



HAL
open science

Singing in physical and virtual environments: how performers adapt to room acoustical conditions

Paul Luizard, Erik Brauer, Stefan Weinzierl

► **To cite this version:**

Paul Luizard, Erik Brauer, Stefan Weinzierl. Singing in physical and virtual environments: how performers adapt to room acoustical conditions. AES Conference on Immersive and Interactive Audio, Mar 2019, York, (Royaume-Uni), United Kingdom. <hal-04910565>

HAL Id: hal-04910565

<https://hal.science/hal-04910565v1>

Submitted on 31 Jan 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization



Audio Engineering Society Conference Paper

Presented at the Conference on
Immersive and Interactive Audio
2019 March 27 – 29, York, UK

This paper was peer-reviewed as a complete manuscript for presentation at this conference. This paper is available in the AES E-Library (<http://www.aes.org/e-lib>) all rights reserved. Reproduction of this paper, or any portion thereof, is not permitted without direct permission from the Journal of the Audio Engineering Society.

Singing in physical and virtual environments: how performers adapt to room acoustical conditions

Paul Luizard¹, Erik Brauer¹, and Stefan Weinzierl¹

¹FG Audiokommunikation, TU-Berlin

Correspondence should be addressed to Paul Luizard (paul.luizard@campus.tu-berlin.de)

ABSTRACT

This study has investigated to what extent and how trained singers adapt to the room acoustic conditions in physical and in virtual acoustic environments. In a first phase, two musical pieces were recorded for four singers in eight different performance venues by means of a near-field microphone. Binaural room impulse responses were recorded by mean of a dummy head together with a directional loudspeaker placed together on a rotating plate, at the position of the singer on stage. In a second phase, the first experiment was replicated in the anechoic chamber, using an interactive auralization provided by dynamic binaural synthesis of the same performances venues, i.e., the singers could hear their own voice in the virtual room while singing the same pieces using real-time convolution and head tracking. The 128 recordings made in the physical and virtual rooms were analyzed in terms of audio features related to tempo, loudness, and timbre of the musical performance. The impact of the room acoustical conditions, as described by established room acoustical parameters, on these performances was analysed by linear regression. The results show rather individual patterns of adaptation, depending on the musical piece and the musical performer. The same interactions between room acoustics and musical performance, however, both in direction and in strength, could be observed in the physical and virtual environment. This seems to indicate the suitability of virtual acoustic environments to investigate musical tasks, not only for music listeners, but also for musical performers and the interaction with their acoustic environment.

1 Introduction

A comparison of physical and virtual acoustic environments has mostly been carried out in terms of task performance such as spatial navigation [1], sound source localization [2], or estimation of egocentric distance [3]. When comparing different qualities of auralization, it was obvious that participants experienced a higher sense of presence and showed better performance for navigation and memory tasks by using higher quality binaural sound rendering [4].

Although virtual acoustic environments are becoming an important tool for musical practice, e.g. for enabling networked music performances [5], they have not yet been considered as a "task" to study the impact of the physical vs. virtual environment on musical performance. Most of the previous research has just *used* virtual acoustic environments, as generated, for example, by multichannel loudspeaker systems, as a technique to investigate the influence of different acoustical conditions on music performances under laboratory conditions [6]. As a result, differences in various perfor-

mance features such as tempo, vibrato, and sound level were observed under the different acoustical conditions. A comparison of these results with the response of the musicians given in questionnaires after the recording indicated that many but not all of the adaptations in the way of playing were made consciously, such as a more pronounced staccato and longer pauses between notes with increasing reverberation time [7].

Another study [8] was carried out in seven different performance venues in which one of the most renowned contemporary violoncello soloists played the same musical program during a concert tour. Room acoustical measurements were performed in the empty rooms, and computer simulations were conducted to estimate the acoustics in the occupied condition. The musical recordings were analyzed in terms of tempo, dynamics, loudness, and timbre characteristics. A statistical analysis of the relationship between room acoustical parameters according to ISO-3382-1 [9] and audio features representing different qualities of the musical performance revealed that more than 50% of the variance in performance features could be explained by the acoustical properties of the environment. With a similar test design, the same authors invited 12 musical performers playing six musical instruments to perform the same pieces of music in different acoustical conditions, this time simulated by dynamic binaural synthesis, thus avoiding the confusion with other influencing factors such as the visual impression, audience reactions, or the architectural design of the space [10]. The analysis again revealed a significant influence of room acoustical parameters on the way of playing. In the simulated environment, however, only 8% of the variance in performance features could be explained by the acoustical properties of the room, if considering the behaviour of the individual musicians. Moreover, quite individual patterns of reaction could be observed, which varied not only between different musical instruments but also between different performers on the same instrument.

The interaction of singers with their room acoustical environment was investigated mainly in the context of choir performance. Marshall and Meyer placed an ensemble of singers in a semi-anechoic room, where they were exposed to synthetic sound fields including discrete sound reflections, combined with an artificial reverberation tail [11]. The experiment demonstrated the strong influence of reverberation on the comfort of singers. Early reflections were appreciated if their time delay did not exceed 40 ms, beyond which they

became unpleasant. A more recent work [12] focused on the effect of early reflections when singing in small ensembles (two to four singers) in a church. Artificial early reflections were added to the natural acoustics of the venue using microphones and loudspeakers. The different groups of singers performed musical pieces with various tempi and as a result, it appeared that singing in fast musical tempi benefits from strong early reflections whereas they do not significantly impact slow-tempo singing.

Previous investigations have thus been mainly focused on the interaction between performance characteristics and room acoustical conditions, using virtual environments as a tool for the experimental control of these conditions. Only Schaerer Kalkandjiev and Weinzierl have conducted such a study in physical *and* virtual environments, finding a much stronger adaptation in the physical condition. In the two studies, however, different musicians, different musical instruments, different rooms, and different musical pieces were considered, and since the reaction patterns turned out to be quite individual, it was difficult to rule out that the effect was due to other factors.

Hence, in the current study we have conducted a parallel investigation of the influence of room acoustical conditions on musical performance, both in physical environments and in the virtual re-synthesis of these environments. For this purpose, we have invited four solo singers to perform an identical repertoire in a wide range of eight different room acoustical conditions. The measurement of binaural room impulse responses at the position of the singer in each room allowed for running dynamic binaural auralization of these rooms in order to replicate the experiment with the same singers and musical pieces under laboratory conditions in the anechoic chamber.

The results should thus not only indicate the impact of the performance space (as described by room acoustical parameters) on the performance of solo singers (as described by audio features), but also the effect of the physical vs. virtual environment on the "task" of adapting the musical performance to the acoustical conditions. The results could show to what extent the performance of musicians in this respect is affected by the virtual environments, and to what extent immersive, interactive auditory environments can be used to study the natural behaviour of musicians in space.

2 Measurements

2.1 Physical rooms

The extent of adaptation of trained singers was estimated after recording the same musical program in six rooms covering a broad range of acoustical characteristics, as presented in Tab. 1. A system of heavy banners could be deployed in two of these rooms, leading to a considerable change of most acoustical parameters beyond their just noticeable difference (JND) [13, 14], and resulting in eight different acoustical conditions for the singers. The measurements were performed by means of a dodecahedron speaker (Norsonic Nor276 and amplifier Nor280) generating logarithmic sine sweeps (3 sec, 20 Hz-20 kHz) and placed on stage at the usual position of a solo singer. A class-1 measurement microphone (NTi Audio M2230) was placed 1 m away from the acoustical center of the source in its horizontal plane. The early interaural cross correlation ($IACC_{\text{early}}$) was measured on stage by means of a dummy head (Neumann KU100) and a small broadband speaker (Fostex 6301B) with a directivity close to the human voice, and positioned just before the mouth of the head. In addition, the dummy head and the speaker being placed on a rotating plate, binaural room impulse responses were recorded for a range of head orientations from -45° to $+45^\circ$ in steps of 5° , with 0° being the orientation of the singer facing the middle of the room. These measurement settings are illustrated in Fig. 1

2.2 Virtual acoustic environments

The experiment conducted in the physical rooms was replicated in an anechoic chamber by means of dynamic binaural synthesis. The experimental setting made it possible for the singers to sing in the eight different environments within a couple of hours, whereas it was not possible to access more than two physical rooms in one day. The virtual acoustic environments were rendered using the recorded binaural impulse responses. The SoundScape Renderer software [15] was used in conjunction with a head tracker (Polhemus Patriot) to generate real-time convolution of the interpolated impulse responses in the horizontal plan with the anechoic signal from the singer's microphone, reproduced with open headphones (AKG K701), which were equalized by the inverse filter of their transfer function, to the singer's ears.

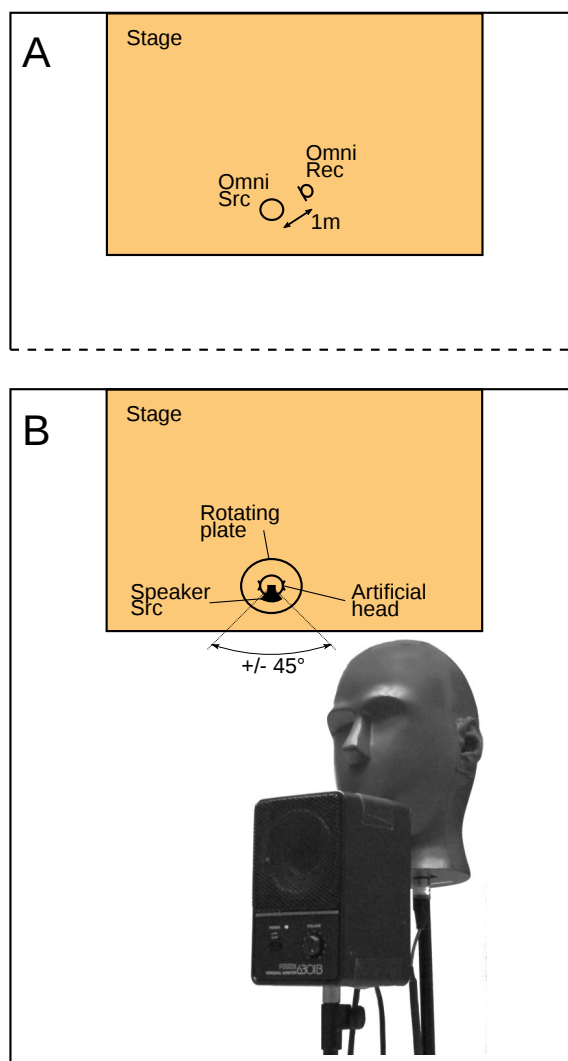


Fig. 1: Room impulse response measurement setups: A (top) omnidirectional sound source and receiver on stage, at the position of the singer. B (bottom) Artificial head and loudspeaker together on a rotating plate.

Room name	Volume (m ³)	EDT (s)	G _{late} (dB)	BR	IACC _{early}
Recording studio (w/o banners)	420	0.7	12.2	1.18	0.73
Recording studio (w/ banners)	420	0.5	8.2	1.55	0.74
Kammersaal	590	1.0	13.8	1.17	0.74
Cabaret theater Distel	1700	0.5	8.5	1.23	0.76
Joseph Joachim Saal (w/o banners)	3660	1.8	10.8	0.94	0.77
Joseph Joachim Saal (w/ banners)	3660	1.3	8.9	1.21	0.77
St Eduard Church	9360	4.3	14.0	0.97	0.54
Philharmonie Berlin	22000	1.2	3.8	0.88	0.74

Table 1: Six musical performance venues representing eight different room acoustical conditions. Four room acoustical parameters, all measured on stage, were selected for the current analysis: Early decay time (EDT), late strength (G_{late}), bass ratio (BR), and interaural cross correlation (IACC_{early}). All parameters were averaged over the 500 Hz and the 1 kHz octave bands, except for IACC_{early}, which was averaged over the 125 Hz to 4 kHz octave bands, according to ISO3382-1 [9].

2.3 Musical recordings

In each of the (physical and virtual) rooms, four solo singers were invited to sing a musical program, including excerpts of two musical pieces with different tempi. The vocal register of the singers and musical pieces are shown in Tab. 2. Recordings were done with a miniature omnidirectional microphone (DPA 4060) positioned 3 cm away from the singer’s mouth edge, with an external audio interface (RME Fireface UFX) allowing to capture the voice as a quasi-anechoic signal in the physical rooms.

3 Analysis

3.1 Feature extraction

Based on the measured impulse responses from both the omnidirectional microphone and the dummy head, standard room acoustical parameters were calculated according to ISO 3382-1 [9]. To analyze the interrelation of room acoustical conditions and performance qualities, four parameters were selected which represent important perceptual properties [16] and which were shown to be largely independent of each other in earlier studies [8, 10]. These include the Early Decay Time (EDT) as a good predictor of the perceived reverberance [17], the early Interaural Cross Correlation (IACC_{early}) as a measure of the perceived source width [18], the level of the late acoustic response (G_{late}) as a measure of the perceived acoustic support of the hall [19], and the Bass Ratio (BR) as a measure of the

tonal color of the room [20]. The indices *early* and *late* denote the time intervals [0;80] ms and [80;+∞] ms respectively. The parameters were calculated using the ITA Toolbox [21] for EDT and custom made codes for BR, IACC_{early} and G_{late} based on the ISO3382-1.

To describe the characteristics of a musical performance, audio features were extracted from the recordings of each singer in each room. As a first step, the onset of each note was detected by generating a MIDI representation from the music score of the performers and using a software which compares a MIDI score with the corresponding audio recording by means of a dynamic time warping algorithm [22]. The obtained onsets were manually corrected whenever necessary on audio-visual inspection with the SonicVisualiser [23] software. As a second step, a collection of 20 audio features were extracted between two consecutive onsets which were shown to be good predictors for musical performance properties, when used in the form of multiple regression models [24]. The present study uses four of these models in order to predict the mean *tempo*, mean *loudness*, mean *timbral brightness*, and *timbral bandwidth* of each musical performance.

3.2 Statistical analysis

To analyze the interrelation of room acoustical parameters and the properties of the musical performance, the dependent variables were z-standardized, according to

$$Xz = \frac{X - \text{mean}(x)}{\text{std}(X)}, \quad (1)$$

Vocal register	Age	Musical pieces	Pace
Soprano	23	Verdi – Aria from <i>Rigoletto</i>	Slow
		Puccini – Aria from <i>Manon Lescaut</i>	Fast
Mezzo soprano	27	Schubert – <i>Nur wer die Sehnsucht kennt</i>	Slow
		Bach – <i>Esurientes implevit bonis</i> (Aria from <i>Magnificat</i>)	Fast
Tenor	21	Schubert – <i>Ganymed</i>	Slow
		Händel – Aria from <i>Tamerlano</i>	Fast
Baritone	26	Fauré – <i>Au cimetière</i>	Slow
		Händel – <i>Honor and arms</i> (Aria from <i>Samson</i>)	Fast

Table 2: Singers and musical pieces

with *mean* and *std* denoting the arithmetic mean and the standard deviation of the series X . The standardization was applied to shift the tempi, the loudness and the timbral characteristics of all musical pieces to the same range, only retaining the effect on their relative variation. All possible pairs of variables from both domains were then tested by linear regression, yielding values between 0 (no correlation) and ± 1 (perfect positive or negative correlation), along with p-values denoting the asymptotic significance of the regression models. These results can be interpreted as quantitative measure to which extent the musical interpretations of the singers are influenced by the acoustics of the room. Comparing the strength and the significance of this interrelation obtained in physical and virtual rooms then yields information about the influence of the environments on this adaptation process.

4 Results

4.1 Physical and virtual rooms

Bivariate Pearson’s correlation coefficients are calculated for each possible pair of room acoustical parameters and performance features, for each singer and each musical piece, over all rooms. Significant correlations ($p < 0.05$) are shown in Fig. 2. The results demonstrate that the strength of the interrelation between room acoustic conditions and musical performance depends on the person, with singer 3 showing a more pronounced adaptation to the room than the other singers. There is, however, no systematic difference between the values for physical and virtual environments. Both the number and the strength of significant interactions is almost equal for both conditions. Thus, the impact of the acoustical conditions seems to be equally effective in both environments.

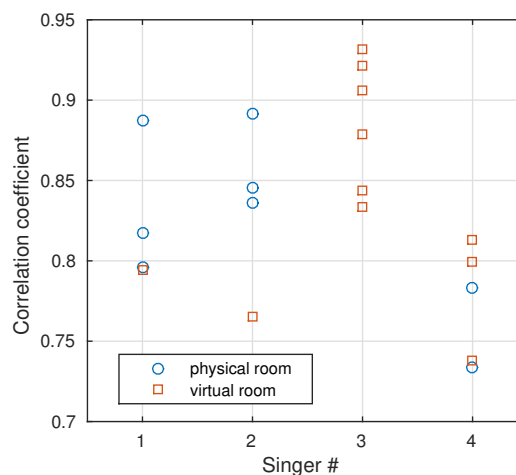


Fig. 2: Absolute value of the Pearson’s correlation coefficients calculated between the room acoustical parameters and the audio variables over all rooms with a p-value such as $p < 0.05$.

4.2 Room acoustics and musical performance

Figure 3, showing the interrelation between the four selected room acoustical parameters and the four musical performance features, cumulated over all singers and musical pieces, seems to indicate the absence of any room acoustical effect: none of the linear regressions are significant, neither for the physical nor for the virtual mode of presentation.

Looking at the significant ($p < 0.05$) interrelations for *individual* singers and musical pieces (Fig. 4), however, demonstrates that this is a premature conclusion. It can be seen that 13 different pairs of room acoustical and musical performance features yield significant correlations, out of a total of 19 possible pairs. This reveals that all of the considered room acoustical features seem to be effective; the manifestation of this effect, i.e. which performance feature is affected by which room acoustical features, however, seems to be quite individual for the musical performer and the musical piece considered.

Some of them even show contradictory trends of adaptation, such as the effect of the Late sound strength (Glate), indicating the ratio between the late energy and the direct sound within the room impulse response as a measure for the sound support of the room, on the timbral brightness of the vocal performance, which was positive for one singer and one musical piece and negative for another.

5 Conclusion

The present study investigates how solo singers adapt their vocal performance to different room acoustical conditions, which was observed both in physical and virtual acoustic environments. Two musical pieces were sung by four singers in eight acoustically different conditions. The results show rather individual patterns of adaptation, depending on the musical piece and the musical performer. Although these individual interrelations can be quite strong, with correlation coefficients (Pearson's r) between 0.7 and 0.95, it is thus difficult to generalize, how "average" performers singing "average" musical pieces react to the acoustical environment. This observation corresponds to the results of earlier research on instrumental performers [10].

It is interesting, however, that the same interactions between room acoustics and musical performance could

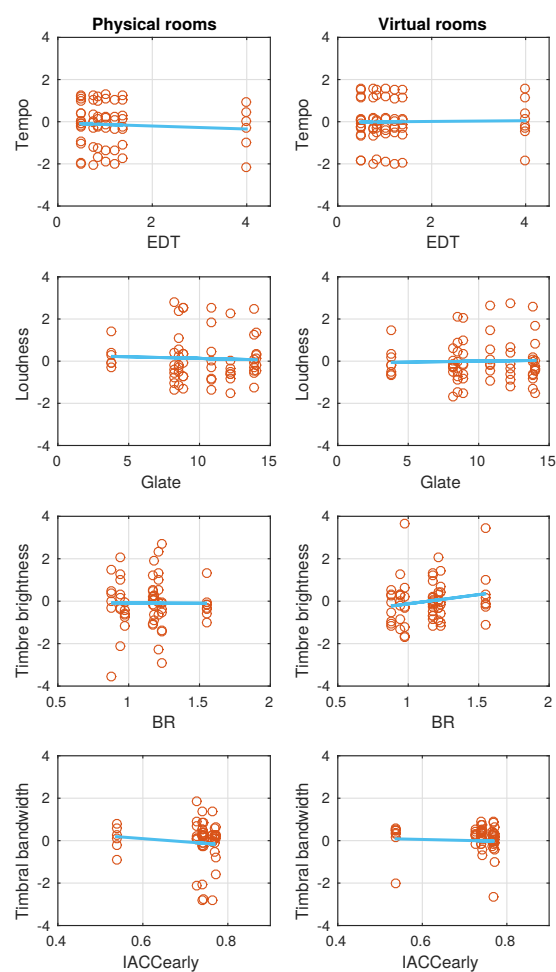


Fig. 3: Interrelation between room acoustical parameters and musical performance features (z-standardized), with linear regressions, observed in the physical rooms (left) and virtual rooms (right). Each point represents one vocal performance.

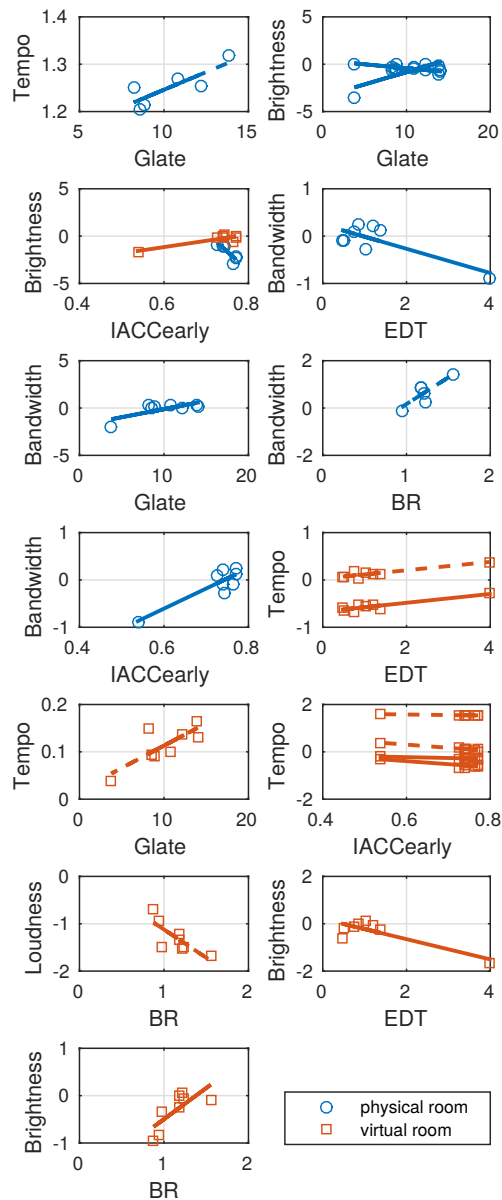


Fig. 4: Significant interrelations ($p < 0.05$) between room acoustical parameters and musical performance features (z-standardized) for individual singers and musical pieces, with linear regressions, observed in the physical and virtual rooms. Each point represents one vocal performance. The solid and dashed lines represent the musical pieces 1 and 2, respectively.

be observed in the physical and virtual environment. This is true both for the strength and the direction of the correlations, and seems to indicate the suitability of virtual acoustic environments to investigate musical tasks. Dynamic binaural synthesis can not only generate authentic acoustic experiences for music listeners [25], but also seems capable of reproducing the interaction of musicians with their acoustic environment in a way that does not affect the influence of environmental conditions on the musical performance itself.

6 Acknowledgements

This work was partially funded by the Alexander von Humboldt Foundation with a grant to P. Luizard. The authors would like to thank the singers for their interest in this study, the room managers for letting us use the venues, and C. Böhm for his help in running the real-time rendering system.

References

- [1] Richardson, A. E., Montello, D. R., and Hegarty, M., “Spatial knowledge acquisition from maps and from navigation in real and virtual environments,” *Memory & cognition*, 27(4), pp. 741–750, 1999.
- [2] Bronkhorst, A. W., “Localization of real and virtual sound sources,” *The Journal of the Acoustical Society of America*, 98(5), pp. 2542–2553, 1995.
- [3] Rébillat, M., Boutillon, X., Corteel, É., and Katz, B. F., “Audio, visual, and audio-visual egocentric distance perception by moving subjects in virtual environments,” *ACM Transactions on Applied Perception (TAP)*, 9(4), p. 19, 2012.
- [4] Larsson, P., Vastfjall, D., and Kleiner, M., “Better presence and performance in virtual environments by improved binaural sound rendering,” in *Audio Engineering Society Conference: 22nd International Conference: Virtual, Synthetic, and Entertainment Audio*, Audio Engineering Society, 2002.
- [5] Rottondi, C., Chafe, C., Allocchio, C., and Sarti, A., “An overview on networked music performance technologies,” *IEEE Access*, 4, pp. 8823–8843, 2016.

- [6] 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*, 96(3), pp. 505–515, 2010.
- [7] Kato, K., Ueno, K., and Kawai, K., “Effect of Room Acoustics on Musicians’ Performance. Part II: Audio Analysis of the Variations in Performed Sound Signals,” *Acta Acustica united with Acustica*, 101, pp. 743–759, 2015, doi: 10.3813/AAA.918870.
- [8] Schärer Kalkandjiev, Z. and Weinzierl, S., “The Influence of Room Acoustics on Solo Music Performance: An Empirical Case Study,” *Acta Acustica united with Acustica*, 99, 2013, doi: 10.3813/AAA.918624.
- [9] ISO3382-1, “Acoustics - Measurement of room acoustic parameters - Part 1: Performance spaces,” 2009.
- [10] Schärer Kalkandjiev, Z. and Weinzierl, S., “The Influence of Room Acoustics on Solo Music Performance: An Experimental Study,” *Psychomusicology: Music, Mind, and Brain*, 25(3), 2015.
- [11] Marshall, A. and Meyer, J., “The Directivity and Auditory Impressions of Singers,” *Acustica*, 58, pp. 130–139, 1985.
- [12] Noson, D., Sato, S., Sakai, H., and Ando, Y., “Singer responses to sound fields with a simulated reflection,” *Journal of Sound and Vibration*, 232(1), pp. 39–51, 2000.
- [13] Seraphim, H., “Investigations on the difference limen of exponential decay for broadband impulse,” *Acustica*, 8, pp. 280–284, 1958.
- [14] Martellotta, F., “The just noticeable difference of center time and clarity index in large reverberant spaces,” *The Journal of the Acoustical Society of America*, 128(2), pp. 654–663, 2010.
- [15] Geier, M., Ahrens, J., and Spors, S., “The Sound-Scape Renderer: A unified spatial audio reproduction framework for arbitrary rendering methods,” in *In 124th AES Convention in Amsterdam, The Netherlands, May 17–20, 2008*.
- [16] Weinzierl, S., Lepa, S., and Ackermann, D., “A measuring instrument for the auditory perception of rooms: The Room Acoustical Quality Inventory (RAQI),” *The Journal of the Acoustical Society of America*, 144(3), pp. 1245–1257, 2018.
- [17] Barron, M., “Subjective study of British symphony concert halls,” *Acta Acustica united with Acustica*, 66(1), pp. 1–14, 1988.
- [18] Beranek, L., *Concert Halls and Opera Houses: Music, Acoustics and Architecture*, 2nd ed., Springer-Verlag, New York, 667 p., 2004, appendix 3 - p 615.
- [19] Dammerud, J. J., Barron, M., and Kahle, E., “Objective Assessment of Acoustic Conditions for Symphony Orchestras,” *Building Acoustics*, 18(3), pp. 207–219, 2011.
- [20] Beranek, L., *Music, acoustics, and architecture*, Wiley, New York, 1962.
- [21] Berzborn, M., Bomhardt, R., Klein, J., Richter, J.-G., and Vorländer, M., “The ITA-Toolbox: An open source MATLAB toolbox for acoustic measurements and signal processing,” in *Proc. of the 43th Annual German Congress on Acoustics, Kiel, Germany*, pp. 6–9, 2017.
- [22] Lerch, A., *Software-based extraction of objective parameters from music performances*, Ph.D. thesis, TU Berlin, 2008.
- [23] Cannam, C., Landone, C., and Sandler, M., “Sonic Visualiser: An Open Source Application for Viewing, Analysing, and Annotating Music Audio Files,” in *Proceedings of the ACM Multimedia International Conference*, pp. 1467–1468, Firenze, Italy, 2010.
- [24] Weinzierl, S. and Maempel, H.-J., *Zur Erklärbarkeit der Qualitäten musikalischer Interpretationen durch akustische Signalmaße [On the prediction of qualities of musical performances by means of acoustic signal features]*, Schott, Mainz et al., 2011, 213–236.
- [25] Brinkmann, F., Lindau, A., and Weinzierl, S., “On the authenticity of individual dynamic binaural synthesis,” *The Journal of the Acoustical Society of America*, 142(4), pp. 1784–1795, 2017.