

Sacred Sound and Space: A Cross-Cultural Quantitative Study of Music–Architecture Acoustic Matching

Alina Yuchen Yang

Beijing National Day School, Beijing, China



Abstract

Religious music is rarely experienced in isolation. It's traditionally performed inside specific sacred buildings. This raises a simple question: are the music and the buildings naturally suited to each other acoustically? Scholars have long suggested such an intimate relationship exists but primarily explored through qualitative approaches. In this study, we tried to analyze this relationship quantitatively by estimating mixed-effect regression models. The findings provide quantitative evidence for the co-evolution of sacred sound and space and offer a new approach to studying the acoustic dimension of religious experience.

Keywords: music; architecture; religion

1 Introduction

A slow Gregorian chant sounds very different inside St. Peter's Basilica than in a small wooden temple. Although the resulting sense of sacredness is partly determined by individual perception, it is also influenced by how architectural acoustics interact with musical structure. This acoustic compatibility between music and its venue is an analyzable factor. We argue that religious architecture, from Tibetan monasteries and Japanese temples to European cathedrals, not only shaped by regional cultures and customs but also serve functions that extend far beyond visual solemnity. This leads to a central hypothesis: have religious music and architecture co-evolved such that the acoustic environment enhances, rather than diminishes the music's expressive power? In this paper, we test this hypothesis quantitatively using acoustic data of some representative religious buildings and corresponding music features.

Existing scholarship has consistently pointed to a close interrelationship between religious music and architectural space. The work of Pentcheva (2011) on Hagia Sophia, for instance, demonstrates the aurality of the church is equally as important as the visual experience. Both the aurality and visuality constitute the immersiveness and transcendence of Byzantine religious experience. However, much of the existing research remains qualitative. Few studies have quantitatively investigated the compatibility between musical features and architectural acoustics across diverse religious traditions. This paper seeks to fill this gap by proposing and empirically validating an explanatory hypothesis based on a supply-demand model. We argue that religious music develops a specific 'demand' for particular acoustic environments, while religious architecture, through its long historical evolution, establishes 'supply' constraints that accommodate these demands. This dynamic relationship fosters a state of mutual acoustic matching, enabling

participants to experience a profound integration of music and environment. Therefore, this study asks one central question: Do sacred musical traditions show measurable acoustic adaptation to the buildings where they evolved?

2 Literature review

The relationship between sound and space is central to both architectural acoustics and the study of religion. Religious architecture functions not only as a setting for ritual but also as an acoustic medium shaping auditory experience (Blessner & Salter, 2007). Research shows that structural features such as material and scale significantly influence reverberation and clarity. For example, large stone cathedrals tend to produce long reverberation times, enhancing sustained tones but reducing speech intelligibility, whereas wooden temples favor shorter reverberation and greater clarity (Martellotta, 2016; Soeta et al., 2013; Foteinou et al., 2025). These acoustic characteristics are closely aligned with cultural practices (Di Loreto et al., 2025).

Religious music adapts to these environments over time to help enhancing sacred feelings. For example, Gregorian chant often consists of slow sustained notes, which remain clear inside highly reverberant cathedrals while more text-centered forms require clearer acoustics. Such adaptations contribute to experiences of immersion and sacredness (Pentcheva, 2011; De Muynke et al., 2024). Recent reviews call for more systematic and quantitative approaches to understanding acoustic performance in religious settings (Romel et al., 2025). However, these studies mainly describe the relationship qualitatively. The quantitative relationship is yet to be discovered.

3 Methodology

3.1 Architectural Samples and Acoustic Parameters

While subtle variations in acoustic characteristics may exist within each religious venue type, the overarching acoustic features within specific religious categories exhibit considerable similarity. Consequently, the acoustic differences observed within Tibetan Buddhist monasteries, Japanese temples, and Christian cathedrals are substantially smaller than the disparities between these distinct religious traditions. Because of this characteristic, we selected three representative venues from each tradition. The rationale for selecting these religious venues was based on both their representativeness within their respective traditions and the availability of associated acoustic data.

Table 1.
Representative Venues Selected for Each Religion

Religion	Venue
Christian	Notre-Dame de Paris
	St. Paul's Cathedral
	St. Peter's Basilica
Japanese Buddhism	Nishi Hongwanji Osaka
	Chion-in Kyoto
	Higashi Hongwanji Osaka
Tibetan Buddhism	Potala Palace
	Jokhang Temple
	Labrang Monastery

In the realm of architectural acoustic evaluation, Reverberation Time (RT60, RT60 measures how long it takes for a sound to fade away after it is produced) and Clarity (C80) stand as the two most critical objective indicators. Given that clarity is intimately related to

and significantly influenced by reverberation time, the latter consistently occupies a pivotal position in numerous acoustic investigations. Accordingly, this study primarily focuses on reverberation time. The data sources utilized are categorized into two principal types:

Measured Data: Acoustic parameters for Japanese Buddhist temples (Soeta et al. 2013) and Christian architecture (Martellotta 2016; Foteinou et al. 2025; De Muynke et al. 2024) were primarily sourced from published peer-reviewed literature and public acoustic measurement reports.

Estimated Data: Due to the lack of on-site measured data for Tibetan Buddhist monasteries, we use estimated acoustic parameters based on existing public information, including architectural geometry, volume, surface areas, and absorption coefficients of main materials (Romel et al. 2025).

Table 2.
Acoustic Properties of Selected Architecture

RT60(f) of Different Religious Venues						
Venue	125Hz	250Hz	500Hz	1KHz	2KHz	4KHz
Notre-Dame de Paris	8.7	8.4	7.4	6.1	4.6	3
St. Paul's Cathedral	8.9	10.5	10.7	9.3	7.2	4.4
St.Peter's Basilica	13.6	12.5	10.9	9	7	4.6
Nishi Hongwanji Osaka	1.4	1.5	1.3	1.1	0.9	0.7
Chion-in Kyoto	1.5	1.6	1.4	1.2	1	0.8
Higashi Hongwanji Osaka	2.8	3	2.7	2.2	1.8	1.4
Potala Palace	3.2	3.5	3.8	3.6	3	2.2
Jokhang Temple	2.8	3.2	3.5	3.3	2.8	2
Labrang Monastery	3.5	3.9	4.2	3.8	3	2

Collectively, the RT60 characteristics observed across these three architectural typologies are not coincidental. They represent an acoustic adaptation co-evolved from the interplay of religious ritual demands, spatial scale, building material traditions, and cultural aesthetics. For instance, St. Peter's Basilica demonstrates the longest reverberation time, with low frequencies exhibiting particular prominence (reaching 13.5 seconds at 125Hz). This reflects the grand yet acoustically challenging sound field of colossal stone churches, where speech intelligibility is inherently low. In contrast, the Kyoto Chion-in Great Lecture Hall features the shortest and most uniform reverberation time across all frequencies (approximately 1.3–1.8 seconds), a typical outcome for wooden architecture optimized for speech intelligibility. The Potala Palace, meanwhile, exhibits an intermediate reverberation time that significantly decreases with increasing frequency (from approximately 4 seconds to 2 seconds), indicative of an acoustic design in Tibetan Buddhist halls that judiciously balances chanting resonance with group intelligibility.

3.2 Music Feature Extraction

The compatibility between music and its architectural environment depends on how sound propagates within a space, governed by three primary musical attributes: tempo, centroid, and energy. Tempo measures the speed of music. If a room's reverberation decays too slowly, the overlapping energy of consecutive notes will blur the music's clarity. Centroid measures the brightness of sound. Sounds with greater high-frequency energy have higher centroid values and are particularly sensitive to the absorption of treble

by room materials and air volume. Finally, energy measures the sonic strength. Together, these three attributes define how a musical piece adapts to a space, determining its specific requirements for reverberation time, early decay, and frequency response.

To extract acoustically and musically significant features from audio, this study implemented a standardized processing procedure on 69 religious music recordings (see dataset for list of recordings and features of each music) acquired from publicly available internet sources covering three religion traditions including Tibetan Buddhism, Japanese Buddhism and Christianity.

The sample size was determined based on three considerations. First, the dataset ensures balanced representation across traditions, allowing meaningful cross-cultural comparison while avoiding dominance by any single category. Recordings were thus selected to capture widely recognized canonical or representative forms within each tradition, rather than aiming for exhaustive coverage. Second, the selected recordings cover a range of stylistic and acoustic variation within each tradition (e.g., chanting styles, tempo ranges, and instrumental textures), thereby capturing intra-cultural diversity. Third, given the study's focus on interaction effects rather than high-dimensional prediction, this sample size provides sufficient statistical power to detect systematic patterns of music–architecture matching, as evidenced by the consistency of results across frequency bands. Together, these recordings provide a balanced comparison across traditions.

Table 3.
Music Recording Types

Type	Sub Type	Number
Tibetan Buddhist	Tantric Chant	9
	Mantra	5
	Ritual Instrumental	3
	Throat Singing	2
	Ritual (Mixed)	4
Japanese Buddhist	Shingon Chant	7
	Shomyo	6
	Zen Chant	4
	Sutra / Dharani	3
	Other / Special	2
Christian	Gregorian Chant	9
	Hymn	8
	Renaissance Polyphony	7

Initially, all audio files underwent conversion to mono and resampling to 22050 Hz, a step designed to unify temporal resolution and mitigate computational load. Subsequently, amplitude normalization was applied to obviate interference from recording loudness disparities on energy-related features. These preprocessing steps were executed utilizing the Python Librosa library (McFee et al, 2015). Librosa is a widely used open-source Python library dedicated to music and audio signal analysis. It encapsulates fundamental audio processing algorithms for waveform loading, time-frequency transformation, acoustic feature extraction, audio visualization and signal manipulation.

4 Results and Discussion

Imagine singing a slow Gregorian chant inside St. Peter's Basilica, each note lingers in the air for several seconds, allowing the echoes to blend smoothly into the next note. The result

is a rich, immersive sound that many listeners describe as sacred. Now, imagine performing a rapid chant in the same building. Because each note is still echoing when the next one begins, the music becomes blurred and difficult to distinguish. This simple example illustrates the basic idea of acoustic matching: different kinds of music work best in different acoustic environments.

To understand why some music fits certain buildings better than others, we examined three core musical characteristics: temporal structure (Tempo), spectral centroid (Centroid), and energy structure (Energy). These characteristics influence how music behaves inside reverberant spaces.

How does the model work? Basically, we compared every piece of music with every building. The statistical model then examined whether music matched its own religious buildings better than buildings from other traditions.

We used a mixed-effect regression model. This model helps us determine whether musical characteristics change systematically with reverberation time while accounting for differences among buildings and musical traditions. The regression analysis leverages 69 musical recordings and a full cross-pairing design with nine religious venues, yielding 621 observations.

Given that religious music is influenced not only by architectural acoustic characteristics but also profoundly by cultural factors, we defined a simple variable called cultural match (Match). If music and architecture came from the same religious tradition, Match=1. Otherwise, Match=0. For example, Tibetan Buddhist music paired with Tibetan Buddhist architecture, then Match is assigned a value of 1; otherwise, Match is set to 0, representing a cross-religious pairing (e.g., Tibetan Buddhist music paired with a Christian church or Japanese temple).

Specifically, the estimation model employed in this study is formulated as: $RT60 \sim Match \times (Tempo + Centroid + Energy)$. In simple terms, the model asks three questions.

- Does tempo relate to reverberation?
- Does sound brightness relate to reverberation?
- Does this relationship become stronger when the music and building belong to the same tradition?

The model is designed to explain patterns rather than predict acoustic properties of buildings. This study estimated a series of regression models incorporating interaction terms for each discrete frequency band (ranging from 125 Hz to 4 kHz).

Findings 1: Religious music fits the acoustic environments of its own tradition significantly better than buildings from other traditions.

The coefficient for the matching variable (Match) was positive and highly significant for all frequencies (β ranging from 0.58 to 0.71, $p < 0.001$). This indicates that culturally or religiously matched music-architecture combinations exhibit significantly higher reverberation times. In contrast, the main effects of individual musical features (tempo, centroid, and energy) were generally found to be small and statistically non-significant (all $p > 0.05$), suggesting that musical features, when considered in isolation and without accounting for the matching relationship, do not effectively predict RT60. Table 4 summarizes the statistical evidence supporting this finding.

Table 4.

Mixed Effects Regression Results Across Different Frequencies

Frequency	125 Hz	250 Hz	500 Hz	1 kHz	2 kHz	4 kHz
Intercept	0.21***	0.23***	0.24***	-0.24***	-0.23***	-0.19***
match	0.58***	0.63***	0.68***	0.71***	0.71***	0.68***
tempo	0.07	0.07.	0.07.	0.06	0.06	0.06
centroid	0.08.	0.08.	0.08.	0.08.	0.08.	0.07.
energy	0.05	0.05	0.05	0.05	0.05	0.05
Match:tempo	-0.34**	-0.34**	-0.32**	-0.30**	-0.29**	-0.28*
Match:centroid	-0.30**	-0.30**	-0.28**	-0.27**	-0.26**	-0.25**
Match:energy	-0.11	-0.12	-0.11	-0.11	-0.1	-0.1

Significance levels: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, . $p < 0.1$

Findings 2: Tempo becomes negatively associated with RT60 only when music and architecture belong to the same tradition.

The regression coefficients of Match $\times$ tempo were consistently significant and negative across all frequency bands (β ranging approximately from -0.34 to -0.28 , $p < 0.01$). This relationship means fast music sounds clearer in shorter reverberant spaces whereas, slow chants benefit from longer echoes. Long echoes blur rapid notes.

Findings 3: Music with brighter sounds performs better in spaces with shorter reverberation.

The regression coefficients for Match $\times$ centroid were consistently significant and negative across all frequency bands (β ranging approximately from -0.30 to -0.25 , $p < 0.01$). This relationship was found to be strongest in the low-frequency range. For example, a Gregorian chant characterized by sustained low-register male voices is more compatible with a highly reverberant cathedral. A piece featuring rapid, bright flute passages is better suited to a less reverberant space, where high-frequency details remain intelligible.

Findings 4: Tempo and sound brightness explain most of the matching relationship.

The regression coefficients of Match $\times$ energy were not significant across all frequencies: This suggests that the acoustic demands related to music's energy might already be adequately explained by the interplay of tempo and spectral centroid.

These results collectively indicate that musical features systematically influence reverberation time exclusively under matching conditions. In such matching contexts, faster tempos and higher centroids are associated with shorter RT60s. Conversely, in non-matching contexts, this relationship is either considerably weaker or entirely absent. Furthermore, the high consistency of this interaction effect across different frequency bands suggests that this relationship is not confined to a specific frequency range but rather represents a holistic acoustic characteristic.

Our results support the central hypothesis proposed in the Introduction: Sacred music and sacred architecture exhibit measurable acoustic matching. Under matching conditions, the negative coefficients of the interaction terms further suggest a compensatory relationship. When music is temporally or spectrally denser (e.g., faster tempo or higher

spectral centroid), the adapted space tends to exhibit a shorter reverberation time to ensure auditory clarity. Conversely, sparser music is better suited for spaces with longer reverberation to enhance spatial presence and the sense of sacredness.

In Christian cathedral, listeners may perceive a continuous and enveloping sound field. Extremely long echoes generated from high vaulted ceiling of Christian cathedral make listeners feel music descending from above. In Tibetan monastery, deep chanting becomes fuller because low frequencies are reinforced by reverberation. But because Tibetan monasteries feature spacious interiors yet relatively low ceilings, sound reverberates repeatedly within these spaces horizontally. In other words, listeners in cathedrals will feel more vertical echoes than horizontal echoes, whereas Tibetan monasteries will feel more horizontal echoes than vertical echoes. For Japanese temples, shorter reverberation preserves silence between sounds, reinforcing stillness which is an important cultural heritage in Japan.

5 Conclusion

The central argument of this study argues that religious music and religious venues are not merely influenced by exogenous factors, but rather exhibit a profound and intricate coupling relationship. Tibetan Buddhist music, characterized by its combination of slow tempo and high reverberation, enables listeners to experience direct low-frequency vibrations, thereby generating a strong, expansive sense of sacredness. Christian music, with its moderate tempo performed within highly reverberant spaces, maintains appropriate rhythmic clarity within a temporally extended acoustic environment, creating a sound field that possesses both continuity and structural distinctiveness. This sound field enhances immersion and transcendental experience, further reinforcing the generation of a structured sense of sacredness. Japanese Buddhist music, conversely, through its profound emphasis on "void" (空, *kū*) and "stillness" (静, *sei*), guides listeners toward self-reflection and introspection, embodying the aesthetic of Void-based Sacredness.

Our work is still very preliminary due to several constraining factors. The most obvious limitation is that only 9 buildings selected for analysis. There are ample acoustics data for Christian churches, but extremely limited for Tibetan monasteries and Japanese temples. Second, the results indicate more of correlations rather than causations between sacred music and venues.

Students interested in this project are encouraged to visit local churches or school halls to experience various musical styles in different acoustic settings. Alternatively, an easy way to conduct a natural experiment is to head to a mountain valley. By shouting at different tempos and spectral centroids, one can physically feel and observe the acoustic effects.

Acknowledgement: We sincerely thank the reviewer for the careful reading of our manuscript and for the thoughtful, detailed, and constructive comments. We greatly appreciate the time and effort devoted to evaluating our work. The review has helped us identify areas that required clarification and improvement, and the suggestions have substantially strengthened the manuscript. We also developed an interactive simulator at **hamusic.netlify.app** that enables users to experience the acoustic effects of different architectural settings using preloaded religious music samples. In addition, users can upload their own audio files to explore how different spaces influence their acoustic characteristics.

6 References

Blessner, B., & Salter, L.-R. (2007). Spaces speak, are you listening? Experiencing aural architecture. MIT Press.

- De Muynke, J., Baltazar, M., Monferran, M., Voisenat, C., & Katz, B. F. G. (2024). Ears of the past: An inquiry into the sonic memory of the acoustics of Notre-Dame before the fire of 2019. *Journal of Cultural Heritage*, 65, 169–176. <https://doi.org/10.1016/j.culher.2022.09.006>
- Di Loreto, S., Pierantozzi, M., Lori, V., & Serpilli, F. (2025). Impact of architectural styles on acoustic characteristics in selected European churches. *Architecture*, 5(1), Article 5. <https://doi.org/10.3390/architecture5010005>
- Foteinou, A., Stevens, F., & Murphy, D. (2025). Revisiting the acoustics of St Paul's Cathedral, London. *Acoustics*, 7(3), Article 49. <https://doi.org/10.3390/acoustics7030049>
- Martellotta, F. (2016). Understanding the acoustics of papal basilicas in Rome by means of a coupled-volumes approach. *Journal of Sound and Vibration*, 382, 413–427. <https://doi.org/10.1016/j.jsv.2016.07.007>
- McFee, B., Raffel, C., Liang, D., Ellis, D. P. W., McVicar, M., Battenberg, E., Nieto, O., et al. (2015). librosa: Audio and music signal analysis in Python. *Proceedings of the 14th Python in Science Conference*, 18–25. <https://doi.org/10.25080/Majora-7b98e3ed-003>
- Pentcheva, B. V. (2011). Hagia Sophia and multisensory aesthetics. *Gesta*, 50(2), 93–111. <https://doi.org/10.2307/41550552>
- Romel, G., Mansour, Y., & Mohamed, A. (2025). Acoustic performance and sustainable restoration in historical religious buildings: A review. *Journal of Umm Al-Qura University for Engineering and Architecture*, 16(4)*. <https://doi.org/10.1007/s43995-025-00218-w>
- Soeta, Y., Shimokura, R., Kim, Y. H., Ohsawa, T., & Ito, K. (2013). Measurement of acoustic characteristics of Japanese Buddhist temples in relation to sound source location and direction. *The Journal of the Acoustical Society of America*, 133(5), 2699–2710. <https://doi.org/10.1121/1.4796116>