SPATIO-SONIC SITE SURVEY USING AMBISONICS RECORDINGS AND VR RENDERING – DESIGNING SOUNDSCAPES IN THE EARLY PLANNING PROCESS

Kristin Nedlich[1], Arne Nykänen[1], and Björn Hellström[2]

[1]Engineering Acoustics, Luleå University of Technology, 971 87 Luleå, Sweden
[2]RISE – Research Institutes of Sweden

Correspondence should be addressed to Kristin Nedlich (`Kristin.nedlich@ltu.se`)

ABSTRACT

As a part of the on-going city move in Malmberget, Gällivare, the inhabitants of Malmberget mining area need not only new housing but also increased communal services. To accommodate this, the municipality is building a new public building called the Multiactivitiy building. The building will house public services such as: library, sports centre, swimming pools and recreational spaces. The central lobby space creates a hub which connects the different parts of the building and serves as a communal meeting place. The proposed building site, and the site of this study, is in the centre of downtown Gällivare. The study focuses on how ambisonics recordings can be used as a part of the site survey in the early stages of the architectural design process. The spatio-sonic qualities of the existing site are integral to the architectural design process, not only as documentation and analysis of the site's conditions, but as an integrated design parameter (on par with lighting, climatic, contextual, and other site-specific conditions). A Virtual Reality model of the site, auralized with ambisonics recordings, is the main artefactual outcome of this study. Additional documentation of the site survey includes spatio-sonic mapping of the site plan, written (autoethnographic) impressions from the main surveyor, photographs, ambisonics recordings and a summary and analysis of respondents' answers to questionnaires in combination with soundwalks. The conclusion of the study is that it is feasible to produce simple, virtual, auralized models that can be used as a base for sketching with sound in architecture.

1 Introduction

Sound and VR in the architectural practice

The spatio-sonic qualities of the existing, outdoor, site could serve the architectural design process in a positive way. Not only as documentation and quantitative analysis of the site's context but as an integrated design driver (on par with lighting, climatic, contextual, and other site-specific conditions) [1].

Virtual Reality is a tool that is on the rise in research on soundscape design for architecture. The aim of this study is to explore the possibility of employing a Virtual Reality model, auralized with ambisonics recordings, for spatio-sonic site surveys as well as

using the Virtual Reality model as a tool in which the designer will be able to listen-move-listen [2] in an early, critical stage of the architectural design and sketch process.

The survey in the study consists mainly of, real and virtual, soundwalks with local respondents. A soundwalk can be summarized as a reflexive research methodology in which active listening, and perception, are given priority [3]. This study explores how ambisonics recordings, presented in a VR-model are perceived when compared to the real-life experience of a soundwalk.

The architectural sketching process
The architectural design process relies to a great extent on iterative sketch making. In that process, ideas are generated and tested. Questions relevant to the design problem are posed and either discarded or resolved through design.

Keeping it simple
To work as a thinking tool, the sketching needs to be quick, easy, cheap, and ambiguous [4]. Typically, sketching is thought of as the use of pen and paper. In architectural practice, a sketch could just as well be a three-dimensional volume, readily made up of scrap material such as left-over cardboard for example.
When sketching with sound and space, the temporal and fleeting nature of sound makes it difficult to work with analogue tools. Therefore, a digital tool such as Virtual Reality, might be used as the "cardboard model" of sound sketching.

Research questions
How can virtual soundwalks, based on ambisonics recordings, be used as a tool for spatial sound design in the early stages of the architectural design process? How do a real and virtual soundwalk correlate in terms of ecological validity, perception and immersion? What are the limitations and benefits of recording ambisonics for spatio-sonic site surveys?

2 Method

Experimental design
The study design combines the ISO standard for Acoustics – soundscapes, part 2 [5], with an architectural mindset on site surveys. The ISO standard triangulates a soundscape by combining quantitative data, such as acoustic properties, with qualitative data, such as questionnaires and interviews, psychoacoustic indices, and information about the context [6] [7].

Soundscape study method
The soundscape method of conducting a site survey is not that different from the architectural method, but it puts more emphasis on the creative synthesis of on-site observations, contextual analysis (spatial, historical and future) interviews with local residents while balancing pragmatic needs with aesthetic values.
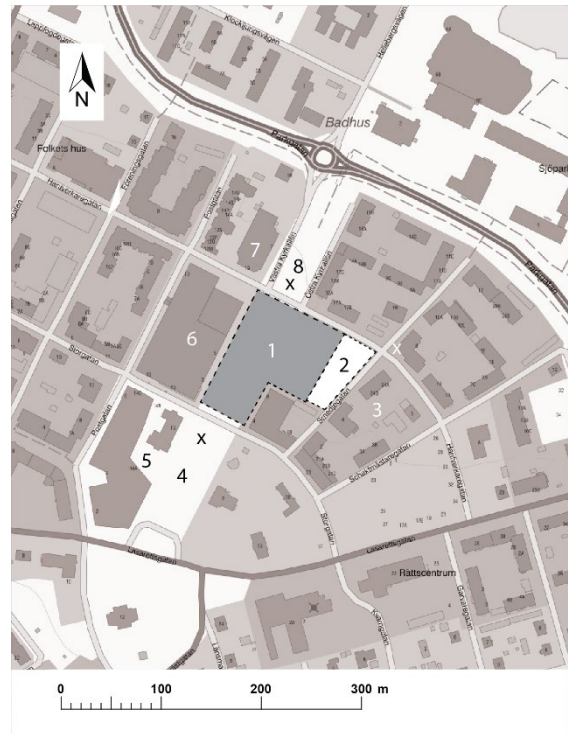

Figure 1. Spatio-sonic triangulation of the site.

1 – the Multiactivity building footprint. 2 – proposed parking lot. 3 – building site of new apartment buildings. 4 – temporary parking lot, planned public park and recreational space. 5 - school and museum buildings. 6 – shopping mall. 7 – Systembolaget. 8 – existing parking lot. X – listening spot/s

Listening spots
Three places, triangulating the site spatially and visually as well as sonically, were chosen as designated listening spots for the soundwalks. The listening spots were chosen based on the focal points of the planned building's main entrances and parking lot. See Figure 1.
The spots were far enough from each other to offer a variety of sonic space, as well as a difference in function, spatial and urban configuration. The spots have been documented and recorded in 360° photos and first order ambisonics recordings. See Figure 2

The equipment used was a Røde NT-SF1 microphone with a Zoom F6 multitrack field recorder. The recordings were done without any respondents being present, but at a similar time, day and weather conditions. The length of the recordings were no less than three minutes and the surrounding SPL levels were measured simultaneously.

The virtual model was rendered in SketchUp and imported into Unreal Engine. The detailing of the 3D model was intentionally kept as simple as possible, in line with the cardboard model theory. See Figure 3

Virtual model

The virtual model was auralized with the ambisonics recordings from the first part of the study.

The Ambisonics recordings were converted from A-format to first order B-Format (AmbiX) using the SoundField by Røde VST plug-in (V. 1.0.2). The binaural synthesis was made using the Resonance Audio plug-in (V. 1.0) in Unreal Engine 4.23. The sound was reproduced using Sennheiser HD600 headphones.

Respondent selection

In total, seven respondents participated in the study. All were volunteers with little or no experience in architectural design practice or 3D modelling. All respondents were permanent residents of Gällivare for at least 5+ years.

Most of them had very good knowledge and experience of the listening spots, in particular the public space in spot #1, and the parking lot at spot #3. The age range was 45-70 years and five were women. All participants had self-reported normal hearing. None of the respondents had any training or experience in active listening, acoustics, architecture or similar.

Study execution

The study was divided into two parts. Between Part One (the real soundwalks) and Part Two (the virtual soundwalks) there was a gap of approximately two weeks during which the virtual soundwalks and ambisonics recordings were produced. Two respondents, out of seven, were not present at the time of the first part, but instead they completed the real soundwalk in conjunction with the virtual. All participants completed the real soundwalk before commencing the virtual.

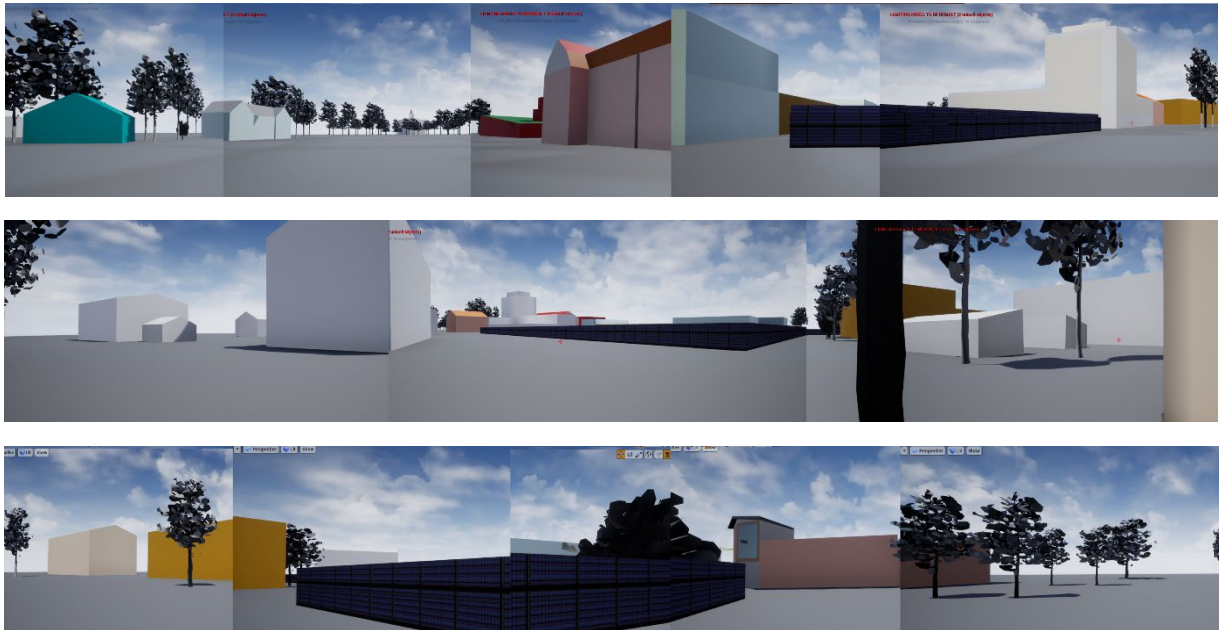Figure 2. 360° photo of listening spots IRL

Figure 3. 360°virtual renders of listening spots #1-3

Instructing the respondents

In the real soundwalks the main surveyor escorted the respondents to the designated spots. If there were more than one respondent, they were asked to keep conversation to a minimum even while travelling between the spots.

At the designated spot the respondents were instructed to stay within a certain limited area, in order not to deviate too much from the exact spot of the ambisonics recordings. They were also asked to, if possible, find a personal space where they would feel comfortable closing their eyes and just listen to the space for approximately half a minute.

3 Data collection

To compare the virtual soundwalk with the real, the respondents were given the same questions for both walks. Their answers have been collated to discern discrepancies and/or commonalities of the real and the virtual soundwalks.

Questionnaire and Interviews

A questionnaire was constructed, based on specified questions in the ISO 12913 standard. A few additional questions were added to collect background information, such as age, hearing deficits and their familiarity with the spaces. The questions in the questionnaire were translated from English to Swedish, as all respondents were native Swedish speakers.

Weather Conditions

The site is currently a building lot, surrounded by other building projects in progress. In the first part of the study, with the real life soundwalks, the weather was sunny, around 13 degrees Celsius and low to moderate wind. For the participants that conducted the real soundwalk in conjunction with the second occasion, climatic conditions were slightly worse with temperatures below 10 degrees Celsius and moderate to strong wind.

Virtual Reality setup

The second part of the study, the virtual soundwalks, took place in the municipality showroom for the city move - Re-Form. The space was a small conference room with a reasonably good sound insulation. When standing outside the room, with the door closed, it was not possible to hear conversations inside. Some noise from outside the room was audible. The most prominent low frequency noise was from weights dropping, at a public training facility in the building. Moderate traffic noise from the adjacent streets was also audible. The most prominent high frequency noise was from young children visiting Re-Form's exhibition hall.

Sound Level calibrations

Sound pressure levels in the model were calibrated using an artificial head (Head Acoustics HMS IV). Other studies indicate that sounds in a virtual setup are perceived as stronger than in real life [7]. This was clearly noticeable in the preliminary audition of the ambisonics recordings; therefore, the SPL-levels were set slightly lower (~3dB) in the model than what was measured on site.

4 Results

The questionnaire included seven questions for each listening spot. Of those, five questions focused on listening perception. Four of those five were answered in full by the respondents, both when listening in IRL (in real life) and in Virtual Reality (VR). At the end of the questionnaire, the respondents could give an overall impression, in free text, of the three listening spots compared to each other.

Data analysis

The answers to Question #1, *"To what extent can you hear the following types of sounds?"*, are reported in Figure 4-6. The figures show a comparison between IRL and VR per sound source for the three listening locations.
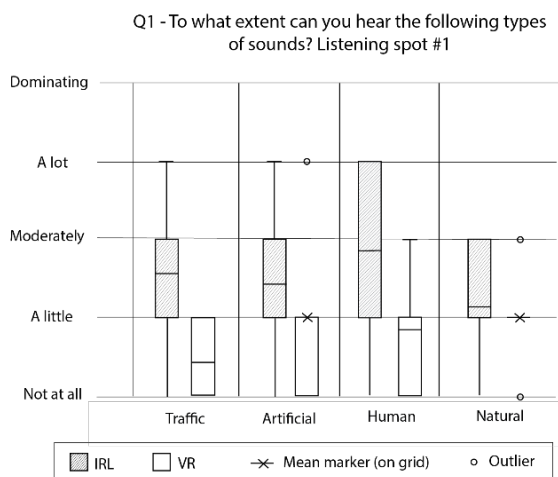

Figure 4. Q1 - Comparison of sound sources IRL and VR in listening spot #1.
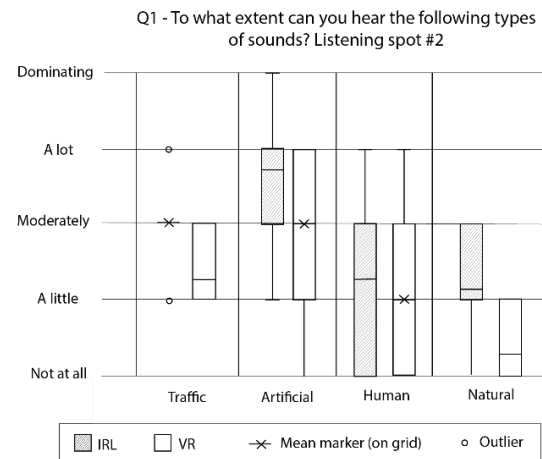

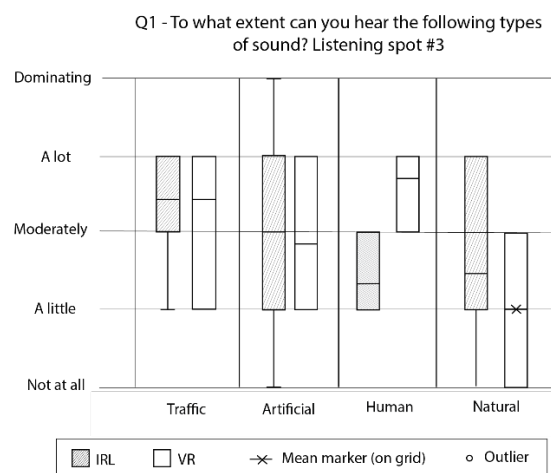Figure 5. Q1 - Comparison of sound sources IRL and VR in listening spot #2.

Figure 6. Q1 - Comparison of sound sources IRL and VR in listening spot #2.

Answers to Question #2 – *"For each of the following 8 descriptions, to what extent do you agree that the sonic environment in this place is…pleasant, Chaotic, Vibrant, Uneventful, Calm, Annoying, Eventful, Monotonous?"* - a question probing "perceived affective quality", based on the SSID (Soundscape Indices Protocol), or SSQP (Swedish Soundscape Protocol). It was not answered in full, neither in IRL or in VR. In real life only a few respondents either missed a few of the categories by accident or didn't feel like they had any response to give.

The most problematic situation for the questionnaire method, and in particular question #2, was in VR. The questions took a long time to read aloud, and some were therefore skipped, or summarized. In this case they are summarized by the respondents' comments in the interview part. Mean values of all three listening spots are illustrated in the same chart. See Figure 7
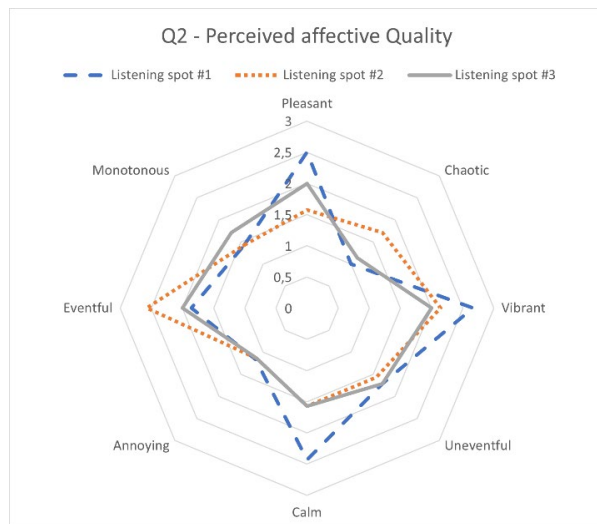

Figure 7. Q2 – Mean perceived affective quality in listening spot #1-3.

For the same question, in the Virtual Reality part of the study, the respondents' comments on perceived quality were noted to gain insight in what they were perceiving. Interestingly, a couple of respondents changed their response from one extreme to another when going from IRL to VR.

Question #3 – *"Overall, how would you describee the current sonic environment? Very bad, Bad, Either or, Good, Very Good?"* - estimating the quality of the sonic environment. The respondents were overall consistent in judging the sonic quality, when comparing IRL to VR. See Figure 8
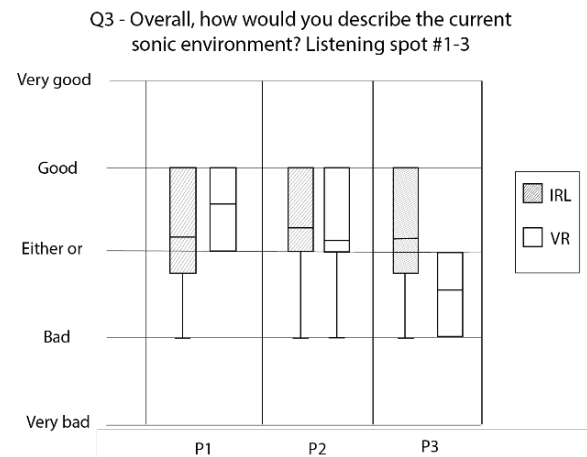

Figure 8. Q3 - Comparison of sound environment quality IRL and VR in listening spot #1-3.

Questions #4-5 *"How often do you visit this spot?"* and *"Would you like to visit it again?"*

These questions were intended as background information on how familiar the respondents were with the spot. Overall, most respondents had visited spot #1 and #3 often and would also like to visit again. Listening spot #2 was less frequently visited and for those that did visit it used it more as a transitional space, such as commuting by car, bike, or foot. The answer to whether they would like to visit again was consistent with what they replied in Q#4. These questions were not asked in the Virtual Reality part.

Question #6 – perceived loudness of the listening spot. The respondents were asked how loud they perceived the three different listening spots to be, both in real life and in Virtual Reality. Listening spot #3 was considered the most eventful, but at the same time as a cause for safety concern with a lot of traffic in a tight space. Noteworthy is that one respondent changed their perception on spot #3 from low to high on loudness, eventfulness, vibrance and chaotic moving from IRL to VR.

Overall, spot #3 was considered the worst both in IRL and VR, while spot #1 was perceived as the most calm and tranquil both in IRL and VR.

Question #7 – *"List, no more than eight, sounds that you can identify at this listening spot.",* both IRL and in VR.

In the real life experience the respondents tended to list more sounds/sound sources and labelled them as

coming from specific, identified sources. In the VR model, the specified sounds tended to be fewer than in IRL and described more as unidentifiable noises. One common confusion was to label a beeping noise from a drainage pump in the construction site as the noise from a car backing up, or a loud ventilation exhaust noise in the public square with strong wind. Those who had identified the correct sources of the sound IRL could, in general, also correctly identify them in the virtual model.

5 Discussion

Interviews
The same questionnaire, as in the first part, was given by the surveyor to a respondent while they were in the virtual model. This was to try and mimic the real-world situation as much as possible where a respondent would continuously be in the sonic environment while answering the questionnaire. This was not a perfect setup as some respondents had difficulties hearing, or rather listening, to the surveyor while wearing the VR gear even though open headphones were used.
A possible way to improve this would have been to have the questionnaire on-screen in the virtual model, with as simple a setup as possible for answering the questions. Just like the first part of the study, a digital questionnaire would hopefully ensure that all the questions would be answered in full. In hindsight, structured interviews should perhaps have been conducted in conjunction with the first part of the study as well, in order to keep the same modus operandi throughout both parts of the study.

Respondents' comments in interviews
When in the VR-model, respondents were allowed to speak freely between answering the questions from the questionnaire. The spontaneous reactions were often verbalized or clearly noticeable in their physical posture and attitude.
Here are some of the comments that was written down verbatim during the sessions.

"This is what one might imagine the spots would sound like, rather than that they sounded like that in real life."

"The sounds feel closer [in VR than IRL]."

"It was the same sensation of calm in spot #1 as in real life."

"I hope there will be more life in spot #1 in the future. It is called the Dead Square since they stopped cars from passing through the centre square [in the late 1980's]."

"I would have wished for more animated sound sources, and people, in the virtual model."

Sources for error
Unreal Engine draws a lot of computational power when played in editor mode. This causes the laptop to turn the ventilation fans up so high that they are clearly audible for the respondent using the VR headset. Although most respondents seem to have been able to sort out the noise, it is still possible that it might have masked some of the key background noises in the virtual environments. For future experiments, the computer should not be placed in the same room where the experiment is done.

The recorded sounds from the listening spots were limited in time to three-minute loops. This makes the results dependent on the events that occurred at that exact moment. However, in real life the respondents did not spend much longer than that in the listening spots, and the experiment had no control over sound events during that time.

One major difference in the virtual model was that the respondents were teleported between the listening spots, instead of traveling both in space and time (walking) between them. This might have had an impact on the perceived changes in the environment when the ears were not exposed to the slower sonic transition in between the listening spots.

Some of the extreme shifts of the perceived quality in certain listening spots might be due to sounds appearing to be closer when listening to them over headphones. The sounds can be perceived to be more present and louder than they feel like in real life. This is particularly evident in listening spot #3 with traffic noise relatively close to the recording spot.

Working in the field
The equipment used is currently not intended for field work. The ambisonics recorder and microphone is relatively light but not easy and simple to use on the go.

Ultimately, such equipment will tend to be left behind in favour of smaller, compacter and less heavy and complicated ones. The downside of such light equipment is that recordings will then serve merely as

memory aid and documentation rather than usable recordings of higher quality.

6 Conclusions

Indications of the study
Although this is a small, preliminary study, the results indicate that a simple, virtual, auralized model can be used as a representation of a site. Keeping in mind that there might have been significant discrepancies in the sonic events between the real and virtual soundwalks, the essence of the spaces was still quite clear and consistent.

Ecological validity
The perception of the same spot in real life and in virtual reality can be considered reasonably similar and achieve a degree of ecological validity. However, recordings of a site may be inadequate, or misrepresentative of the space's sonic quality given the limitations in the length of the recordings. This may happen with site visits in real life as well, which is why it is important to visit a site more than once, and on different occasions of the year, day, and weather.

Negotiating usefulness with accuracy
A more accurate representation of a site could possibly be achieved by using interactive and dynamic sound environments in combination with the acoustic qualities of the space. This is, as of yet, not easy enough to produce to keep the workflow simple and fast. The effectiveness of using virtual reality as a sketching tool with architecture and sound will have to be negotiated with representation and ecological validity.

7 Further work

To further investigate the usefulness of spatio-sonic site surveys and how architects could work with virtual models as tools, in which it is possible to sketch spaces with sound, the next step is to engage respondents from the architectural practice as respondents. The spaces affecting the sound and being affected by the sound will also be more closely examined in detail from a morphological, architectural, spatial and functional perspective.

In future work, the soundscape taxonomy in the ISO standard will need to be adapted to spatio-sonic morphology, in order to narrow the gap between traditional architectural practice and sound-

architecture. Such a study should be conducted with respondents who are architects or with equivalent experience of working with designing living environments.

8 Acknowledgements

References

[1]   J. Llorca-Bofi et al, "Generation of Architectural Designs Using Soundscapes: an Educational Case Study", *DAGA 2019 Rostock* 701 (2019).

[2]   A. Nykänen et al, "Sketching Sounds – Kinds of Listening and Their Functions in Design", *Design Studies* Vol 39, No. C, July (2015)

[3]   A. Radicchi, "A Pocket Guide to Soundwalking", *Universitätsverlag der TU Berlin,* (2017)

[4]   W. Buxton, "Sketching User Experiences: Getting the Design Right and the Right Design", *Elsevier Science* (2007)

[5]   "Acoustics – Soundscape - Part 2: Data collection and reporting requirements", *ISO/TS 12913-2:2018(E),* (2018)

[6]   B.M. Brooks et al, "Exploring Our Sonic Environment Through Soundscape Research & Theory", *Acoustics Today | Winter 2014* (2014)

[7]   A. Mitchell et al, "The Soundscape Indices (SSID) Protocol: A Method for Urban Soundscape Surveys – Questionnaires with Acoustical and Contextual Information", *Appl. Sci.*, *10*(7), 2397; https://doi.org/10.3390/app10072397 (2020)

[8]   T. Oberman et al, "Using Virtual Soundwalk Approach for Assessing Sound Art Soundscape Interventions in Public Spaces", *Appl. Sci.10*(6), 2102; https://doi.org/10.3390/app10062102 (2020)