



**HAL**  
open science

## Immersive Auralisations for Choral Ensembles

Sarabeth S. Mullins, Brian F. G. Katz

► **To cite this version:**

Sarabeth S. Mullins, Brian F. G. Katz. Immersive Auralisations for Choral Ensembles. 11th International Conference on Auditorium Acoustics,, Institute of Acoustics, Sep 2023, Athens, Greece. hal-04226690

**HAL Id: hal-04226690**

**<https://hal.science/hal-04226690>**

Submitted on 3 Oct 2023

**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.

# IMMERSIVE AURALISATIONS FOR CHORAL ENSEMBLES

SS Mullins      Institut d'Alembert, Sorbonne University, CNRS, Paris, France  
 BFG Katz        Institut d'Alembert, Sorbonne University, CNRS, Paris, France

## 1 INTRODUCTION

The last twenty years have seen an expanded focus on the interdisciplinary nature of the study of acoustics, in part driven by resolutions passed by UNESCO on the importance of sound in today's world<sup>1</sup> and the inclusion of types of instrument making, music performance, and other cultural expressions in the list of protected practices acknowledged by the Convention for the Safeguarding of Intangible Cultural Heritage.<sup>2</sup> In parallel, developments within the field of acoustics allow for the applications of modern technology to better preserve, study, and recreate the soundscapes and acoustics of culturally significant sites.

The presented work is a synthesis of ongoing research from the realms of architectural acoustics, auralisation, and musical acoustics. It is undertaken as a part of an ongoing archaeoacoustics project focused on the Cathédrale Notre-Dame de Paris, which has played an important role in the development of western European musical traditions. For years, the interconnection between the cathedral and the musical styles that developed there has been a matter of speculation and intrigue for musicologists and cultural historians.<sup>3</sup> To better understand the relationship between the ancient music practices at the cathedral and its contemporaneous acoustics, a choir-focused binaural auralisation system was developed to support ongoing experimentation.

The study of musician performance variation and room acoustics is well-established, with studies broadly following categorisations of soloist/ensemble performances and actual/virtual acoustic spaces (see Table 1). Within the virtual acoustic rendering approach, studies can be further categorised between loudspeaker-reproduction and headphone-reproduction. Although the computational ability to simulate complex acoustic fields has increased in the last decades, there have been limited attempts to auralise choral ensembles within an individualised, spatialised room acoustic simulation. Fischinger *et al.* studied a choir inside an acoustic field rendered over “partially-open” headphones.<sup>4</sup> This study used the same impulse response for the convolution of each individual singer's microphone. Canfield-Dafilou *et al.* created an auralisation room using a system of four loudspeakers to render auralisations for studying student choirs within various virtual acoustic spaces.<sup>5</sup> The recording system was independent from the auralisation system. Yadav *et al.* studied and developed a system for simulating autophonous room impulse responses (RIRs) in geometrical acoustic models and delivering those simulations over open headphones to soloist singers for use in subjective evaluations.<sup>6,7</sup> Jimenez *et al.* extended these techniques to auralise larger orchestral ensembles within the same theoretical framework, though it appears that the larger ensembles again received spatially-generalised acoustics delivered over the reproduction system rather than individualised acoustic responses.<sup>8</sup> Meanwhile, Boren *et al.* created an auralisation system for a large ensemble using spatially averaged room impulses and closed headphones.<sup>9</sup>

System calibration procedures for interactive auralisation systems are found across a variety of research domains. Laird *et al.* calibrated a loudspeaker-based system by comparing the energy ratios from a reference room's impulse response (RIR) with the same measurement taken inside the experimental room recreating the RIR.<sup>17</sup> Yadav *et al.* calibrated their headphone-based system by recreating a measured oral-binaural room impulse response (OBRIR) through the system, using a dummy head

|         | soloist         | ensemble  |
|---------|-----------------|-----------|
| actual  | [10–14]         | [10]      |
| virtual | [7, 13, 15, 16] | [4, 5, 9] |

Table 1: Categorisation of literature on musician performance studies.

with mouth simulator to measure the output of the auralisation and adjusting gains and delays accordingly.<sup>6</sup> Pelegrin Garcia *et al.* calibrated their headphone-based system in a three-step process. First, an OBRIR was measured in a physical space using a dummy head with mouth simulator. Then, the same configuration was modelled in a geometrical acoustic (GA) software. Finally, the output of the GA software was used in a convolution engine and the delay and gain of the reproduced acoustic was adjusted until the physical OBRIR was reproduced by the overall auralisation system.<sup>18</sup> Amengual Gari *et al.* calibrated the level of loudspeaker-based system by placing microphones near the ears of study participants playing music within the experimental room. The recording was convolved with the direct part of the auralisation signal and played through the system. The gains of the original recording and the reproduced direct sound were then compared and an amplitude offset calculated.<sup>19</sup> Sierra-Polanco *et al.* used a headphone-based system and calibrated the microphone level to 94 dB at 1 kHz, but did not attempt to reproduce a particular room as a part of the experiment.<sup>20</sup>

While the existing literature provides many approaches for calibrating auralisation systems over loudspeakers or for individual users over headphones, there is room for the development of a calibration procedure for a multi-user, headphone-based auralisation systems aimed at recreating specific rooms.

## 2 CONVOLUTION SIGNAL

### 2.1 Creation

The impulse responses used here were created with CATT-Acoustic (v9.1) and TUCT (v2.0e:1.02). GA models of the Cathédrale Notre-Dame de Paris were simulated in TUCT using 1 000 000 rays and algorithm 1 with 1<sup>st</sup>-order diffraction enabled.<sup>21</sup>

Within each acoustic model, four positions were defined for autophonus sources, where the receiver was positioned 10 cm on-axis behind the source's location.<sup>6</sup> These were distributed in a shallow arch, spaced approximately 1.5 m apart, facing towards a central point as with a conductor. A fifth receiver location was at this focal point, a little over 1.5 m away from the singing positions defined around it. The sources were defined with the directivity pattern of a soprano singing at fortissimo,<sup>22</sup> and the binaural receivers used the Neumann KU-100 HRTF.<sup>23</sup> All sources and receivers were directed at the focal point. A passive receiver likewise pointed towards the focal point, facing towards the array of autophonus sources. The BRIRs associated with colocated source-receiver pairs will be called OBRIRs moving forward, to distinguish it from the BRIRs traveling between that source and other receivers in the configuration.

When a soloist plays within a chamber, what she hears can be conceptualised as a combined signal of the direct sound from her instrument as well as a reverberant sound shaped by her instruments' directivity, and her location within the architecture of the chamber. Expanding this framework to two performers within the chamber, Musician A perceives a combined signal of her own direct sound, the direct sound of Musician B, and the reverberant sound of herself and Musician B within the chamber. The RIRs which characterise the path from Source A to Receiver B (*A-B*) and its inverse (*B-A*) are not identical, as they contain spatial characteristics related to both positions in the hall. For any ensemble with  $N$  performers, the total combination of sound paths received by any performer can be calculated as  $N^2$ .

### 2.2 Treatment

After generating the O/BRIRs for each source and all receivers, steps were taken to normalise the level of the direct sound across the entire matrix of sources. To achieve this, the OBRIRs (e.g., Source A to Receiver A (*A-A*) or Source B to Receiver B (*B-B*)) were scaled to have a maximum value of 1 at the peak of the highest sound. The scaling factors were preserved for each OBRIR and applied to the other BRIRs originating from the same source (e.g., (*A-B*, *A-C*, *A-D*)) to preserve the proportionate volume of the direct sound as it disperses with distance. TUCT generates all RIRs with a small amount of leading zeroes before the arrival of the direct sound, so this offset was determined for the OBRIRs and then removed from all O/BRIRs to ensure that the arrival time for each impulse was synchronised between the OBRIRs

and the time of arrival was preserved for the subsidiary BRIRs. This results in a matrix of normalised and time-aligned O/BRIRs ( $g_{ref}$ ) which may be further manipulated to allow for real-time convolution.

Within room acoustics, 0 ms to 10 ms is generally accepted as the time-window containing the direct sound within an impulse response.<sup>24</sup> To exclude the direct sound from the convolution engine, the first 10 ms of the OBRIRs were replaced with a vector of zeroes.<sup>17</sup> For the BRIRs (such as *A-B*, *A-C*, *A-D*), the direct sound arrived between 4.4 ms to 8.1 ms after the beginning of the time-aligned BRIRs. In the case of the BRIRs, the 10 ms-long window began at the direct sound's time of arrival. The signals were then cross-checked to ensure that direct sound had been entirely removed from the standard impulse responses (e.g., *A-D* and *D-A*).

After introducing the zero-vector to exclude the direct sound from the impulses, a 1.1 ms (50 samples) long linear gain ramp was applied to smooth the onset of the convolution signals. The resulting impulses ( $g_{aur}$ ) were then loaded into the convolution system for system calibration and validation.

### 3 SYSTEM

The auralisation system was created in Max 8 on a 3.7 GHz Quad-Core 2013 Mac Pro with 16 GB of memory. An external sound card (RME BabyFace Pro) connected to a microphone preamp (RME OctaMic II), which provided power to four head-mounted cardioid microphones (DPA 4088). The patcher ran at 44.1 kHz with an I/O buffer of 64 samples, and used the `multiconvolve` object from the HISSTools toolbox for efficient partitioned convolution.<sup>25</sup> The reverberant signal was output through the sound card to an amplifier that powered four fully transparent head-mounted loudspeakers (AKG K1000).

Dry audio signals were recorded in Reaper, including each singer's head-mounted microphone, a stereo pair of omnidirectional microphones (DPA 4006), and a 1<sup>st</sup>-order Ambisonics microphone (Core Sound TetraMic) positioned in the middle of the singers' array. A separate computer recorded the video output of four Logitech Streamcams mounted at head height to capture close views of their faces and eyes for further body language analysis. Finally, a wide-angle video camera (Zoom Q2n-4K) captured an ensemble view of the participants to better analyse the nonverbal communication and gestures of the ensembles.

Using this setup, each participant received a personalised binaural audio signal delivered over the open headphones. The virtual acoustic environment (VAE) did not reproduce the direct sound of any musician, instead using the unamplified direct sound travelling within the anechoic chamber. Each participant received a personalised, binaural reverberant signal over their headphones, which accounted for both their physical spatial distribution throughout the experimental room as well as the virtual distribution inside each acoustical model (see Figure 3). The computer handled  $16 \times 2 = 32$  simultaneous real-time convolutions to create the complete binaural convolution engine, with a system latency of 5.8 ms.

### 4 CALIBRATION AND VALIDATION

Previous calibration procedures frequently focus on the calibration of a loudspeaker array, where the loudspeakers can be treated as extending the domain of the experimental room into the virtual domain, or on the recreation of an OBRIR using a dummy head with mouth simulator.<sup>18,26</sup>

This calibration process for the binaural system follows an energy-balance approach previously used within the context of loudspeaker-based reproduction systems and extended in Eley *et al.*<sup>17,27</sup> The approach begins with calculating the stage parameter  $ST_{late}$  as in Eq. (1), which measures the balance of early and late acoustic energy related to a performer's experience of reverberance on a stage. Traditionally,  $ST_{late}$  is measured using an omnidirectional source and an omnidirectional receiver located 1 m apart on a stage<sup>24</sup> and does not allow for cases with directional sources, binaural receivers, and co-located source-receiver pairs. To make a distinction between the standard calculation of  $ST_{late}$  and the compound metric used to adjust the full system's energy balance,  $SysT_{late}$  (System Support) is proposed as an alternative term.

The first step of the calibration was the calculation of a reference value of  $ST_{late}$  for each member impulse in  $g_{ref}$ . These two values (one per channel) provide target parameters for a comparison with the auralisation system.

Using a known delay of  $n$ -samples, a delayed dirac signal  $\delta(n)$  and a delayed version of  $g_{aur}^n$  were created to record the performance of the total system.

A co-axial loudspeaker (Genelec 8331B) was positioned 1.5 m away from a KU-100 dummy head and the volume of the headphones were adjusted to output the same level of sound input into the auralisation microphone (DPA-4088). The auralisation microphone was positioned 7 cm in front of the loudspeaker off-axis from the speaker diaphragm and was connected to the auralisation patcher described in Section 3. Using a synchronised sine sweep<sup>28</sup> covering over the frequency range [20 Hz-22 kHz] and at sampling rate 44.1 kHz, recordings were made of the direct sound at the proximate microphone as well as the dummy head. Having calculated the level offset between the three microphones, a pair of headphones were placed on the dummy head and the auralisation system was turned on. Since the calibration signals include a known delay, the dummy head microphones recorded two sweeps for each system measurement: the first coming directly from the speaker, and the second one delayed through the convolution engine. Using the auralisation system with signal  $\delta(n)$ , the direct and delayed sweeps were recorded and the headphone output was adjusted so that the RMS amplitude of the delayed sweep recorded at the dummy head was equal to the RMS amplitude of the dry sweep recorded with the proximate microphone. By calculating the offset in samples from the onset of the direct sound recorded with the auralisation microphone and the onset of the second sweep recorded by the dummy head, it is possible to account for the total system latency from microphone input to auralised output. Using this offset, an amount of zeroes at the beginning of all  $g_{aur}$  signals corresponding to the total system delay were removed to offset the latency introduced by the convolution engine (in this case, 256 samples). Using the delayed OBRIR  $g_{aur}^n$ , the system gains within Max 8 were adjusted to reproduce the  $ST_{late}$  previously calculated from  $g_{ref}$ , with an adjusted formulation of the calculation (see Eq. (2)).

$$ST_{late} = 10 \log_{10} \left( \frac{\int_{100ms}^{1000ms} h(t)^2 dt}{\int_{0ms}^{10ms} h(t)^2 dt} \right) \quad (1) \quad SysT_{late} = 10 \log_{10} \left( \frac{\int_{100ms}^{1000ms} h_1(t)^2 dt}{\int_{0ms}^{10ms} h_0(t)^2 dt} \right) \quad (2)$$

For the combined term  $SysT_{late}$ ,  $h_0$  is the direct signal received from the DPA 4088 positioned near the loudspeaker, while  $h_1$  is taken from the dummy head's recording of the second, delayed sweep. After adjusting the gains within Max 8,  $g_{aur}^n$  was replaced by the final signal  $g_{aur}$  and the system was again recorded with a sweep for final verification that  $SysT_{late}(g_{aur}) \equiv ST_{late}(g_{ref})$ .

Once the overall volume of the headphones, internal ratios within Max 8, and latency offsets are calculated, the system can be adjusted by varying only the software gains proportionally to each other, preserving the previously calculated energy ratio. By scaling all OBRIRs and BRIRs by the same factor, it is also possible to verify that the gains set for each OBRIR impulse reproduce the same calibration for the BRIRs as well.

As a final verification, the experimental subjects were asked to participate in a final step. In this final step, the singers were fitted with their head-mounted microphones approximately 7 cm from the center of their mouths and wearing headphones. Standing 1.5 m in front of a sound level meter, they were asked to sing a sustained note at 84 dB for a period of 5 s, reproducing the level of the co-axial speaker used in the calibration steps. The RMS level of the recorded sound from the headmounted microphone was then compared to the levels recorded during the calibration, and the software gains in Max 8 were then adjusted to account for level differences due to microphone placement. These adjustments tended to be level decreases ranging from 2 dB to 4 dB below the initial settings, suggesting that this energy based calculation provides a suitable starting point for future binaural calibrations.

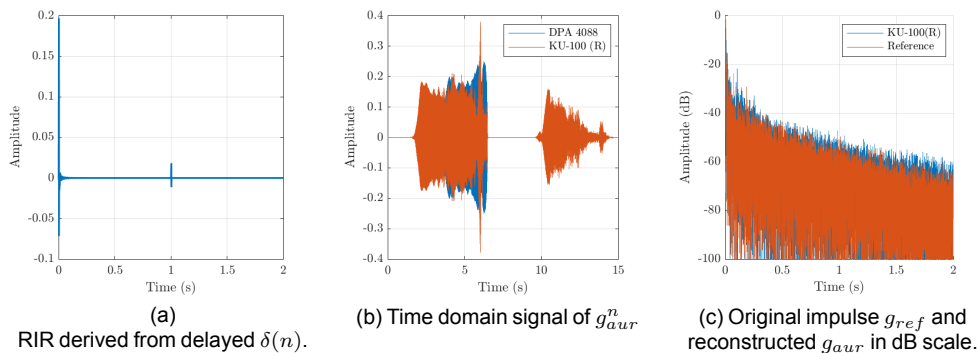


Figure 1: Steps in the calibration process. Final energy balances after calibration are  $ST_{late}(ref) = -29.8$  dB and  $SysT_{late} = -30.4$  dB.

## 5 EXPERIMENTAL APPLICATION

Following the calibration process, two professional choirs participated in a set of singing experiments within the system. The experiments included four test sessions, in which the ensemble would perform 1 min excerpts of three medieval liturgical songs within a five reproduced acoustics of the the Cathédrale Notre-Dame de Paris. Each musical excerpt was representative of a different era of medieval music and corresponded to one of the reproduced acoustics. Members of the ensembles have experience singing inside Notre-Dame de Paris and other significant venues associated with the development of medieval music.

Following the description in Section 2, four performer locations and one listener location were defined in similar positions within acoustical models. For each autophonous source,  $g_{ref}$  and  $g_{aur}$  were generated and then treated as described in Section 2.2. The source and receiver locations within the model replicated the arch the singers standing in within the anechoic chamber. This is slightly wider apart than the preferences for choir spacing established in Daugherty, but still within the bounds of reasonable spacing for choral performances.<sup>29</sup> This spacing was chosen to balance proximity between the singers and the requirement of cross-talk rejection for later performance analysis.

In each acoustic, the ensembles performed each song before filling out a short questionnaire about their assessment of the acoustics. The ensembles were asked to rate the ease of performance within the acoustic per song, as well as the suitability of that acoustic for the song. Finally, they were asked to rate the overall difficulty of performing within the acoustic. At the end of the experiment, Ensemble A and Ensemble B members were each presented with floorplans and elevations of the buildings under study. The ensembles were immersed in each acoustic for 3 min and asked to identify the floorplan of the current acoustic.

Following the experiments with the ensembles, 17 participants were recruited to assess the same auralisation system as the choirs used. This was done to cross-check the calibration system's validity and overall impression of the auralisation system. They were administered a similar subjective listening test and asked to assess the acoustics for the amount of reverberation, the suitability of the acoustics as a music performance place, and to make a guess about which acoustic they were immersed in. Unlike the experiments with the choirs, these participants were given a set of six floorplans to guess from. Analysis of the musical performance is still ongoing.

## 6 DISCUSSION

The calibration was successful in suggesting initial levels for auralisation input and output level, but room remains for further refinements. Due to the amount of time it takes to verify the calibrations for

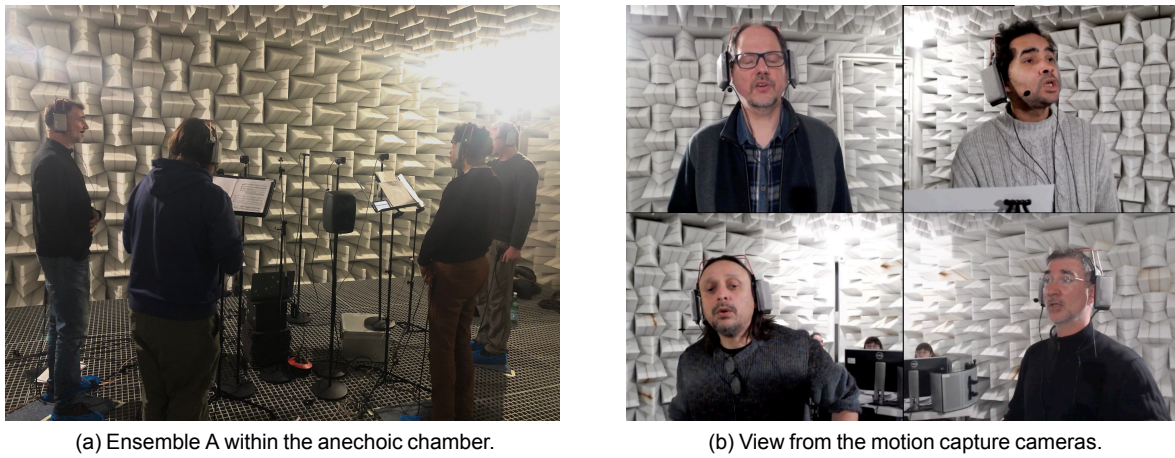


Figure 2: Views of Ensemble A in the final auralisation setup.

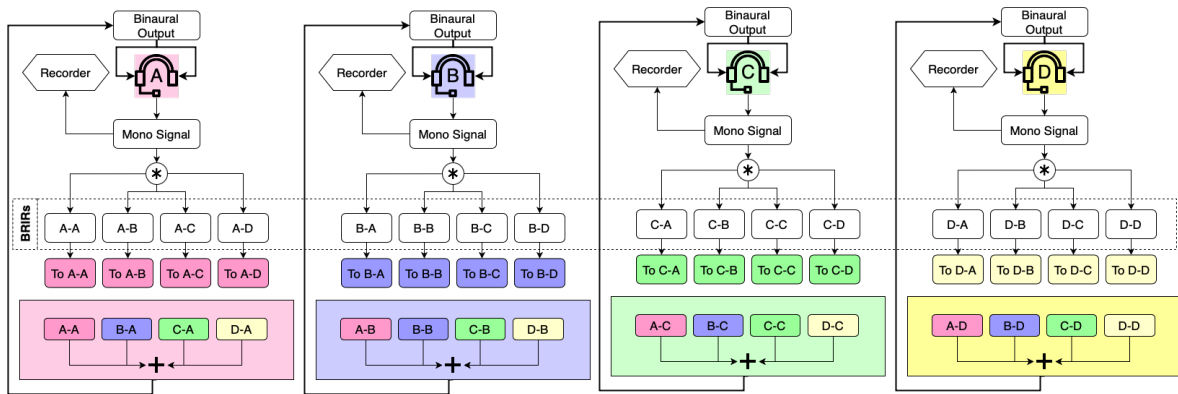


Figure 3: Signal flow of complete choral auralisation system.

$4 \times 4 \times M_{acoustics} = 16 * N$  BRIRs in a four-person choir, it would be preferable to expedite the calibration procedure to allow for concurrent cross-checking of self-to-other levels through the system.

As noted earlier, the adjustments made to the auralisation system during the final calibration step tended to reduce the output volume of the headphones by 2 dB to 4 dB. While this is a satisfactorily small amount of adjustment from the initial calibration level, it points towards a systemic over-estimation of the perceived level of the direct, autophonous sound for each performer. This is likely due to a combination of factors, including the directivity of the head-mounted microphones and their sensitivity to placement and distance, as well as the directivity pattern of the loudspeaker used in the calibration. Finally, the calibration does not account for the bone-conducted sound for the individual performers, which would contribute an additional volume offset from the direct sound to the reverberation in the system. However, after making the final hardware gain adjustments, the members of Ensemble A and Ensemble B noted that they experienced the same difficulties in performing “as in a cathedral,” indicating that the system was set up in a plausible way for their performance.

The responses of the 17 non-specialist assessors likewise indicated an overall amount of plausibility in the calibrated system, with 5 assessors requesting adjustments to the gain averaging 3 dB to 4 dB removed from the original setup. However, unlike the professional choirs used in the experiments, the non-specialists did not uniformly desire a lower level for the auralisations. This indicates that non-specialists may have a different taste profile from musicians<sup>30</sup>, or that non-specialists expect a sense of hyper-reality from the virtual acoustic reality which professional choirs do not.

## 7 CONCLUSION

The calibration procedure used for these experiments was a useful if time-consuming process, ensuring that the properties of a reproduced acoustic are accurately reflected within the signal received at a listener's ear. However, further work is needed to refine the system, including accounting for hardware directivities and a streamlined process to verify the calibration across non-autophonous sound paths. Despite this, the calibration process created a system that both professional choirs and non-professional assessors acknowledged as plausible, and that reproduced key aspects of particular room acoustic for experimental purposes.

## ACKNOWLEDGEMENTS

The authors would like to thank the researchers who have contributed to the development of the auralisation framework over the years and experimentation protocol over the years: Peter Stitt, Nolan Eley, and Elliot K. Canfield-Dafilou. Funding has been provided by the European Union's Joint Programming Initiative on Cultural Heritage project PHE (The Past Has Ears, Grant No. 20-JPIC-0002-FS), the French project PHEND (The Past Has Ears at Notre-Dame, Grant No. ANR-20-CE38-0014), and the Chantier Scientifique CNRS/MC Notre-Dame.

## REFERENCES

1. *The Importance of Sound in Today's World: Promoting Best Practices* Paris, France. <https://unesdoc.unesco.org/ark:/48223/pf0000259172>.
2. *UNESCO - The Convention for the Safeguarding of the Intangible Cultural Heritage* <https://ich.unesco.org/en/convention> (2023).
3. Wright, C. *Music and Ceremony at Notre Dame of Paris 500-1550* 400 pp. ISBN: 0-521-24492-7 (Cambridge University Press, Cambridge, UK, 1989).
4. Fischinger, T., Frieler, K. & Louhivuori, J. Influence of Virtual Room Acoustics on Choir Singing. *Psychomusicology: Music, Mind, and Brain* **25**, 208–218 (Sept. 1, 2015).
5. Canfield-Dafilou, E. K., Callery, E. F., Abel, J. S. & Berger, J. J. *A Method for Studying Interactions between Music Performance and Rooms with Real-Time Virtual Acoustics* in *Proc 146th Conv* (Aud Eng Soc, Dublin, Ireland, 2019), 10.
6. Yadav, M., Cabrera, D. & Martens, W. L. A System for Simulating Room Acoustical Environments for One's Own Voice. *Applied Acoustics* **73**, 409–414 (Apr. 2012).
7. Miranda Jofre, L. A., Cabrera, D. A., Yadav, M., Sygulska, A. & Martens, W. L. *Evaluation of Stage Acoustics Preference for a Singer Using Oral-Binaural Room Impulse Responses* in *J. Acous. Soc. Am.* **133** (May 2013), 3402–3402. (2021).
8. Jimenez, D., Vandenberg, N., Miranda Jofre, L., Yadav, M. & Cabrera, D. *Auralisation of Stage Acoustics for Large Ensembles* in *Proc Acoustics 2013* (Jan. 1, 2013), 5.
9. Boren, B. *et al. Acoustic Simulation of Bach's Performing Forces in the Thomaskirche* in *Proc EAA Spatial Audio Signal Processing Symp* (EAA Spatial Audio Signal Processing Symp, Paris, France, 2019), 6.
10. Chiang, W., Chen, S.-t. & Huang, C.-t. Subjective Assessment of Stage Acoustics for Solo and Chamber Music Performances. *Acta Acustica united with Acustica* **89**, 848–856 (2003).
11. Bolzinger, S., Warusfel, O. & Kahle, E. A Study of the Influence of Room Acoustics on Piano Performance. *Le Journal de Physique IV* **04**, 617–620. ISSN: 1155-4339 (May 1994).
12. Luizard, P., Brauer, E., Weinzierl, S. & Bernardoni, N. H. *How Singers Adapt to Room Acoustical Conditions* in *Proc. Inst of Acous Auditorium Acoustics*. **40** (Institute of Acoustics, Hamburg, Germany, 2018).
13. Luizard, P., Weinzierl, S., Steffens, J. & Brauer, E. *Adaptation of Singers to Physical and Virtual Room Acoustics* in *Proc Intl Symp on Room Acoustics* (Amsterdam, Netherlands, Sept. 2019).
14. Luizard, P. & Henrich Bernardoni, N. Changes in the Voice Production of Solo Singers across Concert Halls. *J. Acous. Soc. Am.* **148**. (2021) (July 2020).



15. Ueno, K., Kato, K. & Kawai, K. Effect of Room Acoustics on Musicians' Performance. Part I: Experimental Investigation with a Conceptual Model. *Acta Acustica united with Acustica* **96**, 505–515. ISSN: 1610-1928. (2023) (May 1, 2010).
16. Amengual Garí, S., Kob, M. & Lokki, T. *Analysis of Trumpet Performance Adjustments Due to Room Acoustics* in *Proc Intl Symp on Room Acoustics* (Amsterdam, Netherlands, 2019). [www.researchgate.net/publication/335688255](http://www.researchgate.net/publication/335688255) (2021).
17. Laird, I., Murphy, D. T. & Chapman, P. *Energy-Based Calibration of Virtual Performance Systems* in *Proceedings of DAFx-12 15th Intl Conf on Dig Audio Effects* (York, UK, Sept. 2012), 163–166.
18. Pelegrin Garcia, D., Rychtarikova, M., Glorieux, C. & Katz, B. F. *Interactive Auralization of Self-Generated Oral Sounds in Virtual Acoustic Environments for Research in Human Echolocation* in *Proceedings of Forum Acusticum* (Krakow, 2014).
19. Amengual Gari, S. V., Eddy, D., Kob, M. & Lokki, T. *Real-Time Auralization of Room Acoustics for the Study of Live Music Performance* in. DAGA (Aachen, 2016), 1474–1477.
20. Sierra-Polanco, T., Cantor-Cutiva, L. C., Hunter, E. J. & Bottalico, P. Changes of Voice Production in Artificial Acoustic Environments. *Frontiers in Built Environment* **7**. ISSN: 2297-3362 (2021).
21. Dalenbäck, B.-I. *CATT-Acoustic v9.1 with TUCT v2.0* Feb. 2016.
22. Shabtai, N. R., Behler, G., Vorländer, M. & Weinzierl, S. Generation and Analysis of an Acoustic Radiation Pattern Database for Forty-One Musical Instruments. *J. Acous. Soc. Am.* **141**, 1246–1256. ISSN: 0001-4966 (2017).
23. Andreopoulou, A., Begault, D. R. & Katz, B. F. G. Inter-Laboratory Round Robin HRTF Measurement Comparison. *IEEE J Selected Topics in Signal Processing* **9**, 895–906 (2015).
24. *ISO 3382-1(E) Acoustics — Measurement of room acoustic parameters — Part 1: Performance spaces* Standard (Intl Organization for Standardization, 2009).
25. Tremblay, P. A. & Harker, A. The HISSTools Impulse Response Toolbox: 38th International Computer Music Conference. *ICMC 2012* (eds Marolt, M., Kaltenbrunner, M. & Ciglar, M.) 148–155. ISSN: 9780984527410 (July 1, 2012).
26. Garí, S. V. A. *Investigations on the Influence of Acoustics on Live Music Performance Using Virtual Acoustic Methods* PhD thesis (Detmold University of Music, Germany, Sept. 2017).
27. Eley, N., Mullins, S., Stitt, P. & Katz, B. F. *Virtual Notre-Dame: Preliminary Results of Real-Time Auralization with Choir Members* in *Proc I3DA* (Bologna, Sept. 8, 2021), 1–6.
28. Novak, A., Lotton, P. & Simon, L. Synchronized Swept-Sine: Theory, Application, and Implementation. *J Aud Eng Soc* **63**, 786–798. ISSN: 15494950 (Nov. 2015).
29. Daugherty, J. F. Spacing, Formation, and Choral Sound: Preferences and Perceptions of Auditors and Choristers. *J Research in Music Education* **47**, 224–238. ISSN: 0022-4294. eprint: 3345781 (1999).
30. Lokki, T., Pätynen, J., Kuusinen, A. & Tervo, S. Disentangling Preference Ratings of Concert Hall Acoustics Using Subjective Sensory Profiles. *J. Acous. Soc. Am.* **132**, 3148–3161 (2012).