Singers’ likability of the acoustic environment and adaptation of singing in contemporary churches

Proceedings of Meetings on Acoustics

Volume 58 http://acousticalsociety.org/

International Symposium on Music and Room Acoustics

24-28 May 2025

Loyola University
New Orleans, Louisiana

Architectural Acoustics

Singers’ likability of the acoustic environment and adaptation of singing in contemporary churches

Louena Shtrepi and Nicolò Somà
Department of Energy, Politecnico di Torino, Torino, 10129, ITALY; louena.shtrepi@polito.it; S310341@studenti.polito.it

Angela Guastamacchia
Politecnico di Torino, Torino, ITALY; angela.guastamacchia@polito.it

Davide Sgro
Department of Electronics and Telecommunications, Politecnico di Torino, Torino, ITALY; davide.sgro@polito.it

Giuseppina Emma Puglisi
University Sustainability, Research Infrastructures and Laboratories, Politecnico di Torino, Torino, ITALY; giuseppina.puglisi@polito.it

Alessio Carullo
Department of Electronics and Telecommunications, Politecnico di Torino , Torino, ITALY; alessio.carullo@polito.it

Arianna Astolfi
Department of Energy, Torino, 10129, ITALY; arianna.astolfi@polito.it

Several studies have reported that singers’ performances vary across different acoustic environments due to the perception variations of the space acoustic characteristics. This study investigated these relationships based on the feedback of six professional singers performing in four contemporary churches. The acoustic parameters according to ISO 3382-1 were measured and singing delivery parameters such as the Power Singing Ratio and Cepstral Peak Prominence Smoothed (CPPS) were analyzed. In order to investigate the singers’ likability of the acoustic environment, the study has been structured in two phases: 1) preliminary subjective feedback of the singers has been collected immediately after the four venues performance, showing that they naturally adapted their singing to auditory feedback of each church, and 2) a subjective listening test with amateur singers has been conducted in the Audio Space Lab at the Department of Energy, Politecnico di Torino, Italy. This laboratory, equipped with a 16-loudspeaker spherical array system, enabled real-time auralization based on 3rd order Ambisonic impulse responses recorded in a single position within each church. The amateur singers were asked to evaluate their preferences after performing in each auralized environment, and their responses were compared to those of professional singers who had performed in the four churches.

Proceedings of Meetings on Acoustics, Vol. 58, 015006 (2025); https://doi.org/10.1121/2.0002144.

1. INTRODUCTION

Although the relationship between acoustics and musicians' performance in a musical venue has been documented for centuries, little attention has historically been paid to how singers adapt to different acoustic environments, with mixed findings in the existing literature.[1] Research has recently focused on the influence of simulated environments on singers, particularly through the use of auralization, an approach that enables the placement of sound within realistic acoustic contexts, as previously measured and reconstructed through rendering systems.[2–4]

In past literature, great attention has been given to the use of Geometrical Acoustics as a tool to simulate acoustical environments. This tool is particularly interesting when dealing with acoustic spaces which no longer exist, but require precise calibration procedures[5] to ensure fidelity to the real characteristics of the reconstructed environments, and grant reliable results. Also, for what concerns real-time auralization of singing, studies have been reported in past literature to use open-back headphones to render the output of the auralization systems,[2] which are a more portable and cheaper solution,[3] but tend to introduce unpredictable filtering which may alter singers' perception of their own voice.[4]

Mixed findings have been reported when attempting to identify general patterns of vocal adaptation to the acoustic characteristics of performance venues. *Luizard et al.* conducted experiments in several venues with measurable acoustic differences. While no consistent trends were found when analyzing the overall behavior of singers, individual adaptations were observed, suggesting that each performer modifies their vocal output in a distinct, personal way in response to different acoustical conditions.[6,7] *Redman et al.*,[8] building upon previous findings by *Bottalico et al.*,[9] which reported notable relationships between vocal and room acoustic parameters, proposed three perceptual metrics potentially related to singers' reported likability of performance venues. Their results indicated that singers tend to prefer rooms with higher Room Supportiveness and Room Noiselessness, each of which correlates with specific room acoustic parameters—Vocal Support and Early Decay Time, respectively.

Few studies have already focused on the effects of real-time auralizations using spherical loudspeaker arrays on singers. One such investigation was conducted in the Virtual Singing Studio,[10] where only limited effects on singing performance were observed.[11]

The objective of the present study is to assess the feasibility of a real-time system for the vocalization and auralization of solo singing in the Audio Space Lab at Politecnico di Torino, Italy. Four modern churches were acoustically measured, and their impulse responses were captured and processed for use in the laboratory's convolution stage. Eighteen amateur singers participated in a subjective investigation to assess their perception of the reconstructed acoustics and to rate the environment in which they most preferred to perform. Finally, recordings of the participants performing a simple song were analyzed to extract vocal parameters and investigate potential adaptation patterns that could be associated with specific room acoustic features.

2. METHODOLOGY

A. CASE STUDIES: OBJECTIVE AND SUBJECTIVE CHARACTERIZATION

Four modern churches were used as the case study. These same buildings were the subject of a previous study by *Shtrepi et al.*,[12] where more details about the constructive and architectural features are provided. The buildings, presented in Fig. 1, are located in the city of Milan, Italy, and are characterized by pretty different styles and shapes. Also, the materials used the walls and ceilings have significant effects on the acoustical features of each environment. To facilitate indexing, the churches will hereafter be identified using numerical indices or a dedicated color code, as illustrated in Fig. 1. Room acoustical parameters were extracted in accordance with the ISO 3382-1 standard,[13] and following the framework proposed by

Figure 1: View of the interior of the four churches

Martellotta et al.,[14] which is specifically designed for measurements in this type of environment. Numerous positions in each church were used to capture omni-directional recordings from which acoustical parameters in different positions could be extracted. However, for the present study, only the recordings in the specific reference positions were considered while computing the room acoustic parameters. The floor plans of the four churches, with the measurement positions highlighted by colored dots, are shown in Fig. 2.

In each church, an omnidirectional sound source (B&K 4296, indicated by red dots) was positioned near the

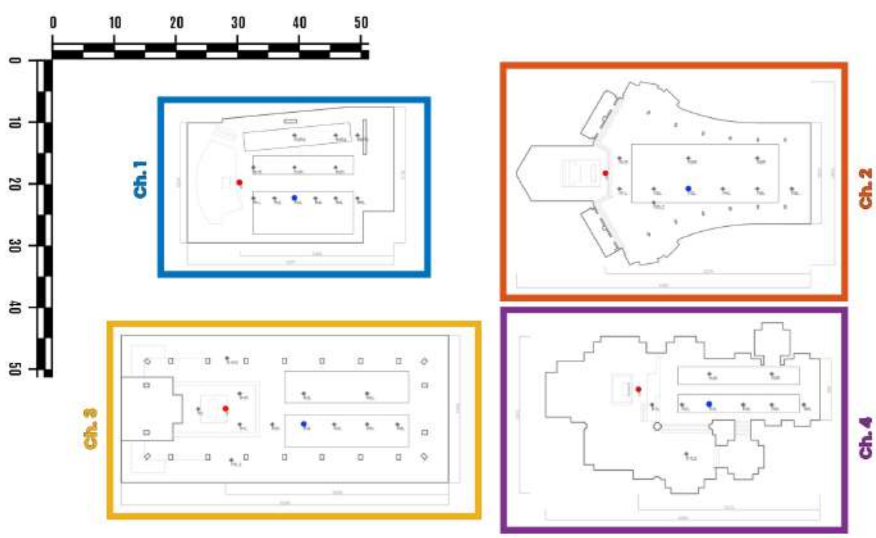


Figure 2: Plans of the four churches. Measurement instruments positioning: omnidirectional source B&K 4296 (red dots); Sound Level Meter NTi Audio XL2 (gray dots); Spherical Microphone Array Zylia ZM-1 (blue dots)

altar. A sound level meter (NTi Audio XL2, gray dots), equipped with a single omnidirectional capsule, and a Spherical Microphone Array (SMA) (Zylia ZM-1) were placed at the reference performing position (blue dots). The omni-directional recordings acquired through the sound level meters were used to extract room acoustic parameters, while the SMA recordings were used to obtain the impulse responses of the venues, as depicted in later sections. For the present studies, some objective room acoustics parameters were extracted, like Early Decay Time (EDT), Reverberation Time (T_{20}), Clarity (C_{80}) and Definition (D_{50}). While EDT is broadly reported to be correlated to the subjective impression of perceived reverberance, C_{80} could be referred to as the perceived quality of transmission from a source to the listener's position, and it's specific for use in musical applications. These parameters were obtained using a MATLAB script, calculating values for each octave band from the ones centered in 125 Hz to 4000 Hz. In order to consider a single value, which may be compared between the four venues, an arithmetical mean of the parameters in the octave bands centered in 500 Hz and 1000 Hz was computed. Results are presented in Fig. 3. It is worth noting that, although Church 4 exhibits reverberation times similar to those of Church 2, it features a narrower nave and several large pillars located near the measurement position. These architectural elements may have a significant impact on early reflections, and consequently on the acoustic characteristics as perceived at that location.

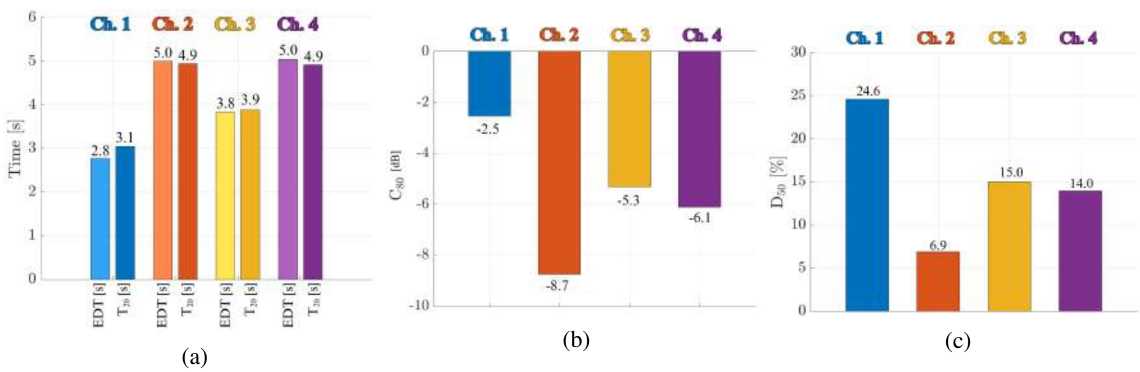


Figure 3: Room acoustical parameters measured in the four churches. (3a): Early Decay Time EDT and Reverberation Time T_{20}; (3b): Clarity C_{80}; (3c): Definition D_{50}

As reported in *Shtrepi et al.*,[12] an ensemble of professional singers, *Faber Teater* , performed in each of the four churches considered in the present study. In particular, during their 2022 tour in these four churches, they presented the *Stabat Mater*, a composition designed to exploit the full spatial extent of sacred buildings to convey emotional content to the audience. After the performances, the singers shared their impressions of each church through open-ended comments, highlighting the characteristics that most influenced their expressive choices and the adaptations they made. They also provided a preference ranking based on the venues in which they felt most comfortable performing. Church 2 and Church 4 were rated as the most favorable environments, as both offered sufficient acoustic support without being overly noisy. Conversely, Church 3 was less appreciated due to a perceived "cut" in the mid-frequency range, which the singers felt compelled to compensate for with their vocal emission. Church 1 received the lowest preference, as it was described as offering little support and making it particularly difficult to perform in synchrony with the other singers.

B. VOCALIZATION AND AURALIZATION SYSTEM SETUP IN THE AUDIO SPACE LAB

i. Preprocessing for singing vocalization and auralization

The real-time vocalization and auralization system was developed to allow singers to perform in the Audio Space Lab (ASL) at Politecnico di Torino, Italy while receiving acoustic feedback as if they were singing in the actual venues. The laboratory is equipped with a spherical loudspeakers array, composed of sixteen 2-way loudspeakers (Genelec 8030B) and two 3-way loudspeakers (Genelec 8351A) dedicated to the reproduction of lower frequencies. The room hosting the system was acoustically treated in accordance with the recommendations of the ITU-R BS.1116-3 standard,[15] and has been validated for the reproduction of audio content up to 3rd Order Ambisonics.[16]

To carry out singing vocalization and auralization in 3rd Order Ambisonic laboratory conditions, the impulse response (IR) of the four churches was sampled using the SMA, which was placed in the positions marked with the blue dots in Fig. 2. The SMA was connected to a PC running a custom patch in the Plogue Bidule DAW, from which 3rd Order Ambisonics IRs were derived for use in the auralization setup in the laboratory. To grant correct vocalization reproduction, the direct sound needed to be removed, as it could naturally propagate from the mouth to the ears of the system's users.
For this purpose, a proper filtering preprocessing is proposed, derived from the one used by *Mullins et al.*[2] and reported in Fig. 4: a time window, with a setup time of 2 ms, was used to cancel the first 8 ms after the arrival of the direct sound, while maintaining the part of the IRs containing the sound which arrives after the threshold of 10 ms. This editing should maintain the majority of the early reflections and the whole reverberant tail, which will be reproduced in the auralization process. In this step, we also took into account the latency introduced by the entire auralization system, which represents a critical factor when aiming to reproduce real environments with the highest possible fidelity. This latency was measured to be approximately 8 ms, with slight variations ($\pm$ 1.5 ms) depending on system load.

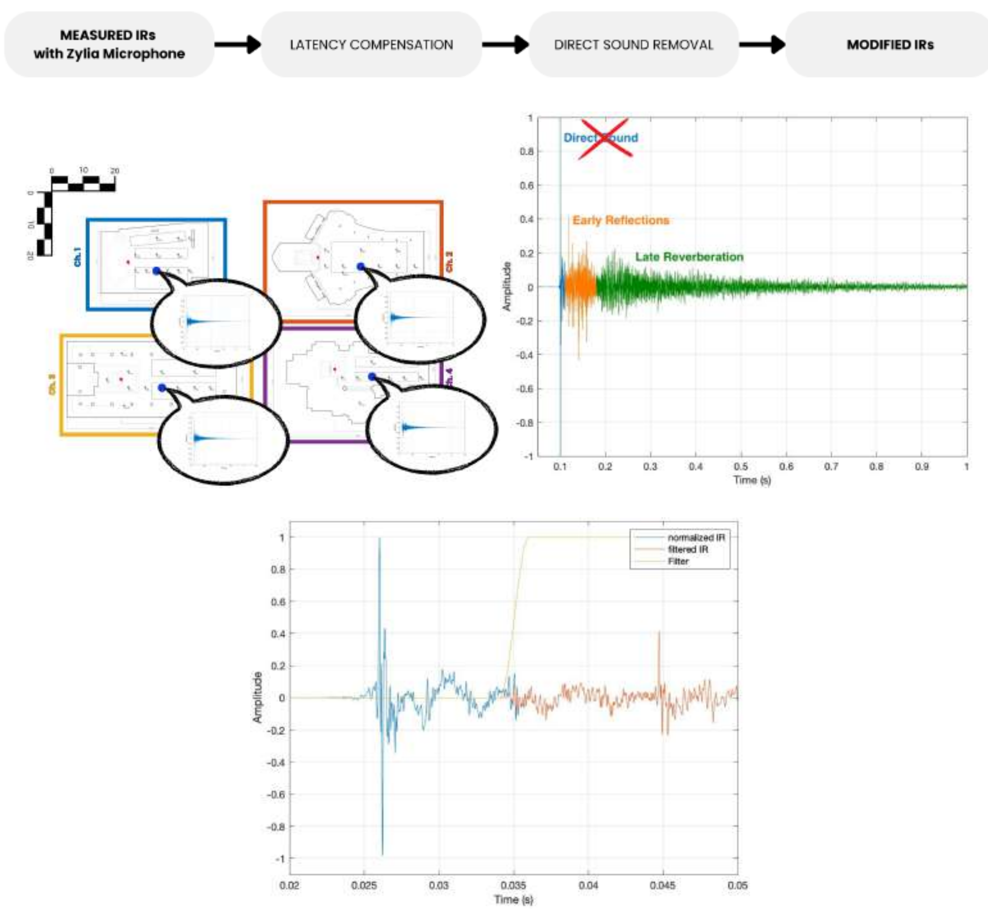


Figure 4: Preprocessing for singing vocalization and auralization in Audio Space Lab

ii. Matching of room acoustical parameters in the four churches and in Audio Space Lab

To evaluate the acoustical characteristics of the environments being reproduced in the ASL, an omnidirectional microphone was placed at the center of the spherical loudspeakers' array. By feeding an exponential sweep signal into the convolution setup, and using the modified IR as presented in the previous section, it was possible to obtain values for both EDT and T_{20}, which could be compared to the values measured in the actual environments. As shown in Fig. 5, the values of both EDT and T_{20} are generally comparable to those measured in the actual environments, with only a few exceptions.

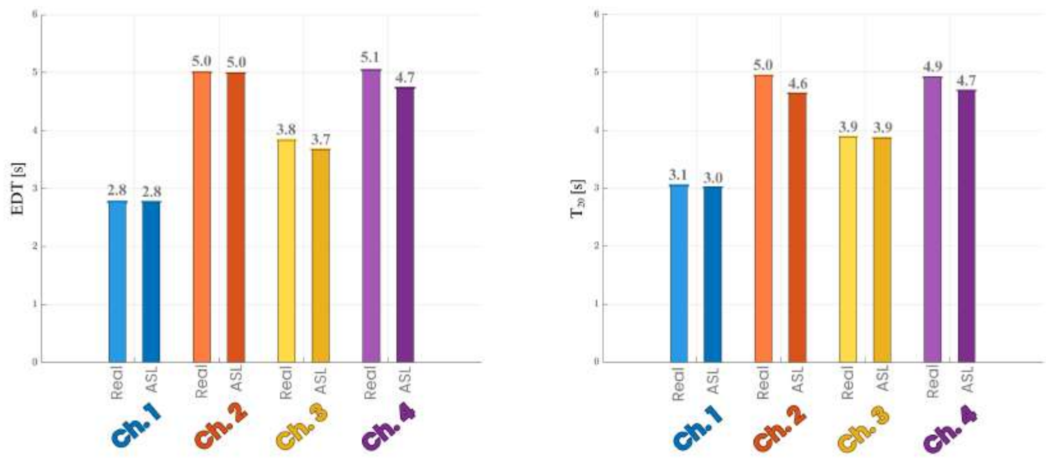


Figure 5: Comparison between EDT(left) and T_{20}(right) measured in the reference position in the four churches (Real) and in the Audio Space Lab (ASL)

As for EDT, the most noticeable discrepancy was observed in Church 4, where the difference between the measured values was approximately 0.3 seconds, corresponding to about 5.9% of the value measured in the actual church. In contrast, for T_{20}, the largest difference was found in Church 2, amounting to roughly 0.3 seconds as well, which corresponds to approximately 6.0% of the measurement obtained in the real environment. Both of these discrepancies exceed the 5% threshold, which corresponds to the Just Noticeable Difference (JND) for both parameters, as defined in ISO 3382-1.[13]

iii. Measurement of Decay Time at the Ears in Audio Space Lab

To obtain a more specific characterization of the perceived reverberance at the listener's ears, the Early Decay Time from Mouth to Ears ($DT_{40,ME}$) was measured using an artificial dummy head simulator (HEAD Acoustics HMS II.3 LN HEC). This parameter has been shown to correlate with EDT and room volume, as well as with self-reported vocal comfort when speaking in unoccupied classrooms.[17] Additionally, values for Room Gain and Voice Support were computed;[18] however, they will not be considered in the present study, as they did not show any significant differences among the four environments.

The dummy head was positioned at the center of the loudspeaker array and equipped with a head-worn microphone (SHURE WBH54), which served as the input transducer for the convolution system. The dummy head was capable of simultaneously emitting a test signal from the mouth simulator and recording the resulting sound at both ear simulators. To measure the $DT_{40,ME}$, the dummy head was used to emit an exponential sweep signal. The signals captured at the ears were then used to calculate the $DT_{40,ME}$ at each ear in octave bands ranging from 125 Hz to 4000 Hz. A single representative value was subsequently derived by averaging the values from both ears across the 500 Hz to 2000 Hz octave bands. It is worth noting that the IRs used in this phase of the convolution system are those obtained from recordings at the reference position in each church, after being processed as described in previous sections to remove the direct sound. Therefore, it can be hypothesized that the results should be comparable to those that would be obtained by placing the dummy head in the corresponding positions within the actual churches. Results are reported in Fig. 6, where the values obtained in each of the four environments show a positive association with both the EDT measured in the real churches and in the ASL.

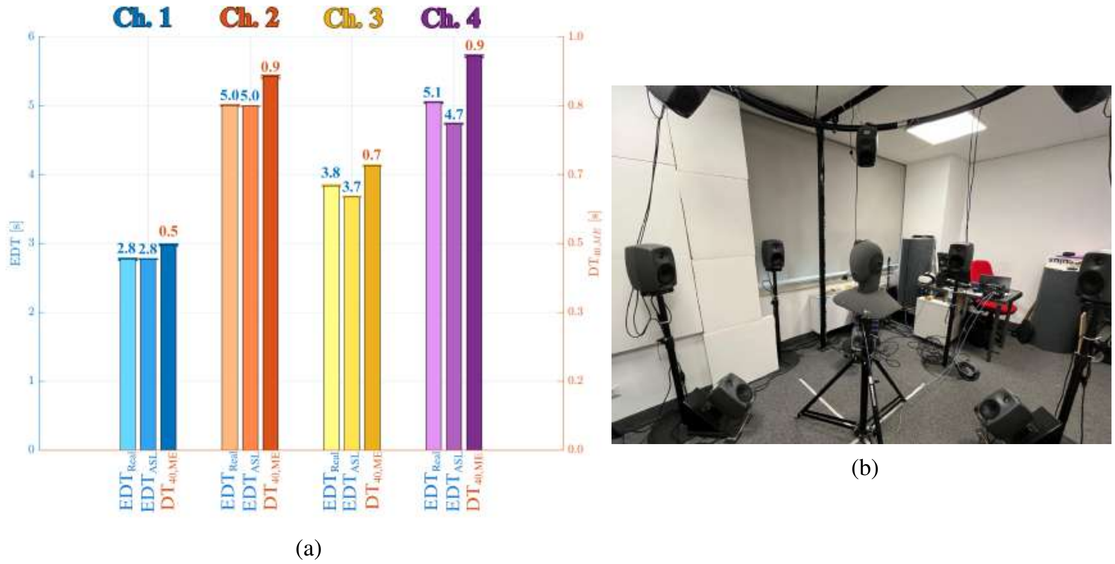


Figure 6: (6a): Comparison of DT$_{40,ME}$, measured using the dummy head in the ASL, and values of EDT from the real churches (EDT$_{Real}$) and in the Audio Space Lab (EDT$_{ASL}$); (6b): measurement setup with the dummy head in ASL

iv. Vocalization Setup

The block diagram of the real-time convolution system used in the ASL is shown in Fig. 7a. The head-worn microphone (SHURE WBH54) employed in this setup was selected for its linear frequency response and its directional polar pattern, which ensures that only the sound emitted from the user's mouth is captured, while remaining largely insensitive to the reverberant tails reproduced by the loudspeakers.
The convolution was performed by a VST3 plugin (X-MCFX Convolver Node[19]), while the decoding stage was calibrated and validated in previous studies[20] to ensure the correct reproduction of 3rd Order Ambisonics audio content.

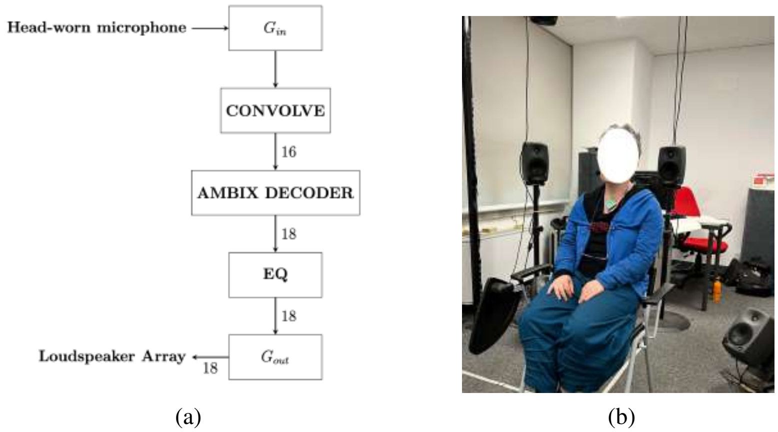


Figure 7: (7a): Block Diagram of the Real-Time Auralization System; (7b): Example of user singing in the ASL with the head-worn microphone

C. SUBJECTIVE INVESTIGATIONS IN AUDIO SPACE LAB

i. Subjective Tests

To investigate the perceptual characteristics of the environments reconstructed in the ASL, a subjective test protocol was developed. Eighteen amateur singers (14 M, 3 F, 1 N.B., 4 to 20 years of practice)—all native Italian speakers or with several years of experience with the language, with little or no prior knowledge of acoustics or music theory—were invited to freely explore the proposed virtual environments. They subsequently rated their perceptions using a questionnaire composed of 11 descriptors, each evaluated on a discrete Likert scale. Following the exploration phase, participants were recorded while performing a simple song, "Happy Birthday to You", in English. To ensure a meaningful comparison across conditions, each participant was assigned a reference note prior to recording in each of the four auralized environments, selected according to their vocal range and provided using a virtual piano mobile application.[21]
They were also asked to indicate the environments in which they most enjoyed performing. No specific instructions were provided by the operators, yet most participants reported that their choices were based on how easily they could hear their own voice and whether they felt adequately supported while singing.
The descriptors used in the questionnaire were freely adapted from terms used in listening evaluation reported in previous literature,[22,23] and were the following: Three-Dimensionality, from *low* to *high*; Plausibility, from *low* to *high*; Spatial Presence, from *low* to *high*; Ease of Singing, from *low* to *high*; Coloration, from *dark* to *bright*; Space Size, from *small* to *big*; Reverberance, from *dry* to *reverberant*; Clarity, from *muddy* to *clear*; Pitch, from *deep* to *high*; Vibrato, from *hard* to *easy*; Velocity, from *slower* to *faster*.

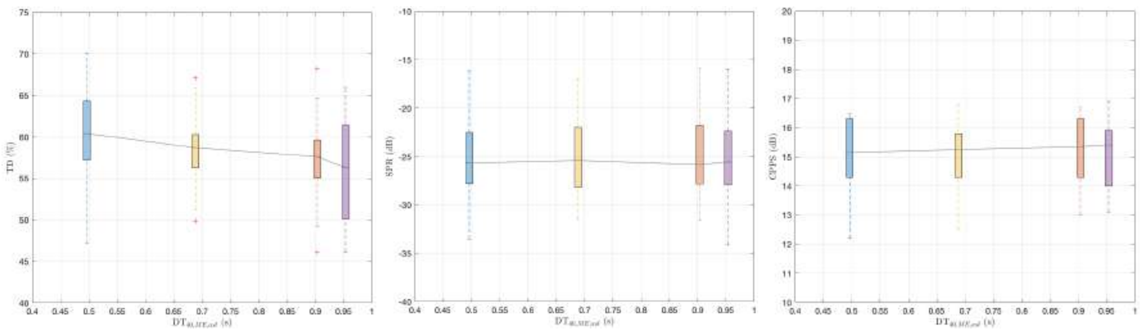
ii. Voice parameters

Voice parameters were extracted from the recordings of the amateur singers while performing "Happy Birthday to You". As the literature has yet to clearly establish a set of parameters specifically describing the features and variations of singing voice behavior, a broad range of vocal metrics was considered. Sound Pressure Level (SPL) [dB] was calculated based on a standard calibration signal. Time Dose (TD) [%], a metric indicating the total vibration time of the vocal folds, was computed as described by *Carullo et al.*[24]
The mean Fundamental Frequency (F_0) [Hz] was calculated as the average value across each recording, normalized by subtracting the frequency of the reference note provided to each singer before the performance. The Singing Power Ratio (SPR) [dB], associated with the perceptual quality of "ringing" in the singing voice, was included as a marker of vocal quality.[25] Cepstral Peak Prominence Smoothed (CPPS) [dB], while widely used to distinguish dysphonic from normophonic voices,[26] has limited validation as a metric of singing quality.[27]

3. RESULTS

A. VOICE PARAMETERS AND ROOM ACOUSTICAL PARAMETERS

Associations between room acoustical parameters and vocal parameters are presented graphically: for each environment, the vocal parameters extracted from the recordings are displayed as dependent variables of the objective room acoustics parameters. Fig. 8a depicts the relationship between Time Dose (TD) and the decay time mouth-ears ($DT_{40,ME}$), where a slight negative association was found. Fig. 8b and Fig. 8c show the relationships between the Singing Power Ratio (SPR) and Cepstral Peak Prominence Smoothed (CPPS) with respect to $DT_{40,ME}$. Analogous comparisons were conducted considering EDT and T_{20}; however, no significant associations were observed.



(a) (b) (c)

Figure 8: Comparison between voice parameters (Time Dose-TD, Singing Power Ratio-SPR and Cepstral Peak Prominence Smoothed-CPPS) and decay time at the ears (DT$_{40,ME}$)

B. SUBJECTIVE IMPRESSIONS AND ROOM ACOUSTICAL PARAMETERS

This section presents only the subjective parameters that exhibited an association with the trends of the objective room acoustics parameters. Other subjective descriptors (Plausibility, Ease of Singing, Coloration, Clarity, Pitch, Vibrato and Velocity), which did not show any meaningful correlation, are therefore not discussed here. It is worth noting that, as expected, parameters that are correlated with each other tend to display similar trends: for example, EDT, T_{20} and DT$_{40,ME}$ show similar trends for all the associations proposed hereafter; conversely, C_{80} and D_{50} showed similar trends for the majority of associations.
The relations are summarized in Fig. 9.

- Figure 9a shows that the perceived *Three-Dimensionality* of the auralized environments tends to increase with the objective reverberation time, as measured by EDT, T_{20}, and DT$_{40,ME}$. Conversely, it tends to decrease as the clarity indexes (C_{80} and D50) increase.
- As shown in Fig. 9b, participants reported that perceived *Spatial Presence* tends to increase with objective reverberation time (EDT, T_{20}, DT$_{40,ME}$) and to decrease as clarity indexes—specifically C_{80} and D_{50}—increase.
- Figure 9c reports that perceived *Reverberance* showed a positive association with the values of objective reverberation time, especially EDT and DT40,ME, while it showed to decrease as D_{50} increases.
- Participants reported that the reconstructed space tended to be perceived as increasingly larger as the objective reverberation time increased. Conversely, they indicated that *Space Size* tended to decrease as the clarity indexes D_{50} and C_{80} increased. These trends are illustrated in Fig. 9d.

C. SUBJECTIVE RANKING BY AMATEUR SINGERS IN AUDIO SPACE LAB

The amateur singers involved in the subjective investigation were asked to express their preference among the experienced auralized environments. Church 2 emerged as the most preferred environment, whereas Church 1 and Church 3 received fewer mentions. Unexpectedly, Church 4 was the least favored setting, as several participants described perceiving dissonances or an overemphasis of higher frequency ranges, which resulted in unpleasant and disorienting auditory feedback.

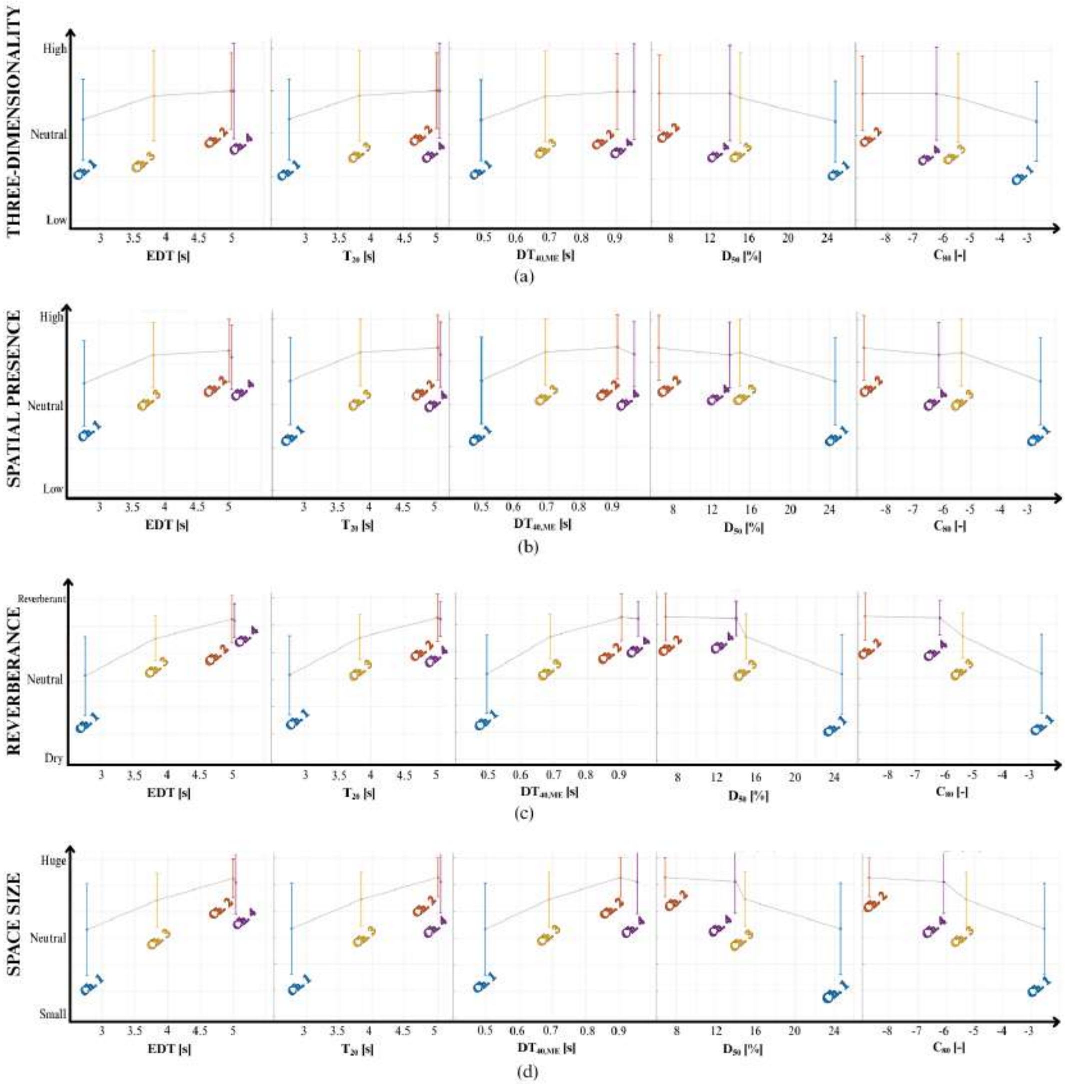


Figure 9: Comparison between subjective impressions (Three-Dimensionality, Spatial Presence, Reverberance and Space Size) and reverberation objective parameters (EDT, T_{20}, $DT_{40,ME}$, D_{50} and C_{80})

4. DISCUSSION

Time Dose (TD) showed an association with room acoustical parameters, in particular with EDT, T_{20} and $DT_{40,ME}$. This vocal parameter was computed using a simple threshold-based technique, as described in previous studies.[24] The selection of the threshold is a critical aspect, as it directly affects the ability to accurately discriminate between voiced and unvoiced segments of the recordings. Furthermore, the effectiveness of this algorithm is highly dependent on the presence of reverberant tails, which may result in the misclassification of unvoiced frames as voiced. In the present setup, this issue is mitigated by the use of a head-worn microphone, which is oriented toward the singer's mouth and therefore captures primarily the direct sound, minimizing the influence of reverberant energy reproduced by the surrounding loudspeakers. Additionally, the ASL is characterized by low background noise levels,[20] which further reduces the likelihood of interference with the discrimination algorithm. The findings reported in Sec. 3.A show that TD tends to decrease as reverberation time increases. This may suggest that singers tend to extend the vocal emission in acoustically dry environments, where reverberant energy decays rapidly, while in more reverberant venues the natural extension of the sound reduces the need for prolonged vocal effort to achieve a similar perceptual outcome. However, this observation may warrant further investigation, as it appears to contrast with previous findings in the literature—such as those reported by Astolfi et al.[28]—where an increase in reverberation time in classrooms was associated with longer voiced segments. The subjective impression of singing Reverberance and Space Size was found to exhibit significant and coherent associations with the objective room acoustic parameters. These findings suggest that participants were generally able to accurately perceive and interpret the characteristics of the real environments being reconstructed in the ASL. Specifically, Church 1 was perceived as less three-dimensional and less immersive, likely due to its perception as a smaller and less reverberant space. In contrast, both Church 2 and Church 4 were rated as highly three-dimensional and immersive, as participants consistently perceived them as larger and more reverberant. Church 3 emerged as a middle ground in terms of these descriptors, being perceived as a compromise between the other environments. All remaining descriptors included in the subjective questionnaire did not show significant associations with the acoustical features of the original spaces. Finally, the comparison between the subjective impression of Preference expressed by amateur singers in the ASL was partially consistent with those reported by professional singers who had performed in the actual churches. Church 2, which exhibited the highest reverberation time, was rated by both groups as the most pleasant environment to perform in, as it provided sufficient support to the singers without resulting in excessive muddiness. Conversely, Church 1 and Church 3, characterized by lower reverberation times, were regarded by both groups as less favorable. Professional singers noted that these spaces required considerable additional vocal effort to achieve the desired projection and expression, while amateur singers reported a lack of sufficient acoustic support, which made it difficult to hear themselves and led to a more uncomfortable singing experience. An unexpected discrepancy emerged with regard to Church 4. As detailed in Sec. 3.C, users of the ASL convolution system reported audio artifacts and effects that interfered with their performance, ultimately resulting in lower preference ratings. In contrast, professional singers who performed in the real Church 4 described a pleasant enhancement of harmonics and rated the space as equally suitable for performance as Church 2. These discrepancies may be attributed to stronger early reflections caused by the narrower nave and the proximity of large pillars near the reference position in Church 4, as previously discussed in this paper. However, further investigation of this aspect remains necessary.

5. CONCLUSIONS AND FUTURE PERSPECTIVES

The objective of the present study was to assess the feasibility of a real-time convolution system for the vocalization and auralization of solo singing within the 3rd Order Ambisonics loudspeaker array of the Audio Space Lab (Politecnico di Torino, Italy). Four auralized environments were created by processing the impulse responses of four modern churches. Reverberation-related acoustic parameters were extracted in both real and virtual conditions, revealing differences comparable to the Just Noticeable Differences (JND). A subjective evaluation was conducted in the ASL to assess the perceptual qualities of the reconstructed environments. The results of the present study are summarized in the following list:

- The amateur singers expressed preference rankings that largely aligned with those of professional singers who had performed in the real churches—except for Church 4, where discrepancies were observed, likely attributable to stronger early reflections;
- Regarding singers' adaptation to different acoustical conditions, Time Dose reflected trends consistent with reverberation metrics across churches, while other parameters—such as Singer Power Ratio and Smoothed Cepstral Peak Prominence—did not exhibit significant correlations with acoustic features;
- The findings demonstrate that the Audio Space Lab provides a viable environment for the vocalization and auralization of singing.

Future research should include direct comparisons of professional singers' experiences in both real and auralized environments. Measurements in the actual churches using a dummy head could help validate the acoustic similarity between real and virtual settings. Moreover, additional room acoustic parameters—such as the Direct-to-Reverberant Ratio—and further voice parameters (e.g., QR, tempo) should be considered to more precisely investigate the discrepancies observed between real churches and their ASL counterparts.

ACKNOWLEDGMENTS

The authors thank Faber Teater and Antonella Talamonti for their collaboration and fruitful exchange; the hosts of the performances in each church; Irene Tartaglione, Farid Jafari and William Calegari Bohnert for their support during the measurements; all the volunteers that participated in the subjective tests.

REFERENCES

[1] N. Eley, S. Mullins, P. Stitt, and B. F. G. Katz, "Virtual notre-dame: Preliminary results of real-time auralization with choir members", in *2021 Immersive and 3D Audio: From Architecture to Automotive (I3DA)*, 1–6 (IEEE) (2021).

[2] S. Mullins and B. F. G. Katz, "Immersive auralisations for choral ensembles", in *Auditorium Acoustics 2023* (Institute of Acoustics, Athens, Greece) (2023).

[3] D. Thery and B. F. G. Katz, "Auditory perception stability evaluation comparing binaural and loudspeaker ambisonic presentations of dynamic virtual concert auralizations", The Journal of the Acoustical Society of America **149**, 246–258 (2021).

[4] J. D. Muynke, N. Eley, J. Ferrando, and B. F. G. Katz, "Preliminary analysis of vocal ensemble performances in real-time historical auralizations of the palais des papes", in *Proceedings of The Acoustics of Ancient Theatres (SAT)*, 1–4 (Verona, Italy) (2022), accessed: 2024-11-03.

[5] B. N. J. Postma and B. F. G. Katz, "Creation and calibration method of acoustical models for historic virtual reality auralizations", Virtual Reality **19**, 161–180 (2015).

[6] P. Luizard and N. H. Bernardoni, "Changes in the voice production of solo singers across concert halls", J. Acoust. Soc. Am. **148**, EL33–EL39 (2020).

[7] P. Luizard, J. Steffens, and S. Weinzierl, "Singing in different rooms: Common or individual adaptation patterns to the acoustic conditions?", J. Acoust. Soc. Am. **147**, EL132–EL137 (2020).

[8] Y. G. Redman, J. D. Glasner, D. D'Orazio, and P. Bottalico, "Singing in different performance spaces: The effect of room acoustics on singers' perception", J. Acoust. Soc. Am. **154**, 2256–2264 (2023).

[9] P. Bottalico, N. Łastowiecka, J. D. Glasner, and Y. G. Redman, "Singing in different performance spaces: The effect of room acoustics on vibrato and pitch inaccuracy", J. Acoust. Soc. Am. **151**, 4131–4139 (2022).

[10] J. Brereton, D. Murphy, and D. Howard, "The virtual singing studio: A loudspeaker-based room acoustics simulation for real-time musical performance", in *Proceedings of the Joint Baltic-Nordic Acoustics Meeting 2012*, 1–8 (Odense, Denmark) (2012).

[11] J. S. Brereton, "Singing in space(s): Singing performance in real and virtual acoustic environments—singers' evaluation, performance analysis and listeners' perception", Ph.D. thesis, University of York (2014).

[12] L. Shtrepi, A. Guastamacchia, M. Masoero, and A. Astolfi, "An investigation on the spatial adaptation of an artistic performance in contemporary churches", in *2023 Immersive and 3D Audio: From Architecture to Automotive (I3DA)*, 1–7 (2023).

[13] International Organization for Standardization, "Acoustics - Measurement of Room Acoustic Parameters. Part 1: Performance Spaces", Standard (2009), iSO 3382-1:2009.

[14] F. Martellotta, E. Cirillo, A. Carbonari, and P. Ricciardi, "Guidelines for acoustical measurements in churches", Applied Acoustics **70**, 378–388 (2008).

[15] International Telecommunication Union - Radiocommunication Sector (ITU-R), "Recommendation itu-r bs.1116-3: Methods for the subjective assessment of small impairments in audio systems", Recommendation BS.1116-3, ITU-R, Geneva, Switzerland (2015).

[16] A. Guastamacchia, R. G. Rosso, G. E. Puglisi, F. Riente, L. Shtrepi, and A. Astolfi, "Real and virtual lecture rooms: Validation of a virtual reality system for the perceptual assessment of room acoustical quality", Acoustics **6**, 933–965 (2024).

[17] D. Pelegrín-García, J. Brunskog, and B. Rasmussen, "Speaker-oriented classroom acoustics design guidelines in the context of current regulations in european countries", Acta Acustica united with Acustica **100**, 1073–1089 (2014).

[18] D. Pelegrín-García, "Comment on "increase in voice level and speaker comfort in lecture rooms"", J. Acoust. Soc. Am. **129**, 1161–1164 (2011).

[19] A. Farina, "X-mcfx: A tool for multichannel acoustic measurements and processing", URL http://www.angelofarina.it/X-MCFX.htm, accessed: 2025-02-21.

[20] A. Guastamacchia, M. Ebri, A. Bottega, E. Armelloni, A. Farina, G. Puglisi, F. Riente, L. Shtrepi, M. Masoero, and A. Astolfi, "Set up and preliminary validation of a small spatial sound reproduction system for clinical purposes", in *Proceedings of the 10th Convention of the European Acoustics Association Forum Acusticum 2023*, 4991–4998 (European Acoustics Association, Turin, Italy) (2024).

[21] Impala Studios, Piano, https://apps.apple.com/it/app/pianoforte-tastiera-virtuale/id1126477246.

[22] R. Spagnolo, *Acustica. Fondamenti e applicazioni* (UTET Università, Torino, Italia) (2015).

[23] A. Kuusinen and T. Lokki, "Wheel of concert hall acoustics", Acta Acustica united with Acustica **103**, 185–188 (2017).

[24] A. Carullo, A. Vallan, and A. Astolfi, "Design issues for a portable vocal analyzer", IEEE Transactions on Instrumentation and Measurement **62**, 1084–1093 (2013).

[25] K. Omori, A. Kacker, L. M. Carroll, W. D. Riley, and S. M. Blaugrund, "Singing power ratio: Quantitative evaluation of singing voice quality", Journal of Voice **10**, 228–235 (1996).

[26] A. Castellana, A. Carullo, S. Corbellini, and A. Astolfi, "Discriminating pathological voice from healthy voice using cepstral peak prominence smoothed distribution in sustained vowel", IEEE Transactions on Instrumentation and Measurement **67**, 646–654 (2018).

[27] C. P. Baker, J. Sundberg, S. C. Purdy, T. O. Rakena, and S. H. D. S. Leão, "CPPS and voice-source parameters: Objective analysis of the singing voice", Journal of Voice **38**, 549–560 (5).

[28] A. Astolfi, A. Carullo, L. Pavese, and G. Puglisi, "Duration of voicing and silence periods of continuous speech in different acoustic environments", J. Acoust. Soc. Am. **137**, 565–579 (2015).