



HAL
open science

Architectural acoustic design: Observation of use cases including audio-only and multimodal auralizations

David Thery, David Poirier-Quinot, Sebastien Jouan, Brian F. G. Katz,
Vincent Boccara

► To cite this version:

David Thery, David Poirier-Quinot, Sebastien Jouan, Brian F. G. Katz, Vincent Boccara. Architectural acoustic design: Observation of use cases including audio-only and multimodal auralizations. Building Acoustics, 2021, pp.1351010X2110455. 10.1177/1351010X211045518. hal-03348293

HAL Id: hal-03348293



<https://hal.science/hal-03348293v1>

Submitted on 5 Oct 2021

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.

Architectural acoustic design: Observation of use cases including audio-only and multimodal auralizations

David Thery¹ , David Poirier-Quinot²,
Sebastien Jouan³, Brian FG Katz² 
and Vincent Boccara¹

Abstract

Auralization technology has reached a satisfactory level of ecological validity, enabling its use in architectural acoustic design. Only recently have the actual uses of auralization in the consulting community been explored, resulting in the identification of a variety of uses, including (1) to present to clients, (2) to test design ideas, (3) as a verification tool, (4) as a verification tool, (5) as a marketing tool, and (6) to improve internal company discussions. Taking advantage of methodologies from ergonomics research, the present study investigates effective uses through the observation of a collaboration project between an acoustic research team and an acoustic consultant, as a case study. Two spaces have been auralized in the context of the conception of a new skyscraper during the design phase of the project. The two spaces faced different problematics: an Atrium for which three different acoustic treatment options were suggested and experienced through multi-modal auralizations and audio-only auralizations of an Auditorium where an intrusive noise was to be acoustically treated. The ergonomic observation and analysis of this project revealed key impediments to the integration of auralization in common acoustic design practices.

Keywords

Auralization use, acoustic consulting, virtual reality, case study

Introduction

In the acoustic consulting community, auralizations are particularly suited for the design of public spaces, restaurants, and art-oriented spaces such as concert halls or museums.¹ These studies show

¹LISN-CNRS, Université Paris-Saclay, Cognition-Perception-Uses (CPU) Team, Orsay, France

²CNRS, Institut Jean Le Rond d'Alembert, Sorbonne Université, Lutheries - Acoustique - Musique, Paris, France

³Theatre Projects, Paris, France

Corresponding author:

David Thery, LISN-CNRS, Université Paris-Saclay, Cognition-Perception-Uses (CPU) Team, Rue John Von Neumann, Orsay 91405, France.

Email: david.theryfr@gmail.fr

the diversity of potential applications for auralizations, as well as actual examples of use in architectural acoustics projects, including design variables and requirements that clearly depend on the type of project. For instance, Hochgraf² shared their experience with auralizations and identified four main uses: (1) communicate with clients, providing a direct listening experience without being confused with acoustic terms; (2) assist design decision-making, enabling the simulation of different configurations and potential uses of the space; (3) eliminate unwanted defects once the space is finished; and (4) build enthusiasm and support fund-raising. The author also insisted on the importance of calibrating levels, the inclusion of Lombard effect modeling,³ and the use of appropriate high-quality anechoic material as audio source,⁴ to create convincing and realistic auralizations.

Milo⁵ provided a comprehensive review of acoustic design practices, concluding that “*the education of future acoustic designers should take place also in architecture schools,*” to help students develop a common vocabulary allowing them to communicate more easily during architectural design tasks.

In the related field of Virtual Reality (VR) technology, Woksepp and Olofsson⁶ investigated the use of visual in situ VR immersion by following large construction projects. They studied how VR scenes were experienced and assessed by the professionals involved, and “*to what extent VR can complement the use of traditional 2D CAD drawings*” and other types of visualizations such as hand-drawings. The authors also investigated how VR immersion was applied and accepted by professionals during both design and planning process. They concluded that VR was very useful for large constructions, being helpful to design decision-making by facilitating information handling.

The degree of adoption of VR technologies in architecture has been investigated in a few studies.^{7,8} Atkins⁸ suggested that these technologies have the potential to improve productivity and reduce costs, but are not yet adopted by the architectural community. To date, no such investigation has been carried out on the role of auralizations in the acoustic design and consultant communities.

In this context, They et al.⁹ presented a survey study that investigated the use of auralization in the acoustic design community. The methodology, coming from the ergonomics field of research, included a questionnaire and semi-directed interviews, enabling the identification of *declared uses* of auralizations by acoustic consultants. A reasonably accurate insight into the practices of acoustic consultants regarding their use of auralizations was obtained from 74 acoustic consultants around the world, enabling to identify the tools used, the main uses, benefits, and difficulties linked to the application of this technology. The adoption rate of auralization has been shown to be quite low, which was mainly explained by a lack of resources (cost and skills). However, these results were only based on declared uses by consultants, potentially omitting undeclared uses or unpredictable external constraints encountered in architectural projects, which can be highlighted by a more situated approach, as presented in this study and documented in the ergonomics research literature.¹⁰

Complementary to that survey, the objectives of the present study are to analyze in a specific context (situated approach): (1) the effective uses of auralization, (2) the effective practical difficulties and impediments to the integration of this technology in common practices, and (3) the effective benefits it can provide during an acoustic design project.

Following a brief introduction of the global research methodology approach utilized in this study in Section 2, this paper begins with a description of the project in Section 3, including the initial auralization request from the acoustic consultant (TP) to the auralization research laboratory (SU), the actors involved, as well as the auralization proposal from SU to TP. Section 4 presents the methodology used to collect and analyze the data. Results are presented in Section 5, subsequently discussed in Section 6, leading to conclusions in Section 7. The technical work performed

for the creation of the auralizations was reported in Thery et al.¹¹ For the sake of clarity, a short glossary is provided in Section 2.

Ergonomics research methodology

This study applies methodologies coming from ergonomics research to the research field of acoustics. Said ergonomics research exploits qualitative research, with methods including questionnaires, interviews, and observation. Some of the terms uncommon to acoustic research that are applied here, originating from social science and qualitative research, are defined here: [noitemsep]

- Thematic tree: hierarchy of categories and sub-categories (or themes and sub-themes) allowing to categorize short text extracts.
- Thematic analysis: qualitative research method consisting in analyzing the proportion of appearance of various themes in textual (potentially transcribed from video and/or audio) data.
- Theme/category: item in the thematic tree.
- Segment: short text extract of variable length to be categorized in a thematic tree. The length of the segment should be sufficient to gather meaning in the extraction.
- Coding: assigning a segment to a category.
- Occurrence: appearance of a category/theme.

This study is conceived as an ergonomic case study of a specific acoustic design project conducted as part of a collaboration between Sorbonne University (SU) and the Theater Projects (TP) acoustic consultant, where auralization was used to support acoustic design choices (to respect a Non-Disclosure Agreement (NDA), sensitive information has been anonymized in this paper). It can be beneficial both to acoustic consultants and researchers, respectively, to include new technologies in their practices as well as gain insight into the practical aspects of using these tools in a project, and to orient future research based on observed impediments to the adoption of these technologies.

While Thery and Katz¹² focused on declared uses and practical acceptability (whose dimensions are defined in Figure 1),¹³ this study focuses on situated acceptability, in a context of use of the technology, during which potential undeclared uses or difficulties can be observed and identified.^{10,14}

Results analysis will be based on the dimensions of practical acceptability, which are defined here:

- Learnability: can be measured by the time it takes to reach a certain level of use. One should keep in mind that users normally do not take the time to learn the entire system before starting to use it. Easy to understand messages make it possible to do useful work before having learned the entirety of the system.
- Efficiency: refers to the expert user's steady-state level of performance at the time when the learning curve flattens out.¹³
- Memorability: time it takes to remember how to use it, depending on the frequency of use.
- Errors: users should make as few errors as possible when using the system; for example, availability of *undo* helps in avoiding errors, confirmation questions should be available before execution of risky commands. Errors could be more catastrophic in nature if they are not easily discovered by the user. Nielsen¹³ defined the term error as "any action that does not accomplish the desired goal." These can be interpreted as various types of difficulties encountered.

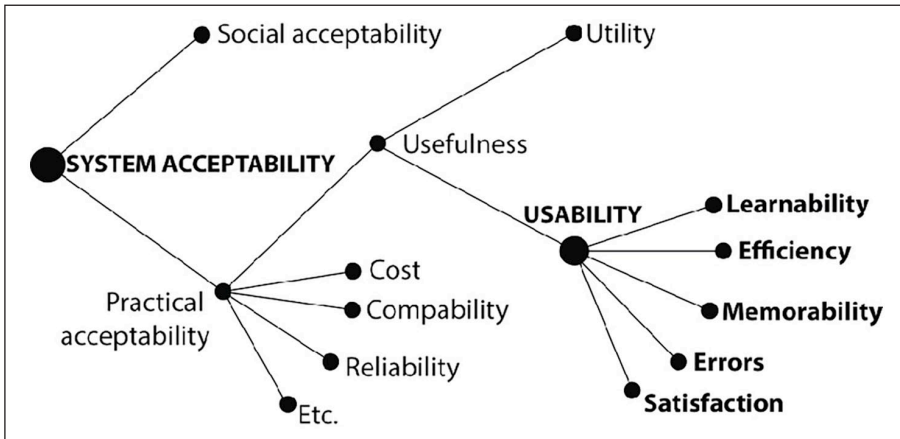


Figure 1. Nielsen's acceptability model and its practical acceptability dimensions.

- Subjective satisfaction: entertainment value (the amount of time users spend with the technology is not of concern if they enjoy using it).
- Cost generated by the use of the technology, whether in time, money, or human resources.
- Compatibility: both from a purely technological point of view (e.g. version match between tools, format compatibility) and with regards to the actual uses, practices, methods of work.
- Reliability: both from the reliability of the obtained results (in terms of uncertainty), and of the tool itself (e.g. robust, solid, stable).

Case study context: Project description

Project actors

While the overall architectural project includes many participating entities, the list of those directly involved are described here. *The Client*, being also the property developer, was based in the US, in direct interaction with the acoustic consulting team, *TP*. *TP* was in charge of the recommendations for the acoustics of the different spaces. The *TP*-team comprises three acousticians: the principal acoustician from the Parisian office (*SJ-TP*) (it is noted that the *SU*-team includes several of the authors), an acoustician of this same office (*SP-TP*), and one of their London directors (*FR-TP*). While *SJ-TP* had prior experience with auralizations, the *TP*-team had no experience in VR. *The Design architect*, provided the visual model to *SU*, and was located in the US. *The Auralization research team*, *SU*, based in Paris, consisted of Dr. *GF* (*GF-SU*) who was responsible, and two research assistants (*AZ-SU* and *DD-SU*), specialists in room acoustics and virtual 3D audio (it is noted that the *SU*-team includes several of the authors). *Additional external actors* included ShowTimeVR (<https://showtimevr.eu/>), who developed the 360 video player with binaural sound with remote control, L-acoustics, who intervened providing a speaker directivity pattern, and Acentech, who provided pseudo-anechoic recordings of suitable restaurant noises.

Initial request: Auralization of two spaces

This project arose from discussions between *TP* and *SU* regarding the use of auralizations for design purposes. In September 2018, an active project of *TP* was selected as a test case to observe

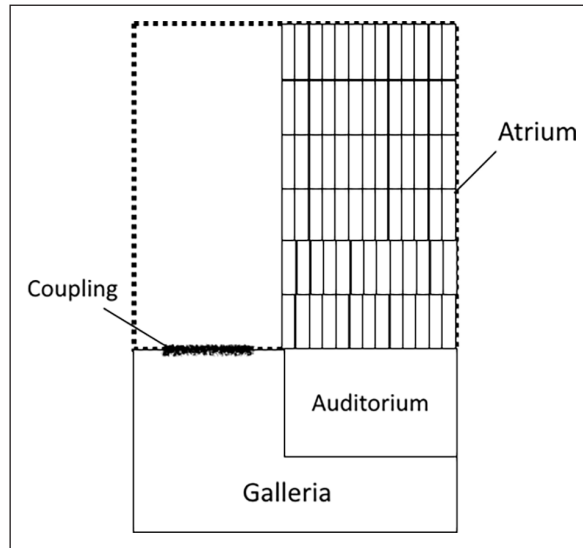


Figure 2. Schematic representation of the auralized spaces (section view), including the Atrium (surrounded by glass) and the Auditorium affected by the intrusive noise coming from the Galleria (sub-space). The coupling between the Atrium and the Galleria is also represented, although this option was not part of the present study.

client/consultants interactions in the context of auralization use. The project concerned two spaces for which auralization was considered meaningful in the design process, schematized in Figure 2. A detailed description of the creation of these auralizations is available.¹¹

The first space was a large glass volume called the Atrium (sometimes called the Glass House), to be used for various purposes (dinner banquet, conference, recitals, jazz, and amplified music). The role of the consultant was to dampen the reverberant acoustic resulting from the glass structure with minimum impact to its overall luminosity. The associated auralization was to be rendered over headphones (binaural), and complimented with an immersive visualization of the space rendered over a Head-Mounted Display (HMD). The second space was an Auditorium, potentially affected by intrusive noise coming from the floor below (Galleria in Figure 2). The role of the auralization was here to provide evidence for the advantage of a proposed structural isolation. This consisted in building the Auditorium as a box-in-box construction. The audio-only auralization was to be rendered over a loudspeaker setup in an acoustically damped room.

A first presentation to the client of these two auralizations was planned for end of October 2018 at SU laboratory, providing 1.5 months to produce the auralizations. After this presentation, TP requested additional auralization configurations, to be presented to the client outside of the laboratory using a portable HMD.

Space 1: Multi-purpose Atrium. The main space to auralize was the Atrium, whose dimensions were 40 m height, 20 m width, and 50 m depth, leading to a volume of 1600 m^3 , with a rectangular footprint. As it is surrounded by glass, it was anticipated to have an undesirable amount of reverberation, hence needing acoustic treatment. The client wanted the Atrium to be multi-purpose, to include dining, conferences, as well as concerts of small ensembles. These uses of the space required stimuli for the auralizations, including speech and several kinds of music. In agreement

between TP and SU, the following three stimuli were selected: (1) anechoic female speech recording, originating from Southampton University (<https://www-mobile.ecs.soton.ac.uk/>); (2) a dry version of the cover *U can't touch this* from singer *Aubrey Logan*; (3) an anechoic recording of *Dave Brubeck's Take Five* from Aalto University Pätynen et al.⁴

Regarding the acoustic treatment, TP specified three options: (1) no treatment, (2) light diffusers only, and (3) light diffusers and low frequency dampening *AQflex*. The light diffusers (149 g/m²) exhibit high absorption performance at mid-frequencies (octave bands 500–4000 Hz), but are less efficient at lower frequencies (below 400 Hz). These were mounted with a 0.15 m air gap behind, on the side walls of the Atrium.^{11,15} In contrast, the recently patented *AQflex* (600 g/m²) provides an almost equivalent absorption at the lower octave bands as the mid-frequencies.

Space 2: Auditorium intrusive noise. The intended use of the Auditorium was to host conferences and musical recitals, leading to Reverberation Time (RT) specifications ranging from 1.0 to 1.4 s respectively. Below the Auditorium would be the Galleria, a space from which amplified music and restaurant activity would potentially leak noise into the Auditorium. For that purpose, TP proposed to the client an isolating structure. The objective was to show the benefit of applying this concept in order to have acceptable acoustics in agreement with the intended uses of the space. TP provided SU with the appropriate noise specifications (in the form of filtered noise octave band level values, containing primarily low frequencies). The selected stimuli included a variety of musical styles and instrumentation, with the following samples: (1) a violin anechoic recording¹⁶; (2) the anechoic female recording used in Space 1; (3) the jazz song *Dave Brubeck's Take Five* used in Space 1; (4) flamenco guitar anechoic recordings (Woirgardt et al., 2012³⁰); (5) two classical trio anechoic recordings,⁴ including *Brahms' Horn trio* and *Schubert's piano trio*, both in *Eb* major.

Project timeline and deliverables

During the period of collaboration, eight phases have been identified, illustrated on the project timeline (Figure 3). P1, P3, P5, and P7 are considered as *creation phases* (lasting respectively 34, 7, 7, and roughly 90 days). P2 and P4 are *review phases* (lasting respectively 30 min and 1.5 h). P6 is the presentation to the client (lasting 2.5 h), and P8 is a technical support meeting (lasting 2.5 h).

Based on the intended uses of the space, the objectives for each auralization, and the specified acoustic treatment options, two auralization deliverables were provided by SU to TP in the context of this study, summarized in Table 1. The first deliverable was handed over in two parts. The first deliverable, that included the auralization of the Auditorium and a series of Atrium configurations, was presented during a meeting with the client and the TP-team at the SU laboratory. The second deliverable only included additional configurations of the Atrium, delivered in the form of 360 videos (no Auditorium auralization).

Auralization framework

Auralizations of the various acoustic treatment configurations were simulated off-line to achieve the highest level of realism. Multiple tools (software and hardware) were used for simulating, composing, and presenting the final auralizations: 11 different software, hardware devices including two types of HMDs, and two types of audio rendering systems. Three different visualization software were used: *Revit*, *Blender*, and *Unity*. RIR simulations were performed in *CATT-Acoustic* while the final audio rendering was handled in *Cycling '74 Max*, communicating with the *Unity* visual rendering via the Open Sound Control (OSC) protocol. Presentation scenario management and visual scene fade-in and out in the virtual model were performed in *Unity* scripts (C#). Other programming tasks including RIR

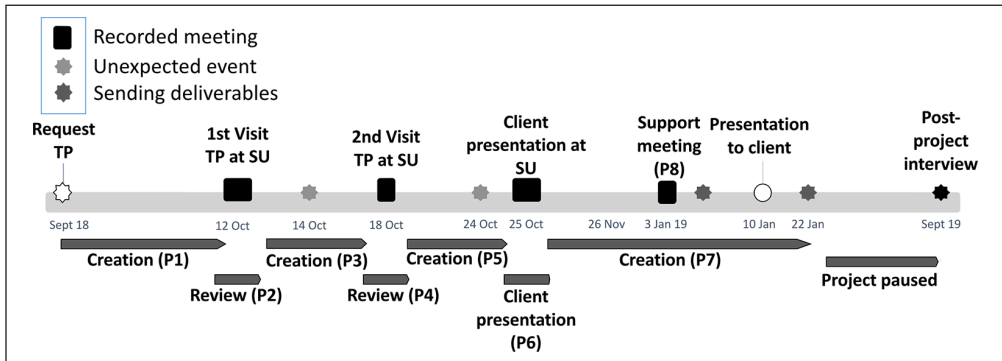


Figure 3. Detailed timeline of the collaboration SU-TP.

convolution, normalization, and level adjustments were performed in *Matlab*. In terms of hardware, several computers were required; for the presentation of auralizations, both *Oculus CV1* and *Oculus Go* were used, the latter allowing for portable presentation of 360 videos. The sound was rendered binaurally over open headphones—with tracking being available via the HMD orientation information—or decoded over a 32-speaker system, necessitating amplifiers and professional sound cards. The creation of the 360 videos involved the use of four other software and developer libraries, including *ShowTimeVR* for remote control of the HMD during presentations, *Facebook Audio Workstation*, *Android File Transfer* for the portable HMD, and *ffmpeg*. Video presentation was performed on an *Oculus Go* and an *iPad*. This framework is schematically represented in Figure 4.

Methodology

The methodology used in this study comes from qualitative data analysis studies, termed *thematic analysis*, and is particularly used in social sciences where questionnaires and interviews are used extensively.^{17,18} This work applies this methodology to the field of architectural acoustics, which is novel, to investigate the effective uses, benefits, and practical difficulties impeding the integration of auralization to common practices in acoustic consulting.

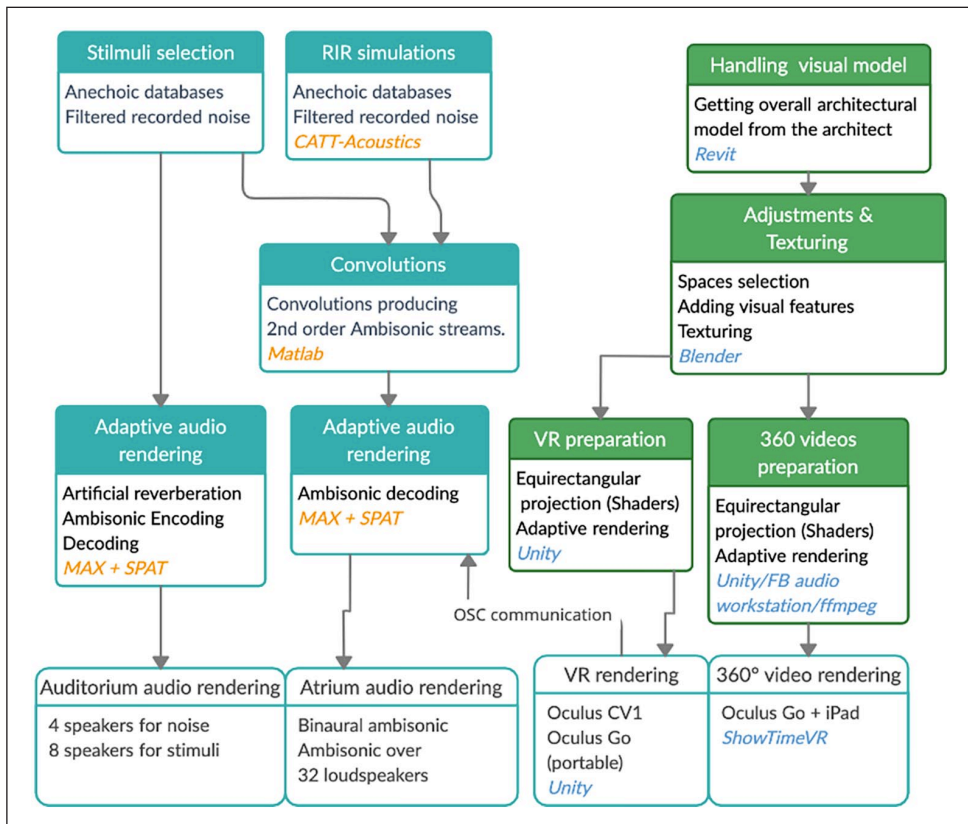
Thematic analysis is a widely used method for identifying, analyzing, and reporting patterns (themes) within qualitative data. The unit of extraction is called a segment and represents the shortest text extract with a self-contained meaning. The length of these extracts is therefore not constant, being selected as soon as they represent a meaning, from a couple of words to entire or even several sentences to recall the context. A categorization of these text segments is performed, necessitating the creation of a thematic tree, whose construction can be an endless iterative and refinement process.¹⁹ This categorization subsequently allows one to quantify the amount of data that concerns a particular topic with regards to the entire corpus. As such, it is possible to analyze the quantity of discussions on various topics. It should be noted that only the number of occurrences of a given category is analyzed in this study. Unless expressed in a sufficiently nuanced way, a given idea mentioned several times in a single phrase was counted once. As a consequence, a given idea shared by different people could be counted several times.

Data collection

In order to perform the analysis of this study, several types of data were acquired following three steps. First, all *emails* were gathered (128 emails), producing a 96 page document (>20,000

Table 1. Auralization proposal.

| | Atrium (first deliverable) | Atrium (second deliverable) | Auditorium |
|-------------------------|---|--|--|
| Acoustic configurations | Two positions (close/far) Three absorption conditions Acoustic seal closed (coupling with Galleria) RT specification: 2.0 s | Two positions Three absorption conditions Acoustic seal open/closed (coupling with Galleria) RT specification: 2.0 s | Single static position RT specification: 1.0–1.4 s |
| Audio stimuli | General public circulation Small music ensemble Amplified music | Banquet (people) Small music ensemble | Small music ensemble Conference (speech) Calibrated levels |
| Deliverables | Presentation to client 360 videos | 360 videos | Presentation to client |

**Figure 4.** Diagram representing the audio-visual auralization framework, used for both auralizations, indicating the audio (left, in blue) and visual/VR (right, in green) elements.

words). Email lengths ranged from 3 words to 856 words (mean=92 words), from simple answers to detailed explanation of acoustic concepts. This document enabled the project progression to be traced. It included exchanges principally between TP and SU (81%), in addition to some minor

external actors (19%). Direct emails to and from the client were not accessible, and are consequently not included in the present analysis.

Second, in addition to notes taken during the *observations* of the meetings, *audio recordings* were made. Four meetings were recorded for a total of 7 h (respectively 2600, 9200, 20,400, and 12,800 words in the transcribed text). For practical reasons, and to remain unobtrusive, only audio recordings were possible. Video recordings would have produced more data, particularly for analyzing non-verbal communication, action, and engagement of the involved actors.

Finally, a *post-project interview* was conducted 8 months after the delivery with SJ-TP, who was in direct contact with the client, to obtain feedback and their intentions of use of auralization after this project. It was a semi-directed interview conducted by DD-SU in video-conference, and lasted 30 min, with the audio being recorded. The three categories were: *confidentiality*, *general feedback about the project*, and *the adoption/integration of auralizations following the acquisition of the portable HMD*. This interview was not transcribed, while key information was extracted.

Treatment and analysis

The first step of the analysis was a chronological concatenation of all gathered emails. Metadata (date, hour, length, presence of attached document, sender, and recipient(s)) were extracted to provide indicators for the analysis. This chronology was used to construct the project timeline (Figure 3) and to analyze the temporal distribution of the phases of the project. Secondly, the audio recordings of the four meetings were transcribed.

Emails and audio recordings were subsequently independently thematically analyzed using the software MAXQDA³¹, using a data driven approach for the creation of the thematic tree.¹⁷ The thematic analysis necessitated four steps:

1. Creating the thematic tree (categories) using all gathered data including emails and meetings transcriptions (see Table 2). This phase makes it possible to obtain a holistic view of the data, needed for its comprehensive understanding ($\approx 65,000$ words in total).
2. Independent coding of each phase material: assignment of text segments to one of the thematic tree categories or sub-categories.
3. Extracting quantitative data (percentages of coded segments, overall, by phase, and by auralization) for each category and sub-category.
4. Synthesizing the content of each category/sub-category, drastically reducing the amount of textual data into a few sentences, from which the main ideas are extracted and reported in the results.

A total of 763 segments were coded: 47.5% from emails and 52.5% from audio recordings respectively. Amongst these, 68% concerned the Atrium, 28% the Auditorium, 4% tackling subjects common to both. Among the 763 segments, 117 were coded as belonging to several categories ($\approx 15\%$). Coding categories are reported in Table 2.

Results

Resulting analysis of the collected and aggregated social interaction data is presented following a hierarchical conceptual division. First, the observed uses are presented in Section 5.1 and analyzed along the high-level variables of this analysis approach: *type of auralization* (multimodal Atrium vs intrusive noise Auditorium) and *temporality* (Creation, Review, and Presentation). The second

Table 2. Summary of the elaborated thematic tree including sub-categories (three level is not included for clarity), illustrated by an example quote.

| Category | Sub-category | Definition | Illustrative quote |
|--------------|--|--|---|
| Organization | Calendar Deliverables/NDA | Appointments/schedule planning Auralization configurations specifications | "The 25th is possible, though of course the following day/week would be better" "The most straightforward deliverable that we can provide is a collection of static position auralization with HOA audio" "We are currently testing on an Oculus Go and a large iPad" |
| Tools | Audio Visual, VR | Related to tools, software, hardware | "Auralizations [. . .] can be used for academic research" "This enables everybody to see on the large screen" "What's not realistic is that there is too many people in the audience" "But if we put speakers throughout the space [. . .] it's not necessarily going to improve" |
| Uses | Research AD-achieve realism AD-acoustic design | Used for research purposes Used for marketing purposes Emphasizing the need for the auralization to be realistic Design choices (sound system, stage, curtains. . .) | "No, he just produced the CLF file for Odeon [. . .]" "For the anechoic sounds, try to gather sounds usable for the different use cases; at least, propose a list of possibilities" "[. . .] the sound recorded Friday, then filtered in Live" "It's too loud, but at least I can feel the reverb" "Our collaborators would still like to have some screen shots of your walkthroughs to be sure of the visual texture" |
| | AD-GA simulations AD-multi-purpose hall | GA model and simulations Importance of this constraint with regards to auralization creation | "I suggest a listening test on a pre-selection of stimuli in the Glass House" "A key point is to know if the visit demo is what design decisions will be made on" |
| | AD-sound design AD-stimuli influence AD-visual models | Use of sound design techniques Related to sound sources Related to visual models | "You should hear a decent benefit for the vocal singer but the double bass should be quite reverberant" "I think we do not start with the public noise. We add it if he makes the comment that he feels the space is empty" |
| | CP-listening/perception CP-negotiation | Listening moments and perception discussions Related to negotiation discussion with the client (program, anticipation of client's expectations) | "If you have a compatibility issue [. . .]" "I just know where we're going on our budget" "We have suffered a major computer problem" "With the A-ponderation, I told you to do that" "Only real issue: the controller app (iOS) crashed on me [. . .]" "I wouldn't spend any time trying to acoustically isolate [. . .]" "Does it leaves you enough time to prepare the demo?" |
| | CP-pedagogy CP-scenario planning | Auralization used to explain acoustic concept to the listener Related to planning the order of stimuli, sequence of events | |
| Difficulties | Compatibility Cost Hardware issues Human errors Lack of skills New technologies Time | Incompatibility issues encountered Related to budget or cost Problems from hardware Related to mistakes, arguments Lack of training or competences Difficulties provided by the novelty Time related discussions | |

AD: auralization design; CP: client presentation.

Table 3. Distribution of the coded segments in the *auralization design* (AD) and *client-presentation* (CP) categories, respectively in Atrium (respectively 167 and 118 codes) and Auditorium (respectively 118 and 91 codes).

| Category—space | Atrium (%) | Auditorium (%) |
|-------------------------|------------|----------------|
| AD-acoustical design | 21 | 32 |
| AD-achieve realism | 4 | 20 |
| AD-GA simulations | 31 | 2 |
| AD-multi-purpose hall | 4 | 10 |
| AD-sound design | 2 | 7 |
| AD-stimuli influence | 29 | 29 |
| AD-visual models | 9 | 0 |
| CP-listening/perception | 47 | 51 |
| CP-negotiation | 36 | 32 |
| CP-pedagogy | 14 | 8 |
| CP-scenario | 3 | 9 |

results section (Section 5.2) presents an analysis of these uses along each dimension of practical acceptability.

Observed uses of auralization

The categories of uses (presented in Table 2) included *Research* (1% of coded segments of uses), *Marketing* (1.5%), *Presentation to the client* (47%), and *Auralization design* (51%), totaling 461 out of the 763 coded segments (60.5%).

Comparing Atrium and Auditorium. An overview of the content of the collected data regarding auralization use, separating the Atrium and Auditorium auralizations, is given in Table 3.

The *Auralization design* category of the Atrium was principally concerned with *GA simulations* (31.7%), the *Stimuli influence* (28.7%), *Acoustical design* (21%), and a non-negligible portion concerned the *Visual model* (9%). Interestingly, the category *Achieve realism* represented only 4.2% of the Atrium data. While a similar important proportion of Auditorium codes (i.e. coded segments) were dedicated to the *Influence of the stimuli* (29.3%), this intrusive noise auralization was also concerned with *Acoustical design* (31.7%), the need to *Achieve realism* (19.5%), and having a *Multi-purpose hall* (9.8%). A significant amount concerning *Sound design* was noted (7.3%), showing the attention paid to the realism of the stimuli in this type of auralization (these codes concerned the design of the recorded and filtered sound simulating the intrusive noise).

The *Stimuli* choice was of particular importance in both cases, though with different aims. In the Auditorium, the level was set based on noise specifications, while the listening level in the Atrium was set perceptually to provide the most realistic listening level. In the Atrium, the stimuli were selected for their capacity to put in evidence the efficiency of the acoustic treatment, while in the Auditorium, the aim was to highlight conditions where the background noise would be noticeable within realistic performance situations, which led to the selection of a solo violin in which the presence of pauses enabled the noise to be more noticeable, as illustrated by the following quote (from AZ-SU, during P6): “*the violin and the guitar are good; piano and ‘take five’ are not the best as they mask everything, and to convince, that’s not so good.*”

Table 4. Distribution of the coded segments in the *Auralization design* (AD) and *Client Presentation* (CP) categories, respectively for the creation, review, and presentation phase, merging Atrium and Auditorium data.

| Category—space | Creation (%) | Review (%) | Presentation (%) |
|-------------------------|--------------|------------|------------------|
| AD-acoustical design | 31 | 19 | 36 |
| AD-achieve Realism | 2 | 28 | 3 |
| AD-GA simulations | 29 | 3 | 18 |
| AD-multi-purpose hall | 10 | 0 | 8 |
| AD-sound design | 2 | 10 | 0 |
| AD-stimuli influence | 13 | 40 | 30 |
| AD-visual models | 13 | 0 | 5 |
| CP-listening/perception | 57 | 69 | 35 |
| CP-negotiation | 36 | 9 | 48 |
| CP-pedagogy | 7 | 5 | 15 |
| CP-scenario | 0 | 17 | 2 |

Creation phase (respectively for AD and CP absolute number of coded segments: 80 codes and 14 only); review phase (respectively 69 and 74 codes); presentation phase (respectively 62 and 122 codes).

The *Acoustical design* category focused on different aspects for the two auralizations. For the Atrium, these included the design of the sound system, the comparison of acoustic features with an existing hall and previous project from TP (the Lincoln center), and the presence of curtains (discussing permanent vs non-permanent installation), as illustrated here (SJ-TP, P1): “*The glass house could be based on the same system, either with line arrays, coaxial boxes or columns with subs.*” For the Auditorium, noise level specifications were given by TP and the intrusive noise was measured with a sound level meter to ensure the requirements were met.

Discussions concerning the *Achievement of realism* were primarily concerned with the target reverberation time (GF-SU to SJ-TP, P1): “*Could you also indicate to us the target RT for the Auditorium?*” and the response (SJ-TP to GF-SU, P1): “*the RT should vary between 1.0 and 1.4s. Take 1.4s in recital mode.*”). In the Atrium, this realism was further supported by audio-visual coherence (in particular for the presence of acoustic treatments both in the auralization and the visual model) and the quality of the ambient noise. In contrast, in the Auditorium the focus was exclusively on the realism of the intrusive noise, and its provenance from lower elevations for increased realism (simulating the sub-space called the Galleria). Therefore, the presence of visuals was a key difference.

Regarding the *Client presentation*, Table 3 shows a consistent division of the sub-categories across auralization types: for both the Atrium and the Auditorium auralizations, the majority of codes concerned the activity of *Listening* and the perception of auralizations of the various actors (46.6% and 51.6% respectively), followed by a significant portion concerning the *Negotiation* aspects with the client (36.4% and 31.9%). The use of auralizations as a *Pedagogical tool* was observed with both auralizations (respectively 13.6% and 7.6%), while the *Scenario setting* was more prominent for the Auditorium (8.8% vs 3.4% for the Atrium).

Temporality analysis. The creation phase included P1, P3, P5, creation phases before the presentation to the client, excluding P2 and P4, as these were considered “*Review phases.*”

Table 4 shows the division of categories of the exchanges in the *Auralization design category*, dominated by *Acoustical design* (31.5%) and *GA simulations* (28.1%), with significant content

contributions of the *Visual model* (13.5%), *Stimuli* (12.4%), and the *Multi-purpose* (10.1%) specificities of the spaces.

The *Acoustical design* category contained information exchanges about: sound system design (type, directivity, and placement/orientation of the speakers in the space to optimize the acoustics, while taking into account the constraints of the space such as screen deployment, and how to integrate into the existing structure); Preferred Noise Criteria (PNC) curves specification for the Auditorium, based on TP calculations.

Exchanges also concerned the architectural *Visual model*, which was provided by the architect, re-textured, and integrated into Unity by SU. Hence, several requests were sent for obtaining screenshots of the interior of the model, details such as the actual spacing between absorbers, or the size of the absorbers, as in the following quote (GF-SU, P1): “*If we could have views with textures and screenshots of the glass house that would be great.*”

The search for suitable *Stimuli* was a significant part of this creation phase, needed for the different intended uses of the spaces: conferences, banquets, and small band musical performances. Recordings of ambient noises (restaurants) as well as shaped intrusive noises were discussed with feedback from the different actors in order to obtain realistic stimuli (2.2% of the *Auralization design* codes), by means of sound design (2.2%) ((DD-SU, P3): “*This is not the final version as I did some spectral adjustments directly in the listening room today*”).

Table 4 shows the distribution of codes within the *Client Presentation* category. Only 14 segments were coded, limiting the amount of data contained here, due to the shorter time spent on this auralization. The *Listening/Perception* category was concerned mainly with feedback between SU-team and TP about the rendering of the visual model and the realism of the auralizations that were exchanged by email (57.1%). The *Negotiation* category (37.5%) was concerned with the preparation of the presentation to the client, making sure the perceptual differences between acoustic treatments were well noticeable and almost “obvious.” The Pedagogical contribution (7.1%) was concerned with the explanation of the concept behind PNC curves as compared to the more commonly used NC curves, by TP to SU-team, who were not all familiar with these curves.

The review phase comprised P2 and P4. These two meetings spaced by 1 week, took place at SU laboratory, and involved the SU-team that has produced the auralizations, and SJ-TP, who was in direct contact with the client. The aim of these meetings was to assess the auralizations and rehearse for the presentation to the client.

Regarding the *Auralization Design* category, Table 4 shows a highly different division of content compared to the Creation phase, as *Stimuli influence* (39.7%) and *Achieving realism* (27.9%) were the main topics, while *Acoustical design* and *Sound design* were also discussed (respectively 19.1% and 10.3%).

Regarding the *Stimuli* category, particular attention was given to the filtered intrusive noise recordings of the Auditorium, which needed spectral adjustments to fit the specified PNC curves provided; the degree of excitation (or intensity) of the crowd noise was also a consideration. The addition of a sub-woofer was decided, as the lowest frequency bands were required in this example auralization, even at the 63 Hz octave band. Similarly, a modification of the loudspeakers layout to simulate the intrusive noise coming from the Galleria (sub-space) was decided to increase this impression of direction provenance. The actual content of the stimuli was debated, as the objective was to have the impact of the noise noticeable for the client: more intense (fortissimo vs pianissimo) and denser pieces were avoided as they applied more masking on the noise, which was not the case with a solo violin excerpt. For the Atrium, much time was spent to adjust the level of the stimuli, and the balance between the stimulus and the ambient noise in the “Banquet” configuration, done perceptually based on the impression of SJ-TP, without objective measurements ((SJ-TP, P6) “*It’s too loud, but at least I feel the reverb,*” and “*What’s the decibel level?*”).

Regarding the *Acoustical design* category, the codes mainly concerned the Auditorium, and the background noise measurements, in relation to the PNC curves and levels, comparing the specifications and the actual perceptual rendering.

Regarding the *Client Presentation* category, Table 4 shows that most of the data concerned *Listening* or *Perception* exchange moments during this review phase, and trying to adjust the stimuli to obtain the most realistic simulations (68.8% of the coded segments in this category). This included A/B comparisons between PNC levels, switching with and without public noises, spectral adjustments of the intrusive noise for the Auditorium, and listening to the different stimuli (Speech male and female voices comparisons, Jazz, Amplified music) in the Atrium, with A/B comparisons at different positions.

At the same time, the *Scenario* for the presentation was planned (17.2%), debating which stimulus to play first, if the demo would start with or without background noise, with the objective to have the client understand and notice the differences between conditions (concerning the recommendation by the acoustician to install the isolating structure for instance), in preparation of the *Negotiation* (9.4% of codes in this category), as illustrated by the following quote from GF-SU: “So in terms of scenario, we start with PNC25.”

The Pedagogical aspect was minor, concerned the Auditorium, in particular concerning the noise design according to the PNC curves, in preparation of the presentation to the client, as illustrated here (SJ-TP, P4): “Ok, this is perfect. We need to play this to the client, telling him that this is what we propose as a conception criteria. Meaning when you are in the room, you hear nothing, and all our recommendations are based on this criteria.”

The presentation phase consisted of a meeting which lasted 2 h and 20 min. The first hour was spent for the Auditorium auralization and the next hour with the multi-modal Atrium auralization.

Regarding the *Auralization Design* category, Table 4 shows that *Acoustical design* (36.1%) and *Stimuli* (29.5%) discussions were the most present categories, followed by a significant (18%) part dedicated to *GA simulations*, and minor contributions from *Multi-purpose hall* (8.2%), *Visual models* (4.9%), and *Achieve realism* (3.3%).

The auralizations, and particularly the Atrium, clearly engaged the client (based on observations), and helped start the *Acoustical design* dialog, related to the *GA simulation* category as well. The client asked for specific conditions ((Client, P6): “Can you put in the absorbent material”), listened and observed all components of the simulation and the model, from the lighting conditions, to the walls and ceiling coverage and distance impression, to the design of the sound system that produced the amplified music that was virtually reproduced.

Constructive discussions about the stimuli and the various acoustic configurations arose: differences between male and female voices, and the impact of the various absorption treatment options on these different stimuli were understood by the client. The importance of the stimuli was notable, even for the client, who asked for a specific (orchestral) sound source at the beginning of the presentation, as illustrated here (Client, P6): “I think jazz is very specific, a very intrusive sound because of the brass, because of the clarity of some of the instruments.”

Regarding the *Client Presentation* category, in contrast to the previous phases, the presentation was mainly concerned with *Negotiation* aspects (48.4%), followed by a large part of *Listening* (35.2%), and interestingly an increased contribution in the *Pedagogy* category (14.8%). Only 1.6% was concerned with *Scenario setting* ((SJ-TP, P6): “Imagine yourself in an Auditorium waiting for a concert”).

The *Negotiation* aspects were prominent during this presentation, although these information remain confidential.

The Pedagogical aspects encompassed the explanation of acoustic/perception concepts such as masking to the client, as well as the existing interactions between modalities, that the visual

Table 5. Distribution of the coded segments in the *Difficulties* category.

| Category—space | Global (%) | Atrium (%) | Auditorium (%) |
|------------------|------------|------------|----------------|
| Compatibility | 14 | 16 | 0 |
| Cost/budget | 15 | 7 | 56 |
| Hardware issues | 9 | 9 | 9 |
| Human errors | 16 | 14 | 26 |
| Lack of skills | 9 | 13 | 0 |
| New technologies | 23 | 24 | 0 |
| Time | 14 | 17 | 9 |

Overall (148 codes); Atrium (123); Auditorium (25).

modality could affect the auditory perception, and that the visual model was only there to provide a reference to support the auralization. The demo was driven in such a way that each condition was explained and successively listened by the client who understood the effect of a given treatment by directly experiencing it, as illustrated here (SJ-TP to the client): “*You should hear a decent benefit for the vocal singer but the double bass should be quite reverberant.*”

How the case study highlights the acceptability of auralization in acoustical design?

This section presents an analysis of the encountered difficulties, according to each dimension of practical acceptability,¹³ including *Learnability, Errors, Satisfaction, Efficiency, Cost, Compatibility, and Reliability*, defined in Section 2. The dimension *Memorability* is not discussed as the acquired data did not provide information regarding this dimension.

The division of sub-categories of *Difficulties* (see thematic tree in Section 4.2), that include *Compatibility, Cost, Hardware issues, Human errors, Lack of skills, New technologies, and Time*, are represented in Table 5, comparing the Atrium and the Auditorium auralizations. However, caution should be taken in the comparison, as only 25 segments were coded for the Auditorium as compared to the 123 segments of the Atrium. Still, this absolute difference shows that many more difficulties were encountered for the multimodal Atrium auralizations, due to the multiplicity of devices, software, and compatibility issues.

Table 5 shows that the most significant difficulty encountered for the Atrium was related to *New technologies* (24.6%), while the Auditorium exchanges mostly concerned *Cost and Budget* discussions (56.5%). The *Lack of skills* and *Compatibility* issues represented significant difficulties for the Atrium (respectively 15.8% and 13.2%), but did not appear for the Auditorium. Finally, the *Time* factor, *Hardware issues*, and *Human errors* were represented for both auralizations.

The difficulties related to *New technologies* were concerned with issues concerning the portable HMD, which had a limited audio output level for headphones, necessitating the use of an external amplifier to obtain the desired level, as well as issues concerning the playing of HOA-360° videos on this HMD. The use of ShowTimeVR in terms of both installation and remote control also proved to be laborious for the TP consultant (*Lack of Skills*), needing technical support from SU. The *Cost* was related to the negotiation during the Auditorium presentation, as well as the cost of the additional software ShowTimeVR. *Compatibility* issues were encountered with GA models and visual models, but mostly for the player of HOA-360° videos. *Time* difficulties concerned the frequent tight deadlines and last-minute requests (and delays for the deliverables), as well as the computation time GA simulations could take. *Hardware issues* concerned the portable HMD (battery, output audio level, HOA-player instability), as well as unexpected computer crash. Finally,

Human errors were related to arguments between project actors, miscoordination, or errors/mistakes in the exchange of data.

Learnability: learning how to use auralizations was previously rated as “rather difficult.”⁹ The current workflow gives some clues of “why.” The difficulty in learning how to produce high-quality auralizations, as is the case for GA-modeling,²⁰ is that it is complex rather than complicated, because there are many parts that need to be considered and mastered, but each part on its own is not exceptionally complicated. This point is supported by the wide variety of tools and programming languages needed to master the entire chain, from GA and HOA creation and rendering tools to visual model texturing and VR interactivity (with 80 codes out of the 763 concerning the tools, i.e. >10% of all data). The second major difficult aspect for learning is that some of these tools are new, representing 23% of the encountered difficulties (see Table 5). Moreover, there are often updates (or new formats, codecs) that should be followed, that could otherwise potentially lead to incompatibilities or unexpected bugs. While researchers and developers often cope with such problems, architects and acousticians will usually rely on more standardized software, saving time to focus on production. The need for training is therefore becoming more important in order to integrate these new technologies in the professional practices, as supported by this quote (FR-TP, P7): “*There is a gap that needs to be bridged between your production and the viewing software. We are unfortunately not experts in the domain and being caught in the middle of this situation is frustrating,*” as also reported in Milo.⁵ Additional quotes, coming from internal SU exchanges, respectively concerning *ShowTimeVR* and issues with the portable HMD, illustrate these problems (AZ-SU, P5): “*Only real issue: the controller app (iOS) crashed on me while I was messing around with settings during playback. No more control over the Go at that point [. . .] had to relaunch Showtime VR app in the Oculus to get control back. The controller app seems stable otherwise.*”

Errors: in this project, two types of “errors” were observed: material and human (related to the categories *New technologies*, *Human errors*, *Compatibility*, and *Hardware issues*, representing respectively 23.0%, 15.5%, 14.2%, and 9.5% of the overall coded difficulties (over 148 codes), see Table 5). One source of technical error was related to the use of recent software or *APIs* that requires advanced programming knowledge: in this project, for instance the use of second order Ambisonic IRs required to know the basics of channel ordering, normalization, or, regarding the visual model, knowledge about shaders and equi-rectangular projections in Unity. Without this advanced technical knowledge, one could easily make mistakes, impacting the realism of the simulation. The creation of the GA model (its level of detail, the associated acoustic measurements and calibration) was another potential source of error. Summing all deviations could significantly impact the overall ecological validity of the auralization.^{21,22} As software, hardware can also cause unexpected problems, for example, when the main computer running the auralizations broke down a few days before the visit of TP to the laboratory for rehearsals, necessitating the support from the manufacturer, introducing an additional time delay. Another observed error was categorized as *human errors*, as a result of organizational determinants. These included careless mistakes in emails, or sending the wrong data (two occurrences of this problem were observed: when sending the architectural visual model, some elements were missing, and when sending the PNC frequency curves, although they were corrected afterwards by email).

Satisfaction: the presentation meeting clearly showed the engagement of the client in the simulations, particularly for the multi-modal Atrium auralizations. This observation was supported by the numerous questions asked by the client during the simulation, to listen to a particular stimulus, as illustrated by the following quotes (Client, P6): “*Can you hear what I hear?*,” “*What’s on the roof?*,” “*And with the full absorbers now?*” The post-project interview (P8) revealed two points: first, that the client had not selected any of the options presented with both

auralizations, neither the structural isolation for the Auditorium, nor any of the acoustic treatments in the Atrium. Secondly, TP now proposes auralization as a service to their clients, in collaboration with SU. They purchased the portable HMD that was used during the collaboration, for its portability and its ease of use, demonstrating their interest and overall satisfaction. Still, more projects would be needed for a better evaluation of the satisfaction of each actor, ideally over a longer period.

Efficiency: in terms of productivity, the creation of all these auralizations involved three researchers working during >1 month for the first deliverables for presentation to the client, and three additional months at a less intense rhythm for the creation of the remaining multimedia files. Still, the process of creation could be optimized for future projects (by standardizing scripts, employing command line interfaces or batch mode processing for launching sets of GA simulations, or already having a wider anechoic ambient noise and stimuli database).^{2,12} During the post-project interview (P8), SJ-TP mentioned the difficulty of introducing auralizations to the client, primarily because it is still time consuming.

Cost: while these data are kept anonymized (such as the total amount of money spent), these auralizations represented a significant investment by TP, not affordable in usual projects. Time was seen as the most limiting factor to the use of auralizations during P8. The *Time* category contained 20 coded segments, representing 13.5% of the encountered difficulties, mostly from the email exchanges (see Table 5). The GA computation time was discussed, which can take several days depending on the Algorithm used and the complexity of the GA model, but most often the Time factor concerned last-minute demands or unanticipated deadlines due to the client, which is a common constraint in such architectural projects. During the presentation to the client, half of the exchanges concerned the negotiation with the client. This latter mentioned 10 times the need to respect the program and the defined budget, as illustrated here: “*I think you as designers, me as architect and client, have to have a discipline regarding the budget.*” This was particularly prominent for the Auditorium (see Table 5), where the cost of the addition of the structural isolation would have increased the budget significantly.

Compatibility: incompatibilities appeared in this project: first, during the *Creation phase*, for the reception of the GA model: initially created in ODEON by TP, the material file was not present at the opening by SU, due to a version mismatch. A second compatibility issue concerned the 360° videos and the associated second order Ambisonic audio: while spatial audio formats are starting to be standardized in the audio industry (e.g. ADM or SOFA), the formats used in this project (.tbe for audio, .mkv for video) were still evolving between hardware and software architectures. For instance, no stable 360° video with HOA audio player was readily available at the time for iOS devices or for the Oculus Go, requiring to buy another software.

Reliability: from a technological point of view, it has been shown that a significant amount of problems can occur, due to the *New technologies* and *Compatibility* factors, respectively representing 23% and 14.2%, a total of 37.2% (55 occurrences) of the encountered difficulties. These included the following sources of problems, or unreliability: audio and video formats (.tbe, .mkv unsteady video formats or codecs), Ambisonic channel ordering, the absence of adapted 360° audio-video reader on tablets, application contents' tree view dependent on the Operating System (meaning the architecture is dependent on the OS). Regarding the realism of the auralization, precision in the process of creation and calibration of the models gives confidence in the produced aural results. The need to achieve realism was prominent during the review phase (27.9% of the codes related to *Auralization design*; RT specifications were given with the aim to have a realistic space for instance). The *influence of the stimuli* appeared to be crucial for the realism of the simulations (≈29% overall, Table 5), but is also a key element that helps put in evidence particular elements, for instance the influence of a male versus female voice in the perception of the effect of the

low-frequency absorbers (SJ-TP, P2): *“This is good because we show that the AQflex have an effect on the low frequencies of the male voice.”*

Discussion

Previously obtained results⁹ showing the diverse types of use of auralization were confirmed in this usability case study. Two different types of auralizations were realized, driven by two different objectives: the first aimed at assessing the influence of different options of acoustic treatment in a large glass Atrium, while the second aimed at showing the impact of an intrusive noise in a smaller Auditorium, both spaces being multi-purpose. In both cases, previously identified uses² were observed, with situated details regarding the creation of auralization and interaction with the client. This study, however, only concerns a specific project, and results should be validated on several other projects.

The analysis of the temporality of the project led to the definition of three phases: creation, review, and presentation phases, each encompassing different uses. The creation phase necessitated numerous email exchanges including various information and data exchange (from GA and visual models, to material absorption frequency curves, or background noise measurements curves). The review phase consisted of many informal perceptual evaluations to reach the most plausible simulations. The scenario of presentation was also a significant consideration in the two review meetings. Finally, the presentation to the client turned out to be very different for the two auralizations. The Auditorium resulted in a lively discussion based on budget restrictions, while the Atrium auralization was beneficial as it started a dialog and let the client really understand the impact of the acoustic design choices.

The choice of appropriate stimuli and precise settings were required to achieve the desired and needed realism. Existing anechoic stimuli were used, and recordings of ambient noise were performed to simulate intrusive noise from a sub-space in the auralized Auditorium. Stimuli were chosen for their capacity to highlight the efficiency of the proposed acoustic treatments in the Atrium and to put in evidence the annoying intrusive noise in the Auditorium in order to inform the client on the risks of discarding the proposed acoustic treatment. Hence, the selection of the stimuli was driven both by the aim of the auralization and by the intended uses of the space, as similarly reported in previous experiences of acoustic consulting using auralizations.¹ An additional aspect to reflect on, in the context of auralization and the choice of stimuli, are the differences and similarities between the use of auralization in (fundamental) research and in architectural projects. Regarding the selection of stimuli, Kuusinen²³ proposed a method to produce a large set of stimuli in a systematic way, providing a representative sample of sounds (which include the different sources of variance) to evaluate a set of factors of auditory perception in a more controlled experiment. On the other hand, at least in the context of the case study, there are additional factors to take into account concerning the choice of the stimuli. The situation here involves a negotiation aspect between the consultant and the client, and consequently the stimuli are partly chosen to bring pleasant emotions to the client, which facilitates the negotiation phase, or at least can be considered as an asset to support the convincing/persuasion stakes (including the choice of acoustic design material or other design options). This is in contrast to research experiments where there are potentially a large panel of listeners, and the main stake is to understand the influence of a given factor that we vary, as discussed in Kuusinen.²³ There is also a compromise between (1) the representativeness of the selected stimuli, which is a critical point, (2) the selected acoustic configurations (source and listener placement, absorption material), and (3) the level of experience of the client, and the degree of familiarity between this latter and the acoustic consultant—who, if he or she knows the client well, is able to arbitrate auralization choices to better drive the demonstration—based on what he or she wants to

highlight in terms of acoustic characteristics. The attention of the client can also be oriented towards other aspects such as visual features. In this way, the selection of stimuli is limited by the constraints that exist in architectural projects, in contrast to the greater freedom available when designing auditory experiments. To conclude, there is probably an incompatibility between the desired representativeness of stimuli and the constraints of architectural projects, which requires compromises and arbitrarily selected stimuli.

The use of auralization in these cases enabled the client to directly experience the space and understand better the direct impact of the acoustical design choices. The actual presentation to the client allowed a pedagogical guide through the different acoustic conditions, supported by verbal explanations of the acoustic configuration they were presently listening to. This pedagogical role of auralization was also mentioned by the acoustic consultant, who would use it with the support of 360 videos rendered for a portable VR HMD, in formation, or during conferences on music and architecture.²⁴ However, it should be noted that current auralization tools are not designed to integrate this pedagogical component; they are built for technicians/acousticians, and require a non-negligible technical background to be able to produce them. This pedagogical function could therefore drive the conception and design of the next generation of better integrated auralizations tools.⁵

On the other hand, the observation of the technical aspects (the creation process) revealed that much improvements or optimizations are needed, as many difficulties were encountered due to these unsteady technologies, and the lack of an ecosystem that integrates all the needed components (Visual, VR, IR simulations, Convolutions). The *Learnability* of this technology remains an impediment to its integration, rendered difficult by the diversity of tools, software, and programming languages required to produce high-quality auralizations. These are amplified by the regular updates that need to be followed, and the constant evolution of tools, notably the rapid evolution of VR technologies (e.g. newly available HMDs yearly). Standardization of 3D formats (both for visual models and for audio) is yet to be reached, and wastes time through incompatibilities. The adoption of auralization by acoustic consultants, which remains quite low,⁹ may benefit from the increasingly democratized VR and AR technologies.^{25–28} Tech companies (GAFAMI/BATX) will also play a key role in their diffusion with the role of media, fashion, and technology.²⁹ In architecture, *Revit* has integrated VR in its options, and *Unity* has announced the integration of BIM information directly linked with real-time updates in addition to a much better integration of AR devices via *FoundationAR*.

Conclusions

This work presented a usability case study of auralization use, conducted in collaboration between an acoustic consultant and an acoustic research team, which was observed and described based on the analysis of email exchanges and audio recordings of meetings. Effective uses were categorized in two families: *Auralization design* and *Client presentation*. The *Auralization design* family encompassed *GA simulations*, *Acoustical design*, *Sound design*, the need to *Achieve Realism*, *Multi-purpose hall*, *Visual models*, and *Stimuli influence*. The *Client presentation* family encompassed *Listening*, *Negotiation*, *Pedagogy*, and *Scenario setting*. These uses were distributed differently during the creation, review, and presentation phases identified during the project.

The analysis of practical acceptability highlighted the unpredictable character of events in such projects, whether due to material issues, human errors, or organizational difficulties. Difficulties were highlighted due to the novelty of auralization tools, whether software or hardware, producing compatibility issues. The multiplicity of tools required to produce these (especially VR) auralizations increases the complexity of the task and the range of skills needed. Auralization technology

would therefore benefit from the homogenization of VR tools, formats, and 3D software, as well as a straightforward or unique ecosystem enabling the production of high-quality VR auralizations within the architectural (acoustics) community. Furthermore, pedagogical functions could be integrated into such tools. Acousticians and architects would benefit from training on the still emerging technologies of VR/AR and spatial audio, with for example, specific classes, tutorials, or workflow guides.

Auralization adoption is highly dependent on the adoption of VR technologies by the architectural community. Efforts are needed to deploy them by informing the community of the potential benefits. Collaborations between acoustic consultants and researchers would be a good starting point, and would benefit both, allowing for sharing infrastructure facilities and knowledge transfer, while at the same time providing researchers with the possibility to conduct in situ studies and have access to large-scale projects, furthering our understanding of the uses and their evolution following technology developments.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was conducted as part of a PhD funded by Paris-Sud University.

ORCID iDs

David Thery  <https://orcid.org/0000-0001-9431-0605>

Brian FG Katz  <https://orcid.org/0000-0001-5118-0943>

References

1. Azevedo M and Sachs J. Auralization as an architectural design tool. In: *EAA joint symposium on auralization and ambisonics*, Universitätsverlag der TU Berlin, Berlin, 18 March 2014, pp.162–168.
2. Hochgraf K. Auralization as a tool for acoustical design of restaurants and public spaces. *J Acoust Soc Am* 2017; 141: 3476.
3. Stowe LM and Golob EJ. Evidence that the lombard effect is frequency-specific in humans. *J Acoust Soc Am* 2013; 134(1): 640–647.
4. Pätynen J, Pulkki V and Lokki T. Anechoic recording system for symphony orchestra. *Acta Acust United Acust* 2008; 94: 856–865.
5. Milo A. The acoustic designer: joining soundscape and architectural acoustics in architectural design education. *Build Acoust* 2020; 27: 83–112.
6. Woksepp S and Olofsson T. Credibility and applicability of virtual reality models in design and construction. *Adv Eng Inform* 2008; 22(4): 520–528.
7. Castronovo F, Nikolic D, Liu Y, et al. An evaluation of immersive VR system for design reviews. In: *13th International conference on construction applications of VR* (eds N. Dawood and M. Kassem), Teesside University, London, 30–31 October 2013, pp.22–29.
8. Atkins J. *Bridging the gap between reality and virtuality: an analysis of VR technology adoption within architectural practices*. Master's thesis, Plymouth University, Plymouth, 2017.
9. Thery D, Boccara V and Katz BFG. Auralization uses in acoustical design: a survey study of acoustical consultants. *J Acoust Soc Am* 2019; 145(6): 3446–3456.
10. Bobillier-Chaumon E. Conditions d'usage et facteurs d'acceptation des technologies de l'activité: questions et perspectives pour la psychologie du travail. Habilitation à diriger des recherches, HDR, Ecole Doctorale Sc. de l'Homme, du Politique, et du Territoire, Grenoble, 2013.

11. Thery D, Katz BF and Poirier-Quinot D. Auralization creation: collaboration between Sorbonne Université and theatre projects. Research Report (HAL). <http://hal.archives-ouvertes.fr/hal-02511092hal.archives-ouvertes.fr/hal-02511092>, Sorbonne Université (2020).
12. Thery D and Katz B. Anechoic audio and 3d-video content database of small ensemble performances for virtual concerts. In: *International congress on acoustics*, Aachen, Germany, September 2019, vol. 23, pp.739–746. International Commission for Acoustics (ICA).
13. Nielsen J. *Usability engineering*. Burlington, MA: Morgan Kaufmann, 1994.
14. Barcenilla J and Bastien JM. L’acceptabilité des nouvelles technologies : quelles relations avec l’ergonomie, l’utilisabilité et l’expérience utilisateur ? *Trav Hum* 2009; 72(4): 311–331.
15. Werner Adeelman-Larsen N, Jeong CH and Stofringsdal B. Investigation on acceptable reverberation time at various frequency bands in halls that present amplified music. *Appl Acoust* 2018; 129: 104–107.
16. Japan Audio Society, GES 9080. Various, CD compilation. Rights society: JASRAC. In: Impact, CD compilation. 1985.
17. Braun V and Clarke V. Using thematic analysis in psychology. *Qual Res Psychol* 2006; 3(2): 77–101.
18. Boyatzis RE. *Transforming qualitative information—thematic analysis and code development*. New York, NY: SAGE Publications, 1998.
19. Gilbert LS. Going the distance: closeness in qualitative data analysis software. *Int J Soc Res Methodol* 2002; 5(3): 215–228.
20. Dalenbäck BI. What is geometrical acoustics? *CATT-Acoustic Whitepaper*. CATT-Acoustics, Gothenburg (Sweden), 2018, pp.1–10.
21. Katz BFG and Wetherill EA. Fogg art museum lecture room, a calibrated recreation of the birthplace of room acoustics. In: Forum acusticum, European Acoustics Association, Budapest (Hungary), 2005, pp.2191–2196.
22. Postma BN and Katz BF. Perceptive and objective evaluation of calibrated room acoustic simulation auralizations. *J Acoust Soc Am* 2016; 140(6): 4326–4337.
23. Kuusinen A. An anechoic audio corpus for room acoustics and related studies. In: *Proceedings of the EAA joint symposium on auralization and ambisonics*, Universitätsverlag der TU Berlin, Berlin, Germany, 3–5 April 2014, pp.113–118.
24. Brandā L, Amaral D, Daniel W, et al. Virtual reality as a tool for teaching architecture. In: Aaron M, Wentao W, (eds) *Design, user experience, and usability: designing interactions*. Cham: Springer International Publishing, 2018, pp.73–82.
25. Bellini H, Chen W, Sugiyama M, et al. *Profiles in innovation: virtual and augmented reality*. The Goldman Sachs Group, Inc. New York, 2016. pp.1–30.
26. Cohen L, Duboé P, Buvat J, et al. *Augmented and virtual reality in operations*. Capgemini Research Institute. Paris, 2018. pp.1–36.
27. Lee J, Kim J and Choi JY. The adoption of virtual reality devices: the technology acceptance model integrating enjoyment, social interaction, and strength of the social ties. *Telematics Inform* 2019; 39: 37–48.
28. De Klerk R, Duarte AM, Medeiros DP, et al. Usability studies on building early stage architectural models in virtual reality. *Autom Constr* 2019; 103: 104–116.
29. Herz M and Rauschnabel PA. Understanding the diffusion of virtual reality glasses: the role of media, fashion and technology. *Technol Forecast Soc Change* 2019; 138: 228–242.
30. Woirgardt M, Stade P, Amankwor J, et al. *Anechoic recordings*. Cologne University of Applied Sciences. Institute of Communication Systems, Cologne (Germany), 2012.
31. MAXQDA. Verbi software. MAXQDA 2020 [Computer software]. Berlin, Germany, <http://www.maxqda.com> (2019, accessed 9 September 2021).