TOWARDS AN EX-SITU ECOLOGICALLY VALID SOUNDSCAPE EVALUATION METHODOLOGY USING IMMERSIVE AUDIO

Giorgos Dedousis
Laboratory of Music, Acoustics & Technology (LabMAT)
National and Kapodistrian University of Athens, Greece
gdedousis@music.uoa.gr

Areti Andreopoulou
Laboratory of Music, Acoustics & Technology (LabMAT)
National and Kapodistrian University of Athens, Greece
a.andreopoulou@music.uoa.gr

ABSTRACT

The design of soundscape evaluation protocols, realized either in-situ (real-world) or in laboratory conditions, and the assessment of their ecological validity is gaining interest among researchers. This paper reports on an experiment that aims to simulate the listening conditions of interacting with a travel-guide application presenting users with recreations of historically-informed soundscapes, designed to be used in-situ, in a laboratory setting. This was achieved by simulating the real-world soundscapes, which would have surrounded users of the app when used outside, from a periphonic speaker arrangement in a sound-treated space using an 2D ambisonic reproduction of the corresponding data. The evaluation results showed that participants interacted with the travel-guide content in a manner that resembled a real-life situation. That is, they were in general able to recognize and separate the auditory content of the travel-guide from that of the real-world simulated auditory space they were placed within, with some occasional confusions. This study adds to the work on the ecological validity of soundscape evaluations and tries to highlight the boundaries between the two simultaneous soundscape auditions.

1. INTRODUCTION

Since its original introduction in 1969 [1], the term Soundscape has grown to encompass various scientific research and creative applications. It refers to the foreground and background acoustic energy present in a landscape, relative to the orientation and position of the listener in that space [2]. The soundscape of a physical environment has been shown to have a stronger impact on people’s mood and a restorative effect, than its visual qualities do [3]. In general, natural sounds are suggested to promote health and well-being, albeit to a different degree, among people from different sociocultural backgrounds [4]. Furthermore, the acoustic qualities of indoor spaces have been found to affect people’s mental health either positively or negatively depending on their demographics, working from home conditions, and other factors, as shown in research conducted during the COVID-19 lockdown [5, 6]. Such results offer a chance to rethink the design of urban and indoor areas in terms of acoustic design.

Audio-walks, guided-tours, and virtual-audio tours are examples of soundscape design and composition. Audio-walks and guided-tours are a form of outdoor/ indoor exploration, augmented via auditory information often presented over headphones. They are becoming a ubiquitous way to experience history and heritage of any form, and are often associated with visitor experience in museums, archaeological sites, galleries, etc. [7]. Similarly, virtual-audio tours offer users the opportunity to virtually travel and experience the soundscape characteristics of various places in the world in an immersive manner [8], enriching their awareness of the plethora of unique auditory qualities that can be found around the globe, from the convenience of their homes.

The ecological validity of soundscape reproduction in laboratory conditions has been examined in terms of auditory scene recreation and reproduction as well as assessment methodology and user evaluation data collection. The term ecological validity, when introduced in the context of auditory perception, refers to the extent to which findings from a laboratory experiment can be applied in real-world situations [9, 10]. The process requires choosing representative participants and realistic experimental conditions while simulating real-life situations, in order to evoke similar perceptive reactions [11].

Virtual and Augmented Reality technologies can potentially address the limitations of laboratory spaces and enhance efficiency in soundscape assessment and noise control, allowing the realization of experiments in non-acoustically treated rooms, designed for multiple simultaneous users, thus forming a framework for future optimization of soundscape designs, which can improve their acoustical and overall environmental quality [12].

2. STATE OF THE ART

The first step towards the standardization of soundscape-related research was taken in 2014 with the release of ISO 1291301:2014 [13]. Since then, more standards have been released dealing with data collection protocols and reporting [14] and accurate analysis and interpretation of the results [15]. The collection of subjective assessments of soundscapes in their original location (in-situ) can be time-consuming and complex logistically-wise. As a result, research has long been exploring alternative evaluation methodologies, using portable systems [16], and soundscape reproduction technologies in laboratory settings [17].

Research in soundscape reproduction has been extensive, especially when spatial audio technologies are used [18]. Such tech-

https://doi.org/10.21785/icad2023.4479
nologies have the advantage of offering greater control over variables compared to in-situ studies, but at the possible risk of introducing artifacts and cognitive experiences that are different to those perceived in real-world conditions [19]. To that end, it has been demonstrated that, quality assessments of "overall pleasantness" of in-situ soundwalks do not differ significantly to those collected in laboratory (ex-situ) settings [20].

The binaural reproduction of soundscapes has been found to lead to systematic qualitative and quantitative assessments [21], and to higher quality experiences in terms of "realism", "reverberance", and "directivity" [22]. The augmentation of binaural reproduction with active head-tracking has been also found to lead to more accurate perception of soundscape characteristics, despite the limitations introduced by the use of non-individualized HRTFs [23]. In general, soundscapes experienced in-situ are being perceived as more pleasant, possibly because of the complementary visual information present, while ex-situ reproductions are preferred for more targeted analyses of soundscapes. These observations seem to be indifferent to the level of expertise of the assessors [24]. Similarly, because ambisonic reproduction prioritizes immersion over localization of sound sources [25] the auditory experience of soundscapes in laboratory conditions tends to be similar to in-situ experiences [26, 27]. Hence it can be argued that, immersive audio can be confidently used for soundscape reproduction and evaluation [19].

Immersive Virtual Reality has also been suggested as an alternative, ecologically valid soundscape reproduction technology. Most such systems employ dynamic rendering using head-mounted displays, tracking, and spatial audio reproduction [10]. Nevertheless, researchers remain in disagreement as to whether in such a setting the auditory modality is less [28] or more significant than the visual one [29], or even whether the two contribute to the experience independently [30].

This paper extends past research on the ecological validity of ex-situ soundscape evaluations, through an experiment that aims to simulate the listening conditions of interacting over headphones with a travel-guide application, intended to be used in-situ, in a laboratory setting. This was achieved by recreating the real-world soundscape, users would have experienced in-situ, from a periphonic speaker arrangement in a sound-treated space, using ambisonics, and running a user-evaluation study of the travel-guide auditory content within that setup. The topology and content of the soundscapes, the evaluation study, as well as its results, are presented and discussed in the following sections.

3. SOUNDSCAPE DESIGN

TRACCE 1 (TRavelogue with Augmented Cultural & Contemporary Experience) was a national collaborative research project that led to the design of a novel mobile platform for cultural route exploration. With this application users can explore routes of historic interest while having access to both past and contemporary information, hence enriching their experience through the inevitable comparison between historically documented and contemporary data. This work falls within the scope of Informal Learning Environments (ILEs). As such, the created content must be interesting, appropriate, and informative for a wide audience of various age groups, educational backgrounds, and sensory abilities. In order to address this issue, the designed auditory content (soundscapes) for the app followed a multi-layer topology, introduced in 3.2.

3.1. Data Collection

Extensive research is necessary in order to effectively design historically informed soundscapes. That is soundscapes which imitate the auditory environment of specific landscapes as predicted to have existed in the past. In the case of the TRACCE project, our main source of information was the published notes and diaries (travelogues) of travellers, who had visited these places in the past. These texts were scanned for direct descriptions or indirect information of the auditory characteristics of each soundscape. Additional sources of information included iconography and sketches of the specific locations of interest dating back to that era, contemporary photographs and cartography of the same locations today, as well as any other sources of information we could find on these locations. These additional sources were necessary as they provided invaluable information on the morphology of each place and its acoustic characteristics [31].

The audio used for the design of the historically informed soundscapes was based on selected samples from publicly available databases and newly recorded material. Each soundscape was designed in such a way to acoustically match and describe excerpts from the travelogues. These travelogue excerpts were also narrated, when needed, to complement the auditory content. The designed experience, which consisted of soundscape compositions and narrated text excerpts, was further enriched with abstract musical elements, composed from stochastic interpretations of the remaining auditory content. Each historically informed soundscape design was conceived in such a way so as not to exceed 90 seconds in duration.

3.2. Soundscape Topology

Soundscapes were designed for audition over headphones featuring 4 layers of auditory content with which users can interact independently or in any combination they prefer or need. Each layer contains different types of auditory content and carries different levels of information [31]. The Background layer comprises sounds that reconstruct the main auditory content of the landscape (such as wind, air etc.) and the background noise of the travelogue writer. The audio track is stereophonic. The Storytelling layer consists of sounds introduced in the environment by the traveler’s interaction with the surrounding area, as well as other sounds which are directly related to the story. This audio track is binaural. The Narration layer consists of a monaural audio track with narrated excerpts from the original text that are directly related to the action of the story, yet difficult or impossible to communicate only through the auditory reconstruction of the soundscape. The Music layer contains a composition of electroacoustic music which is always based on a stochastic interpretation of the remaining 3 layers of each soundscape, ensuring a smooth integration of all auditory content. This track is also rendered binaurally.

4. EVALUATION

The aim of this study was to explore the possibility of designing an ex-situ evaluation experiment for the assessment of soundscapes, which would be ecologically valid. As previously mentioned, in-situ evaluations can be influenced by both ambient sound interference and other extraneous factors [19], while laboratory simu-
lations, if designed properly, can lead to results and conclusions that are transferable and relevant to the in-situ experience. Thus, a framework was developed within which listeners would experience ex-situ the auditory content of the TRACCE application over headphones (as they would have if using the app in the real-world) while being immersed by a recorded excerpt of the corresponding real-world soundscape, presented through a periphonic speaker configuration, imitating the in-situ (real-world) experience.

The experiment took place in an acoustically treated space at the facilities of the Laboratory of Music Acoustics and Technology (LabMAT), of the Music Department at NKUA. 25 individuals (N=25) participated in the experiment (14 female), aged from 20 to 45 years old (mean: 23, std: 6.31), all undergraduate or graduate students of the Music Department. Eight KRK ROKIT RPG2-5 loudspeakers, equally space on a circular configuration of 1.7 m radius, were connected to a Steinberg UR-824 audio interface on a Windows laptop. This periphonic speaker setup was used for the ambisonic reproduction of the Contemporary soundscapes, achieved using the SPARTA audio suite.

Two additional laptops (one per participant), each connected to a Steinberg UR-22c audio interface, were available for listening to the historically informed soundscapes of the TRACCE mobile application. Two listeners at a time were situated at the acoustic sweet-spot of the aforementioned loudspeaker configuration, facing away from each other, and were asked to complete the evaluation experiment. Participants were instructed to bring their own preferred set of headphones which they would normally use if they were to use the TRACCE app in-situ. The historically informed soundscapes were reproduced over headphones through a Max/MSP patch (Fig. 1), while excerpts of ambisonic recordings of the corresponding real-world contemporary soundscapes were playing back over the surrounding speakers. The Max/MSP patch offered users control of the auditory content played back over their headphones (the TRACCE app historically informed soundscapes), and tools for its evaluation. Each participant evaluated two soundscapes, selected randomly from a collection of four.

Figure 1: Soundscape user-evaluation interface, designed in Max/MSP.

The average duration of the evaluation process was approximately 30 minutes.

5. RESULTS

The first evaluation task concerned the identification of the historically informed (app based) soundscape locations, out of four possible answers: (i) soundscape next to / inside the sea, (ii) soundscape next to / inside a river, (iii) rural area soundscape, and (iv) indoor soundscape. Participants could select one or multiple answers through the provided Max/MSP interface. As can be seen in Fig. 2, participant responses were relevant, appropriate, and in most cases anticipated, fact which strengthens our ecological validity hypothesis for the assessment process.

More specifically, the app-based soundscape of Chania was mostly identified correctly as a "soundscape inside / next to the sea".

Chania is a city on the northwest coast of Crete, known among others for its historic harbor and its stores and street market on the waterfront.

Figure 2: Distribution of responses to the question “where do you think this soundscape is located at”.

https://doi.org/10.21785/icad2023.4479
Figure 3: Per soundscape location venn-diagrams depicting the relationship between the sets of sound-sources identified in the historically informed and real-world soundscapes. Red color refers to the sounds present in the former soundscapes, blue color to those in the latter, purple to those present in both, and black (peripheral to the venn diagram) to the ones that were misidentified and were not present in either. Font size is proportional to word frequency.

"sea", suggesting that the sea sound was easily perceivable and possibly predominant. A small percentage (∼8%) of the participants seems to have mistaken the sea for a river, which may be an indication of the need for sound-design improvement, while others (∼31%) had chosen the "rural area soundscape" option, which might suggest that, to them, the sea was not the prevailing sound in that soundscape.

The characterization of the Samaria Gorge\textsuperscript{3} soundscape was the one that had led to the highest degree of disagreement among participants. The answers were almost equally divided between options i (∼31%), ii (∼31%), and iii (∼38%). The latter would have been the obvious choice for this famous gorge location. Yet, the sound of the distant river present in the designed soundscape seems to have prevailed, and sometimes even to have been mistaken for a sea sore, fact which had led to the selection of the other two locations too.

Even though the Kolymvari\textsuperscript{4} soundscape was expected to be the most easily categorizable, being a monastery next to the sea sore, as either option i, or iv, or even both, only one of the participants opted for multiple answers. Most of the participants (∼82%) seem to have ignored the fact that a significant portion of the action takes place inside a church, where one could clearly hear monks chanting as the outside world sounds fade away, and focused on the sea-related sounds present at the beginning of the soundscape. A significant percentage of those answers (∼46%),

\textsuperscript{3}The Samaria Gorge, located in southwest Crete, is a National Park of Greece. It was created by a small river running between the White Mountains and Mt. Volakias. The gorge is 16 km long, starts at an altitude of 1,250 m at the northern entrance and ends at the shores of the Libyan Sea in Agia Roumeli.

\textsuperscript{4}Kolymvari is a coastal town at the southeastern end of the Rodopou peninsula on the Gulf of Chania. It is known for its Greek Orthodox monastery Gonna, dedicated to the Assumption of the Virgin, which is situated next to the sea-sore.
though, concerned option ii: "soundscape next to / inside a river", which is a strong indication of a need for possible sound-design improvement.

The Melidoni Cave 3 soundscape was also mostly correctly identified as an indoors location. The occasional water-drops from the stalactites inside a small pond must have been mistaken for a river by some of the participants. Interestingly enough, while there exist strong auditory and acoustic cues throughout the duration of this audio example indicating that this soundscape is taking place indoors, such as absence of outdoors / nature-related sounds, absence of wind, high reverberation times etc., approximately one third of the participants had chosen to focus only in the very beginning of the soundscape, which describes the entrance to the cave, and selected iii as their only answer. This behavior remains to be further investigated.

Later on in the evaluation process, participants were asked to identify the most prevailing and characteristic sounds in the historically informed soundscapes, reproduced over headphones, while, the excerpts from recorded real-world soundscapes from the corresponding physical locations in modern-day Crete were being reproduced from the surrounding loudspeakers, virtually placing assessors at a realistic replication of an in-situ audition state. The intention behind this task was to observe whether there was a clear cognitive separation between the two co-present soundscapes, or whether auditory information from the soundscape imitating contemporary real-word audition conditions was being "leaked" into the auditory soundscape experience of the travel application. User responses, per evaluated location, are presented in Fig. 3 as word-clouds in venn diagrams, depicting identified sound sources that belong to the historically informed and real-world contemporary soundscapes, as well as their intersection. Misidentified words appear outside the two circles. As can be observed, participants were in general successful in correctly identifying characteristic sound sources of the historically informed soundscapes and in isolating the contemporary auditory content.

More specifically, for the Chania soundscape, approximately half of the participants had successfully identified the sea, wind, and boat sounds as prevailing sound sources, with some misinterpreting the former for running water. However, a large percentage noted that they had heard music, speech, footsteps, and children’s voices, all of which were sound sources that belong to the modern, real-world soundscape, which was being reproduced over the loudspeakers. A small minority had also mentioned the presence of loud percussive sounds and traffic noise, both of which were also part of the real-world auditory content. The misidentified sound sources, depicted in the top-left plot of Fig. 3 in black, even though not really present in any of the two presented soundscapes would natural fit in their context. For example the clack, airplane and garbage track sounds would very naturally fit into the real-world soundscape context, while bird sounds could easily be a part of the historically informed soundscape, designed for the TRACCE mobile application.

The Samaria Gorge (top-right plot of Fig. 3) soundscape consisted of only a few characteristic sound sources, as most of its content was ambiences. The most commonly identifies sources were the wind and footsteps and the occasional sounds of running water from afar. These sources can be heard in both the historically informed and the contemporary, real-world soundscape. The sound of birds singing, which was indicated by a small percentage of the participants was only present in the latter, as the hike in the historically informed soundscape one was taking place during the winter, when birds are not a common encounter in the gorge. A small minority of the assessors had also mentioned hearing rain and sea sounds, which actually were not present in either of the two soundscapes.

According to the assessors, the most prevailing sounds for the Kolymvari location (bottom-left plot of Fig. 3 were the wind, trees / leaves, the sea, and running water, all of which were correctly identified sound sources present in both the historically informed and the contemporary real-world soundscapes. Chanting horses and carriage sounds were also correctly identified by most of the participants as content of the historically informed soundscape. However, some participants also indicated that they had heard the sounds of birds, footsteps, speech, traffic and chimes, all of which were only present in the real-world soundscape. Once again the misidentified audio sources (river, thunder, puddle, anchor, and rain) could very easily fit context-wise in either of the two soundscapes.

For the Melidoni Cave soundscape, a lot of participants had indicated the presence of sound sources which were not in reality part of neither the historically informed nor the real-world soundscapes. Most of the participants had correctly identified footsteps and the wind, present in both soundscapes, along with water-drops, only present in the historically informed content. Some participants had also noted the presence of speech, which was in reality only part of the real-world soundscape. Out of the many misidentified sound sources, which can be seen in the bottom-right plot of Fig. 3, especially of interest is the reference to a "subway ambiences", which was probably mentioned due to the distinct highly reverberant acoustic conditions in this indoors space.

As a final task, participants were asked to assess the auditory correlation between the historically informed and contemporary, real-world soundscapes. Before completing this task, they were asked to create their preferred mix of the 4 layers in the historically informed soundscape, which were then not allowed to change. Since they had all used their preferred set of headphones, each offering a different degree of isolation from the surrounding world, participants were given the option to state that they could not hear any sound, which was indicated by a small percentage.

---

3The cave is located 1.8 km northwest of the village Melidoni, in Rethymno Crete. The arched entrance to the cave leads to several cave chambers filled with stalagmites and stalactites some of which are not currently open to visitors.

---

https://doi.org/10.21785/icad2023.4479

Figure 4: Correlation between the auditory content of the historically informed and contemporary real-world soundscapes.
the real-world soundscape content. The collected responses can be seen in Fig. 4.

The correlation between the two soundscapes of Chania as well as that between those of the Samaria Gorge location were characterized as average or good. For the case of Chania, the historically informed and the contemporary, real-world soundscapes differed greatly due to the vast differences between the historic and contemporary auditory content of this urban location. As a result, an average (38%) or good correlation assessment was rather justified and expected. While one would have expected that the correlation between the historically informed and the real-world soundscape in the Samaria gorge would have been very high, due to the minimal changes in the landscape over the decades, a difference in seasons (the former soundscape was designed to replicate a winter day, while the real-world one was recorded during the late summer) between the two had severely impacted this link, leading to an average correlation characterization by 69% of the participants.

For the Kolymvari location, the correlation between the two soundscapes was mostly characterized as either average (37%) or poor (28%). This comes to no surprise, as the two soundscapes have very little auditory content in common (see Fig. 2). This can be attributed to the fact that nowadays next to the monastery exists a high-traffic road, which was not present in the historic landscape. Finally, the Melidoni Cave soundscapes seem to lead to average (36%) and good (36%) correlation ratings due to the fact that the soundscape of the cave has remained mostly unchanged over the decades.

6. CONCLUSIONS AND FUTURE WORK

The aim of this study was to extend past research on the ecological validity of evaluations in controlled environments such as a laboratory. Through a subjective assessment task it was demonstrated that the interference between virtual reconstructions of real-world auditory content simulating an in-situ listening experience and the TRACCE mobile application auditory content of historically informed soundscapes was equivalent to those reported in similar studies in the literature. Controlled ex-situ auditions of soundscapes can lead to a valid way of listening to and interacting with auditory content. Participants in this study were not only able to recognize and, frequently, separate the auditory content of the travel-guide from that of the virtual space they were placed into, but also to correlate the two in an individual manner.

More specifically, participants had managed to separate the historically informed from the contemporary, real-world auditory content they were simultaneously presented with over head-phones and speakers respectively, with occasional albeit anticipated "leaks" between the two. This can be attributed to the fact that, the two were oftentimes highly correlated, from an auditory content perspective, sharing auditory information, and, thus, seamlessly blending into one another.

The correlations between the historically informed and contemporary, real-world soundscapes, which were in general characterized by the participants as above average, suggest that the designed soundscape content, although mainly emanating from 18th and 19th century travelogues, matches harmonically with the corresponding contemporary real-world soundscapes.

This work takes a step towards an ex-situ ecologically valid soundscape evaluation methodology. However, more experiments need to be done predominantly with a larger number of participants and potentially a 3D speaker configuration [26]. A comparison between in-situ and ex-situ user reactions, perceptions, and expectations would also be very informative towards this goal.

7. ACKNOWLEDGMENT

8. REFERENCES


