Preliminary analysis of vocal ensemble performances in real-time historical auralizations of the Palais des Papes

Julien De Muynke[1,2]; Nolan Eley[1,3]; Julien Ferrando[4]; Brian F.G. Katz[1]

[1] Institut Jean le Rond ∂'Alembert, Sorbonne Université/CNRS, France, [julien.de_muynke, nolan.eley, brian.katz]@sorbonne-universite.fr
[2] Eurecat, Centre Tecnològic de Catalunya, Multimedia Technologies Group, Spain
[3] ETIS Laboratory, CY Cergy Paris University, ENSEA, CNRS, France
[4] AMU-CNRS, PRISM, UMR 7061, Marseille, France, julien.ferrando@univ-amu.fr

ABSTRACT

In the middle of the 14th century, the recently constructed Great Clementine Chapel of the Palais des Papes had a flourishing reputation for the composition and interpretation of polyphonic singing in the emerging Ars Nova musical style. In modern times, the space is still employed for musical performances. However, the acoustic conditions between the two periods vary greatly, and as such, can be expected to have an impact on vocal performances. As part of the IMAPI and PHE projects, the impact of the acoustics of the Great Clementine Chapel on the performance of a conducted vocal ensemble specializing in medieval music was examined for these two periods. A numerical simulation of the medieval acoustics was developed, based on a calibrated geometrical acoustics model of the modern-day chapel which was then regressed in time to a historically informed medieval state. Experiments were carried out with singers performing repetitions of several pieces in a Virtual Acoustic Environment (VAE) using close-mics and headphone renderings. Recorded performances were analyzed using various metrics, with objective results paired with questionnaires acquired for each VAE condition. Preliminary analysis of these results is presented in this study.
Keywords: cultural heritage acoustics, virtual acoustic environment, vocal ensemble performance

1. INTRODUCTION

1.1 The Palais des Papes in the 14th century

In 1309, the Holy See relocated to Avignon, France, where Pope Clement V established his residence, remaining there until 1403 [1]. Construction of the Palais des Papes, which is the largest Gothic edifice ever built, started in 1335 under the pontificate of Benoît XII, was continued by Pope Clement VI from 1342, and was completed in 1352. Masses accompanied by music performance, especially polyphonic singing, were usually performed in the Great Clementine Chapel (a.k.a. Great Chapel). The Great Chapel attracted music composers, cantors, and musicians, particularly those belonging to the movement known today as the Ars Nova style. Ars Nova is a polyphonic musical style that developed in France in the 14th century as the successor of the Ars Antiqua exemplified by the School of Notre-Dame. It allowed for a higher degree of musical expressiveness and for more elaborate rhythmic modes due to a new standardized system of musical notation, even though some studies have shown that interpreting Ars Antiqua and Ars Nova as two radically different styles is probably excessive [2].

1.2 The impact of room acoustics on musical performance

Practitioners of choral music have long been aware that room acoustics play a significant role in musical performance [3]. However, despite this awareness, there has been no unified approach or theory to guide performance practice in response to different acoustic environments. In fact, while empirical studies have shown measurable effects on musical performance as a result of changes in acoustics, these effects tend to be rather small and/or individual [4, 5]. In short, it is still not well known precisely how room acoustics affect musical performance, and the evidence within the context of historical music is even less sufficient.

This study aims at assessing the impact of room acoustics on the musical performance of the conducted vocal ensemble *Diabolus in Musica*, consisting of three male vocalists (one baritone and two tenors) trained in medieval performance methods and with familiarity singing in the modern-day Great Chapel of the Palais des Papes.

2. VIRTUAL ACOUSTICS OF PALAIS DES PAPES

The acoustics of the Great Chapel of the Palais des Papes in two historical states, namely medieval (ca. 1362) and modern (ca. 2020) states, were generated through geometrical acoustic (GA) models designed in CATT-Acoustic v9.1e.

For that, the geometry of the room was first designed with reference to a 3D laser scan point cloud. Then, the definition of the construction materials of the modern-day room was carried out through an acoustic calibration procedure [6], in which the acoustic properties of the materials were adjusted until various acoustic metrics fell within the range of ± JND of the measured acoustic values, based on recorded room impulse responses.

Next, a GA model of the medieval state of the Great Chapel was obtained by carrying out a time regression of the

peer reviewed

10.58874/SAAT.2022.204

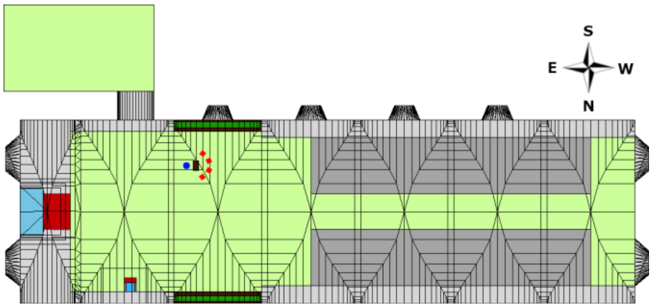

Figure 1 – Top-down view of the modeled Great Chapel in the medieval state. The various materials are represented by different colors. ■: singers; ●: conductor; ⌷: lectern.

modern-day GA model with the help of historical records of the interior furnishing and decoration from that time era [7]. The presence of absorbing materials such as wall tapestries, floor rush matting, stalls, pews, the pope's throne, cloth on top of the altar and the stalls, and canopies above the altar and the pope's throne (represented by different colors in Fig. 1) make the medieval Great Chapel significantly less reverberant than the modern-day Great Chapel which is, essentially, a large empty shoe box made of dense limestone. A comparison of the reverberation time between the two historical states of the Great Chapel is shown in Fig. 2, including the presence of a simulated audience which tends to reduce the reverberation time.

3. SINGING EXPERIMENTS

The experiments were aimed at assessing the impact of the acoustics of the Great Chapel under different conditions on the musical performance of the vocal ensemble *Diabolus in Musica*. It was therefore important to provide the singers with performance conditions that were as close as possible to a real concert situation, while allowing for an audio recording quality sufficient to be used for further objective analysis (see Section 4). The experimental setup used in this study was guided by findings in [8].

3.1 Hardware setup

The singing experiments were carried out in a hemi-anechoic room at the PRISM laboratory in order to reduce interference of the recording room as much as possible. Singers were each equipped with a close microphone (head-band cardioid microphone, DPA4088) in order to reduce the level of inter-singer crosstalk while having them distributed close together in a usual concert configuration. Virtual Acoustic Environments (VAEs) were reproduced for each individual singer over open-back headphones (Sennheiser HD650). Open-back headphones were chosen as they allow the direct sound from one's own voice and from other singers to pass through relatively unobstructed, while the reverberated voice sound is reproduced inside the headphones. The ratio between direct and reverberant sound levels was adjusted prior to the experiments through a calibration procedure to achieve realistic balance [9].

3.2 Auralization system

The VAEs were auralized via convolution with pre-rendered Binaural Room Impulse Responses (BRIRs) from the GA models with the direct sound removed. Each singer playing the role of a source and a receiver at the same time, the reproduction system included a total of $3 \times 3 = 9$ BRIRs for 3 singers, to which 3 extra BRIRs intended for the conductor were added, for a total of 12 BRIRs (i.e. 24 convolutions). The computational cost of such a reproduction system is of concern, especially considering that the BRIRs in the modern-day Great Chapel are 10s long in accordance with the reverberation time in the lowest octave band (see Fig. 2). The auralization architecture was created in MaxMSP to facilitate real-time processing. The convolution was done using the object `multiconvolve` from the HISS Impulse Response Toolbox[1] which employs a fixed partitioning scheme. The system was configured with an internal audio buffer size of 64 samples at 48 kHz, corresponding to an I/O delay of 1.3ms, without artifacts. This delay was compensated for by removing the 64 leading zeros in the BRIRs, providing correct time synchronization between the natural direct sound and the virtual reverberated sound.

3.3 The VAEs

The VAEs used in this study differ in their historical state, namely medieval (ca. 1362, when the Great Chapel was actively used for papal masses) and modern-day (the Great Chapel is still used as a performance space for concerts of vocal ensembles). The choir was positioned in the third bay starting from the east (the *parcus cantorum*, which included stalls in the medieval era), halfway between the chapel symmetry axis and the southern wall. The singers were distributed along an arc spanning 90° centered on the position of a virtual lectern and at a distance of 1.2 m. Sources were simulated with the singing voice directivity pattern `singer.SD1` from CATT-Acoustic, pointing at the

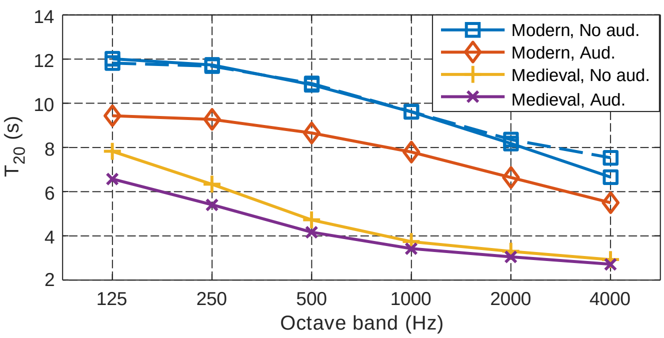

Figure 2 – Estimated T_{20} as a function of frequency band, averaged across sixteen positions evenly distributed along the Great Chapel. For the modern, unoccupied state, the measured (– –) and modeled (—) T_{20} s are shown.

[1] http://eprints.hud.ac.uk/id/eprint/14897/

virtual lectern with the conductor facing them. The positions of the singers and conductor are shown in Fig. 1.

3.4 Experiment protocol

The repertoire comprised two pieces of polyphonic music from the Ars Nova style: "Petre Clemens" by Philippe de Vitry, and "Kyrie Rex Angelorum" (anonymous). The recordings were organized in separate sessions, each rendering either the medieval or modern-day room. In each session, the music pieces were interleaved and repeated 3 times. To compensate for a slightly different positioning of each individual singer's microphone, each microphone gain was adjusted at the beginning of each session to ensure a good consistency between individual singers' voice levels in the overall rendered audio scene. After each session, the participants answered a questionnaire on their subjective experience regarding the simulated acoustics and their performance in that particular space.

4. MUSIC PERFORMANCE ANALYSIS

Features were extracted from recordings of the performances which can be broken down into four musical categories: timing, dynamics, timbre, and pitch.

To represent timing, the note-level tempo was calculated by taking the inverse of the time interval between the onsets of adjacent notes weighted by the written note duration. The note onsets were obtained by manual annotation of one performance followed by audio-to-audio alignment using the `Match` plug-in[2] in Sonic Visualizer[3] followed by manual verification and adjustment.

A-weighted RMS was chosen to serve as a measure of musical dynamics or loudness. As a simplified measure of timbre, the spectral centroid was calculated. The spectral centroid represents the center of gravity of the spectrum and has been shown to be strongly correlated with the perception of a signal's "brightness" [10]. Both the spectral centroid and A-weighted RMS were extracted as time-series vectors with a window size of 2048 samples and a hop size of 10 ms. These vectors were later shortened to a grid of 8th note durations utilizing the note onset information necessary for calculating tempo. The 8th note segments corresponding to rests in the score were removed prior to analysis.

The fundamental frequency of each singer's performance, which was necessary to calculate higher level pitch-related features, was extracted using the `pYin` algorithm [11] in Sonic Visualizer. Because vibrato is a common expressive tool for singers, both the vibrato rate (mean pitch variation rate, in Hz) and vibrato extent (mean absolute distance to the note pitch center, in cents) were calculated on all notes with a duration ≥ a dotted-quarter.

[2] www.eecs.qmul.ac.uk/~simond/match/index.html
[3] www.sonicvisualiser.org

Other researchers have used pitch drift, or the amount the pitch center changes throughout the course of the piece, as an indicator of overall ensemble intonation [5]. However, some amount of pitch drift is normal and may simply be the result of an unaccompanied ensemble singing in a non-equal temperament [12]. So, rather than using pitch drift as a measure of intonation, normalized pitch error was used. Normalized pitch error describes individual note intonation compared to its nominal pitch adjusted for overall pitch drift; a slight modification of the methodology outlined in [5].

5. RESULTS

As a preliminary analysis of the data, box plots were produced for each feature to examine whether or not there was a significant difference between the two acoustic settings. No significant differences were found with the exception of the loudness feature which indicated that each singer sang louder overall in the modern acoustics (see Fig. 3), however, the average difference was only 1.2 dB± 0.2 dB in "Kyrie Rex Angelorum" and 0.8 dB $\pm$ 0.5 dB in "Petre Clemens". A Friedman test with singers as blocks and acoustics as group variable showed that these differences were statistically significant ($p < 0.001$ for both pieces). This greater vocal effort may be partly as compensation for the more reverberant nature of the modern acoustics, however, given that the difference is so small, too much emphasis should not be put on this finding at this time.

Rating questionnaires were given to the participants which asked about the following categories: reverberation, ease of ensemble singing, sound support, quality of the space, and size. No broad consensus was reached in any of these categories with the exception of reverberation, in which all the participants correctly ascertained that the modern state was more reverberant than the medieval state.

In addition to the questionnaire, participants were also encouraged to provide commentary freely which indicated some preferences. Two participants reported that the acous-

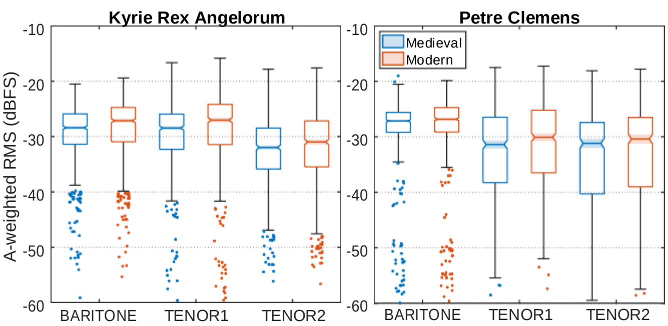

Figure 3 – Box plots of loudness measure in medieval (blue) and modern (red) acoustics by voice, for "Kyrie Rex Angelorum" (left) and "Petre Clemens" (right). Whiskers extend from upper & lower quartiles to non-outlier maxima and minima.

tics of the medieval state were more satisfactory than that of the modern state to the interpretation of complex polyphonic music. One participant reported that singing in the modern state was less demanding than in the medieval state. In accordance with his previous singing experiences, this individual claimed to sing with less diligence in the modern state than in the medieval state, and that a longer reverberation time is more forgiving of small inaccuracies and defects in a performance, as heard from the audience. He also mentioned that as a listener, he would prefer the medieval state because it was "musically more satisfying". Despite their open design, two participants reported an unwanted "filtering" effect of the headphones , producing experimental conditions somewhat less comfortable than normal situations. Although open headphones let external sounds pass through, they still attenuate high frequencies, coloring the sound of one's voice, and the direct sound of the other singers.

6. CONCLUSION AND FURTHER WORK

In this study, all participants were able to correctly identify the more reverberant VAE indicating some perceptual validation of the auralization. However, there were some complaints about the usage of headphones which leaves room for improvement in future performance auralizations. There was no consensus as to which acoustic setting was optimal for the performance of music in the Ars Nova style.

Timbre, tempo, intonation, and vibrato did not seem to have been significantly affected by the acoustics. There is an indication, however, that the singers sang louder in the more reverberant modern configuration of the Great Chapel, but more data would be needed to strengthen this conclusion as it may be somewhat dependent on the style of the musical composition. Additionally, the impact of the acoustics on ensemble-specific features like inter-singer synchrony and intonation could be of interest in future analysis.

Finally, there are still recordings from this experiment which have not been analyzed, including those of a 4-voice ensemble interpreting a repertoire comprising one piece of monodic Gregorian chant and an additional choir configuration and position in the two acoustics. Analysis of these recordings could strengthen some of the preliminary findings of this paper as well as shed light on additional trends which may be a function of these additional variables.

ACKNOWLEDGEMENTS

Funding has been provided by UMR PRISM and Avignon Tourisme project IMAPI, the European Union's Joint Programming Initiative on Cultural Heritage project PHE (phe.pasthasears.eu), and the Paris Seine Graduate School Humanities, Creation, Heritage, Investissement d'Avenir ANR-17-EURE-0021 – Foundation for Cultural Heritage Sciences.

REFERENCES

[1] J.-M. Poisson, "Le palais des Papes d'Avignon : structures défensives et références symboliques," in *Les palais dans la ville*. Presses universitaires de Lyon, 2004, pp. 213–228.

[2] D. Tanay, "The transition from the ars antiqua to the ars nova: Evolution or revolution?" *Musica Disciplina*, vol. 46, pp. 79–104, 1992.

[3] K. Schiltz, "Church and chamber: the influence of acoustics on musical composition and performance," *Early Music*, pp. 64–80, Feb. 2003.

[4] P. Luizard, E. Brauer, and S. Weinzierl, "Singing in physical and virtual environments: how performers adapt to room acoustical conditions," in *Proc AES Conf: Immersive and Interactive Audio*, York, UK, Mar. 2019.

[5] T. Fischinger, K. Frieler, and J. Louhivuori, "Influence of virtual room acoustics on choir singing." *Psychomusicology: Music, Mind, and Brain*, vol. 25, no. 3, pp. 208–218, 2015.

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

[7] G. Colombe, "Au Palais des Papes : la chapelle Clémentine vue de l'intérieur," in *Mémoires de l'Académie de Vaucluse*. Forgotten Books, 1935, vol. XXXV, pp. 79–95.

[8] S. Mullins, V. Le Page, J. De Muynke, E. K. Canfield-Dafilou, F. Billiet, and B. F. G. Katz, "Preliminary report on the effect of room acoustics on choral performance in Notre-Dame and its pre-Gothic predecessor," *Meeting Acous. Soc. Am., J Acoust. Am.*, 2021.

[9] 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)*, 2021.

[10] P. Iverson and C. L. Krumhansl, "Isolating the dynamic attributes of musical timbre," *J Acous Soc Am*, vol. 94, no. 5, pp. 2595–2603, Nov. 1993.

[11] M. Mauch and S. Dixon, "PYIN: A fundamental frequency estimator using probabilistic threshold distributions," in *IEEE Intl Conf on Acoustics, Speech and Signal Processing (ICASSP)*, Florence, Italy, May 2014, pp. 659–663.

[12] D. M. Howard, "Intonation Drift in A Capella Soprano, Alto, Tenor, Bass Quartet Singing With Key Modulation," *Journal of Voice*, vol. 21, no. 3, pp. 300–315, May 2007.