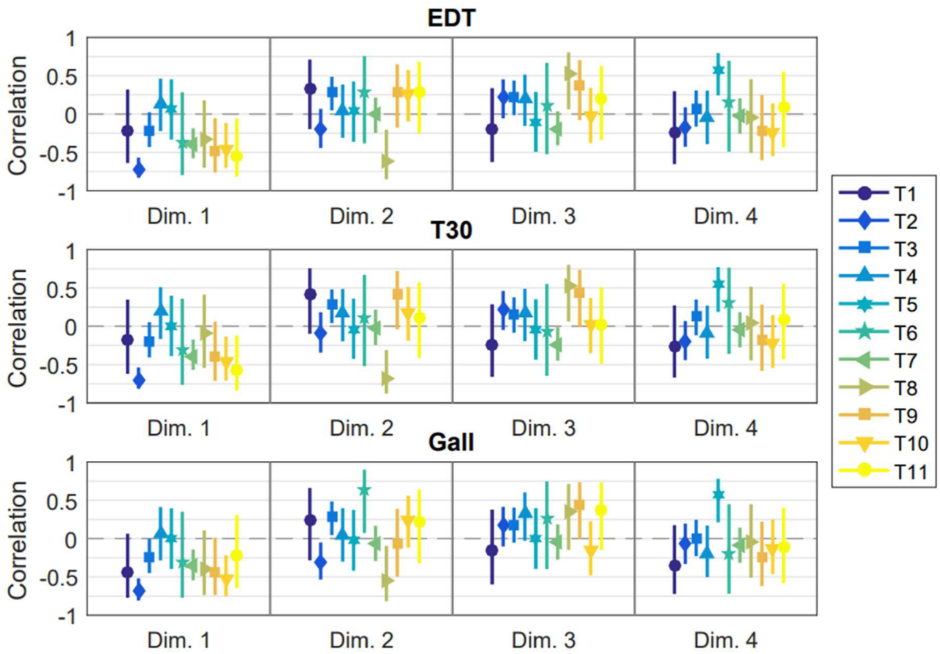

Figure 5 – Correlation between the extracted dimensions (Dim.1 – Level/Timbre, Dim. 2 – Dynamics, Dim. 3 – Overall tempo, Dim.4 – Tempo variations) and room acoustic parameters. Vertical bars represent 95% confidence intervals [extracted from (7)].

of the room. Players T3 and T6 seem to significantly increase the extent of dynamic variations in more energetic conditions, while players T2 and T8 reduce them. The adjustments of players T3 and T8 seem to significantly correlate with EDT and T_{30} as well.

5.3 Overall tempo – Dimension 3

Although the overall tempo does not appear to be significantly adjusted in a generalized manner, an important portion of the players seem to slightly reduce it in more energetic environments. Players T8 and T9 present a significant decrease of overall tempo in rooms with a longer T_{30}, while player T7 seems to slightly increase it as a function of T_{30}. Note that player T7 is the only jazz musician, and it is possible that the adopted strategy to adjust their performance differs from classical musicians.

5.4 Tempo variations – Dimension 4

Finally, only player T5 presents a significant adjustment of tempo variations, presenting an increase of tempo variations as a function of EDT, T_{30} and G_{all}. From a musical perspective, the adjustment of tempo variations might be strongly influenced by the character of the piece, e.g., while a rhythmic piece might not allow for extensive adjustments, a cadenza provides much more freedom in the interpretation.

6. Listener Impressions

In an attempt to investigate the perceptual impact of performance adjustments on listeners, we conducted an online test evaluating several recordings from the analyzed data set. In addition, the test aims at evaluating the perceptual validity of the DMFA dimensions and their correlation with listener subjective judgments.

Four recordings of three pieces were selected and compared in a quantitative scale rating on 7 axes: overall loudness, overall tempo, articulation, dynamic variations, sound color, tempo variability, and expressiveness. In each of the trials, one of the pieces was used as a reference and the others were rated relative to the reference in a bipolar continuous scale. All the pieces (including the reference) could be heard and compared against each other without time limitation in each trial. A total of 24 listeners participated in the test.

The results of the listening test suggest that listeners are generally able to perceive performance changes induced by room acoustics (see Tab. 2). Both loudness and sound color judgments are significantly correlated with the values of Dimension 1. Moderate correlations are found between per-

Table 2 – Correlations between subjective listener judgments of musical aspects and DMFA dimensions. Bold numbers refer to Holm-Bonferroni corrected correlation coefficients with p<0.01.

Player/Piece	Dim. 1 Loudness	Dim. 1 Timbre	Dim. 2 Dynamics	Dim. 3 Tempo	Dim. 4 Tempo variations
T2/Hoehne1	**0.59**	**0.44**	0.13	-0.02	0.15
T5/Mussorgsky	0.49	0.43	0.61	-0.52	0.35
T9/Charlier34	**0.72**	**0.61**	0.46	-0.042	0.42

ceived dynamics and Dimension 2, although they are not statistically significant. For both temporal dimensions (3 and 4), tempo and tempo variations present weak to moderate non statistically significant correlations. However, the signs of the correlations match with the expected behavior (negative for Dim. 3 and positive for Dim. 4). Given that the changes in Dim. 1 are generally much larger than on the rest, the amount of variability in Dimensions 3 and 4 might not be enough to investigate underlying correlations. Further tests with other pieces with larger variations on those dimensions are needed to confirm whether they correlate with perceptual judgments.

7. Discussion

The results suggest that musicians actively adjust their performance to suit room acoustics and the correlation analysis suggests that the adaption strategies of trumpet players are partially in line with the results presented in previous studies (1, 2, 4, 6). The results presented here agree with previous studies in the fact that generalized trends are present regarding the reduction of the overall level in more energetic environments, and a semi-generalized reduction of the tempo in environments with longer reverberation times. However, the fact that we examined only linear correlations does not allow for the comparison with studies that found a decrease in tempo for very short reverberation times. In addition, and contrary to past experiments, in this experiment we found that the timbre properties are strongly related to the performance level, thus following the same trend. This was not observed in other studies, and it might be partially explained by the acoustic properties of the trumpet, which naturally renders a brighter sound at higher levels due to non-linear sound propagation (19). Thus, musicians cannot isolate one feature from the other.

It is not clear to which extent those adjustments are conscious and deliberate i.e. musicians determine what adjustments are necessary in order to match their preconception of the performance, or they are the result of an underlying more instinctive approach i.e. musicians have a target concept of the performance, but do not actively consider what are the adjustments that they do to achieve it. This question is motivated by the interviews conducted after the recordings, which confirm that musicians are aware of consciously adjusting the performance, but in some cases the reported adjustment strategy does not entirely match with the results obtained in the objective analysis (see 7 for a more comprehensive discussion).

The results presented in this paper allow the comparison of adaption strategies of individual players versus generalized trends. However, in the study, each player recorded multiple pieces with contrasting musical character. An in depth analysis on the relationships between musical character of each piece and the extent of the adjustments could provide a greater insight in terms of evaluating the adaption strategy of each individual player. This would contribute to determining whether differing strategies or degrees of adjustment are influenced by the musical character or by underlying differences between musicians. However, the classification of pieces by musical character is not straightforward, given that the same excerpt could be a combination of several phrases with contrasting character. In addition to that, while techniques for genre or music mood classification are available, to the knowledge of the authors, there are no reliable automatic approaches to labelling the musical features of a classical or jazz piece.

The implementation of the virtual acoustic environment has proven to be efficient for use within the investigations presented here, but more extensive studies could be performed to investigate the effect of variations in the sound reproduction such as the direct-to-reverberant sound level ratio. Also a correlation between acceptance of the virtual environment and the amount of musical adjustments could be a topic of further investigations.

Finally, the use of more complex room acoustic parameters and the addition of more rooms could contribute to establishing more subtle relationships. For instance, the use of spatial parameters accounting for the distribution of energy could be considered. In addition, the determination of single values for the studied parameters (EDT, T_{30}, G_{all}) could be refined by weighting them with using the spectra of the recordings, instead of averaging over an extended frequency range.

8. Conclusions

A study investigating the effects of room acoustic conditions on live music performance adjustments of solo trumpet players is presented. To isolate the acoustic conditions from environmental factors, the experiments were conducted in a virtual acoustic environment. Results from the experiment were partially in line with previous research, suggesting that musicians generally tend to decrease their playing level and render a darker timbre in more reverberant and energetic environments. In addition, a subset of players tended to decrease their overall tempo in increased reverberation. Adjustments of dynamics and tempo variations were observed as well, although they are more individualized and do not follow general trends. A listening test showed that listeners are able to perceive differences between recordings done under differing acoustic conditions, and their subjective judgments correlate well with the objective signal analysis. Inconclusive results are found for the perceptual judgments of dynamic and tempo variations, and further tests are required to investigate whether the dimensions used to characterize these musical aspects correlate well with perceptual judgments of these musical dimensions.

ACKNOWLEDGEMENTS

The authors wish to thank the musicians and listeners participating in the experiments. This research work was funded by the European Commission within the ITN Marie Curie Action project BATWOMAN under the 7th Framework Programme (EC grant agreement no. 605867).

REFERENCES

1. Schärer-Kalkandjiev, Z. The Influence of Room Acoustics on Solo Music Performances. An Empirical Investigation. PhD thesis. TU Berlin, 2015:227.
2. Ueno, K, Kato, K, and Kawai, K. Effect of room acoustics on musicians' performance. part I: Experimental investigation with a conceptual model. Acta Acustica united with Acustica 2010;96:505–515.
3. Bolzinger, S and Risset, JC. A Preliminary Study on the Influence of Room Acoustics on Piano Performance. Journal de Physique IV 1992;2:93–96.
4. Bolzinger, S, Warusfel, O, and Kahle, E. A study of the influence of room acoustics on piano performance. Journal de Physique IV 1994;4:617–620.
5. Kawai, K, Harada, K, Ueno, K, Sakuma, T, and Kato, K. Experiment on the Adjustment of Piano Performance To Suit Room Acoustics: Analysis of Performance Coded Into Midi Data. In: *Proceedings of the International Symposium on Room Acoustics*. Toronto, 2013.
6. Kato, K, Ueno, K, and Kawai, K. Effect of Room Acoustics on Musicians' Performance. Part II : Acoustic Analysis of the Variations in Performed Sound Signals. Acta Acustica united with Acustica 2015;101:743–759.
7. Amengual Garí, SV. Investigations on the Influence of Acoustics on Live Music Performance using Virtual Acoustic Methods. PhD thesis. Detmold University of Music, 2017:180.
8. Marshall, AH and Meyer, J. The directivity and auditory impressions of singers. Acta Acustica united with Acustica 1985;58:130–140.
9. Gade, AC. Investigations of Musicians' Room Acoustic Conditions in Concert Halls. II: Field Experiments and Synthesis of Results. Acustica 1989;69:249–262.
10. Laird, I, Murphy, DT, Chapman, P, and Seb, J. Development of a Virtual Performance Studio with application of Virtual Acoustic Recording Methods. In: *130th Convention of the Audio Engineering Society*. New York, 2011:12.
11. Brereton, JS. Singing in Space(s): Singing performance in real and virtual acoustic environments – Singers' evaluation, performance analysis and listeners' perception . PhD thesis. 2014:312.
12. Amengual Garí, SV, Lokki, T, and Kob, M. Live performance adjustments of solo trumpet players due to acoustics. In: *Proceedings of the International Symposium on Music and Room Acoustics 2016*. La Plata, 2016.

13. Amengual Garí, SV, Eddy, D, Kob, M, and Lokki, T. Real-time auralization of room acoustics for the study of live music performance. In: *Fortschritte der Akustik – DAGA 2016*. Aachen, 2016.

14. Tervo, S, Pätynen, J, Kuusinen, A, and Lokki, T. Spatial decomposition method for room impulse responses. Journal of the Audio Engineering Society 2013;61:17–28.

15. Dammerud, JJ. Stage Acoustics for Symphony Orchestras in Concert Halls. PhD thesis. University of Bath, 2009:210.

16. Tervo, S, Pätynen, J, Kaplanis, N, Lydolf, M, Bech, S, and Lokki, T. Spatial Analysis and Synthesis of Car Audio System and Car-Cabin Acoustics with a Compact Microphone Array. Journal of the Audio Engineering Society 2015;63:914–925.

17. Amengual Garí, SV, Brimijoin, WO, Hassager, HG, and Robinson, PW. Flexible binaural resynthesis of room impulse responses for augmented reality research. In: *EAA Spatial Audio Signal Processing Symposium*. Paris, 2019.

18. Abdi, H, Williams, LJ, and Valentin, D. Multiple factor analysis: Principal component analysis for multitable and multiblock data sets. Wiley Interdisciplinary Reviews: Computational Statistics 2013;5:149–179.

19. Rendón, PL, Ezeta, R, and Pérez-López, A. Nonlinear Sound Propagation in Trumpets. Acta Acustica united with Acustica 2013;99:607–614.