7 - 12 SEPTEMBER 2014, KRAKÓW
FORUM ACUSTICUM

Perception of room size and the ability of self localization in a virtual environment. Loudspeaker experiment

Marko Horvat
University of Zagreb Faculty of Electrical Engineering and Computing, Zagreb, Croatia
Kristian Jambrošić
University of Zagreb Faculty of Electrical Engineering and Computing, Zagreb, Croatia
Hrvoje Domitrović
University of Zagreb Faculty of Electrical Engineering and Computing, Zagreb, Croatia
Juraj Francetić
Brodarski Institute, Zagreb, Croatia
Monika Rychtáriková
KU Leuven, Leuven, Belgium
Vojtěch Chmelik
Slovak Technical University, Bratislava, Slovakia

Summary
The paper presents the results of two experiments carried out in the Auralization Laboratory at the University of Zagreb as a part of a research conducted in cooperation with KU Leuven. The experiments were designed to investigate the ability of a normal sighted person to assess room size and localize its own position within the room based only on available acoustical information, i.e. the room response to predefined stimuli. A total of 36 listeners took part in each experiment. For assessment of room size, four virtual rectangular rooms were defined: a small room with a volume of 30 m^3, a medium-sized room and an elongated hallway of 252 m^3 and a large room of 2016 m^3. The investigation of self-localization was performed in a medium-sized room only. Acoustic treatment of the rooms was manipulated, i.e. the average absorption coefficient of a room and scattering coefficients of certain faces in a room were changed in a controlled manner. Two types of sources were used: a series of hand claps as a stationary source and footsteps as a moving source. All virtual environments were defined in ODEON® room acoustics software and the necessary sound files were encoded for 2^{nd} order Ambisonics reproduction. The results of the experiments are statistically analyzed to determine whether or not the abilities of self localization and assessing room size can be confirmed at a given level of statistical significance.

PACS no. 43.55.Hy, 43.66.Lj

1. Introduction

For a normal sighted person, auditory information is a valuable complement to visual information received from the environment. A blind or visually impaired person is forced to perceive their environment with no visual information available. For such a person, auditory information becomes the main source of information about that environment. With no visual information to process, a blind person can process auditory information more thoroughly, and extract information about the environment a normally sighted person would usually ignore.

The difference in perception of the environment by normal and visually impaired people, or even the difference between early- and late-onset blind people has been studied extensively [1-4]. The ability of self-localization by echolocation is also an investigated topic [5, 6].

The work presented in this paper is an extension of the research done by Chmelik [7], who has investigated self-localization and room size assessment using auditory cues extracted from acoustic responses of virtual environments. The work described in this paper was based on recreating these experiments, but in a virtual acoustic environment recreated with a multi-channel loudspeaker system, rather than using headphones.

2. Experimental setup

The experimental investigation described in this paper was divided in two parts. The first part was focused on the ability of self-localisation inside a closed space, i.e. the listeners were asked to assess their position in a room based on its response to given sound stimuli. The second part dealt with the listeners' ability to differentiate between different room sizes, again based solely on the response of the room to a sound stimulus.

2.1. Stimuli and scenarios

Two different sound stimuli were used in the experiments; hand claps produced by a stationary virtual listener and footsteps produced either by a moving virtual listener or by another virtual person moving inside the room.

Different scenarios were devised by changing the values of absorption coefficient of the walls and the ceiling in the room; specifically, the chosen values were 0.1, 0.2 and 0.4. The floor was left perfectly reflective in all scenarios. For each value of absorption coefficient, the scattering properties of the surfaces were varied in three steps by changing the scattering coefficient, as follows: 0.05 on all surfaces, 0.9 on the ceiling and 0.05 on the remaining surfaces, and 0.9 on a side wall with 0.05 maintained on other surfaces.

All environments used in the investigation were defined, simulated and auralized in ODEON room acoustics modelling software. Further processing and sound reproduction were performed with Reaper audio software package. The auralization was made for 2nd order Ambisonics reproduction. The reproduction level was kept constant throughout the duration of the experiment. All recordings were RMS-normalized to the same level, without going into clipping.

2.2. Listeners

A total of 36 listeners took part in these experiments, with their age ranging from 21 to 28. About 70 % of the listeners are familiar with acoustical concepts to a certain level, and they have gained this knowledge through interest in music and/or courses on acoustics taken at the faculty. About 10 % of the listeners had good knowledge on acoustics or musical education.

2.3. Procedure

All tests were performed in the Auralization Laboratory at the University of Zagreb, equipped with a multichannel loudspeaker system in a quasi-spherical 4-8-4 configuration, making it capable of handling up to 2nd order 3D Ambisonics recordings and up to 3rd order 2D Ambisonics recordings.

The listening tests have been divided in two stages due to sheer quantity of test material the listeners were asked to listen to, with 18 evaluations in each stage. The listeners were instructed on where to sit and how they have to be oriented, so that their orientation would match the one of the virtual listener, as defined in the simulations. The light in the laboratory was dimmed, so that the listeners would not be influenced by anything from their real environment. However, a certain amount of illumination was necessary so that the listeners would be able to see the questionnaire and the images of the rooms provided for them as the supplement to oral and written explanations.

The task defined for the listeners was to listen and evaluate the sequences of recordings for different acoustic treatments (defined absorption and scattering) of a room or rooms. Specifically, in the self-localization experiment the listeners listened to recordings on all three positions (A, B and C) for a given acoustic treatment of the room and were asked to assign the positions in the room to the recordings they heard, thus yielding one of six possible permutations (e.g. BAC). In the room size assessment experiment the listeners listened to recordings in all four rooms for a given acoustic treatment, and then were asked to assign room numbers to the recordings they heard, thus yielding one of twenty four possible permutations (e.g. 3124).

Before any actual evaluation, a training sequence was reproduced to the listeners with solutions, so that they would understand the nature of their task. After that, the listeners proceeded with the case-by-case evaluation for each of the nine different acoustic treatments and with two different sound stimuli in each of the experiments.

2.4. Self-localization

The self-localization experiment was performed in a virtual room 12 m long, 7 m wide and 3 m high, having a volume of 252 m^3. In the first part of the experiment the virtual listener stands still in one of

the three predefined positions in the room, as shown in Figure 1, and claps his hands to produce the sound stimulus. Six consecutive claps are used as the sound stimulus. The height of the virtual listener's ears was set to 1.7 m, and the source, i.e. his hands were set to a height of 1.5 m and 0.7 m in front of the listener's position. In this case the wall with adjustable scattering is the left side wall.
The reverberation times in the room are shown as a function of frequency in Table 1 for different absorption coefficients of relevant surfaces, i.e. all but the floor.

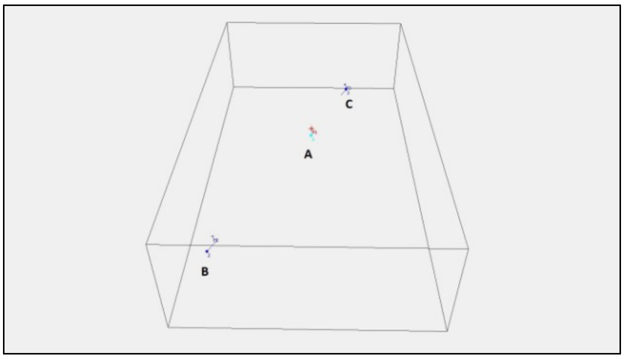

Figure 1. Room 1: medium-size room, dimensions 12 m x 7 m x 3 m, volume 252 m3; setup for the first part of the self-localization experiment: A = central position, B = corner position, C = frontal position

Table 1. Reverberation times in room 1 vs. absorption coefficients of relevant surfaces

RT_{60} (s)	α ( )		
f (Hz)	0.1	0.2	0.4
63	1.98	0.97	0.41
125	1.97	0.97	0.41
250	1.95	0.96	0.41
500	1.92	0.95	0.41
1000	1.88	0.94	0.41
2000	1.77	0.92	0.40
4000	1.46	0.82	0.38
8000	1.07	0.60	0.32

In the second part of the self-localization test the listeners attempted to determine their position in the room by listening to the sound of footsteps of a virtual walker walking along the length of the room, as shown in Figure 2. The length of an individual step is 0.6 m and the height of the sound source was set to 0.1 m. The listener is oriented towards the middle of the room and does not follow the movement of the virtual walker. The listening positions were modified with regards to the first part of the experiment. In this case the wall with adjustable scattering is again the left side wall (the farthest wall in Figure 2).

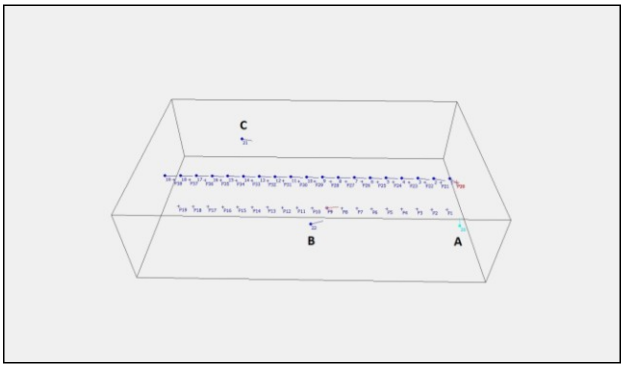

Figure 2. Room 1: setup for the second part of the self-localization experiment: A = corner position, B = central right lateral position, C = back left lateral position

2.5. Room size assessment

The experiment on room size assessment was performed in four different virtual rooms. Room 1 is the same one used in the experiment on self-localization. Room 2 is a long hallway, as shown in Figure 3, with dimensions 35 m x 2.4 m x 3 m, and the volume of 252 m^3, the same as in room 1. The reverberation times are shown in Table 2.

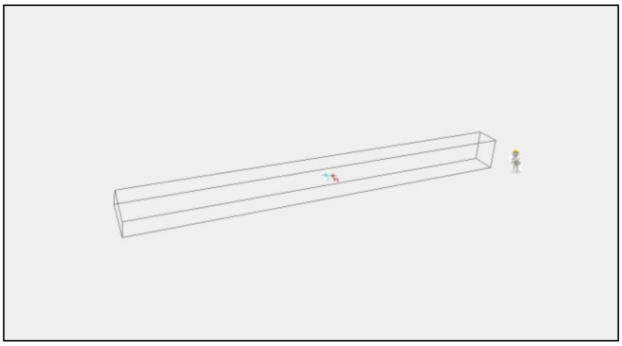

Figure 3. Room 2: hallway, dimensions 35 m x 2.4 m x 3 m, volume 252 m^3

Table 2. Reverberation times in room 2 vs. absorption coefficients of relevant surfaces

RT_{60} (s)	α ( )		
f (Hz)	0.1	0.2	0.4
63	1.27	0.61	0.28
125	1.27	0.61	0.28
250	1.26	0.61	0.28
500	1.25	0.61	0.28
1000	1.23	0.60	0.28
2000	1.18	0.59	0.28
4000	1.03	0.55	0.27
8000	0.70	0.44	0.24

Room 3 is a small room with dimensions 4 m x 3 m x 2.5 m, yielding the volume of 30 m^3, as shown in Figure 4. The reverberation times are shown in Table 3.

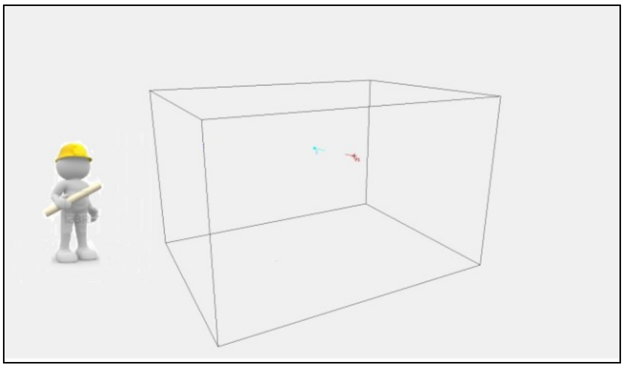

Figure 4. Room 3: small room, dimensions 4 m x 3 m x 2.5 m, volume 30 m^3

Table 3. Reverberation times in room 3 vs. absorption coefficients of relevant surfaces

RT_{60} (s)	α ( )		
f (Hz)	0.1	0.2	0.4
63	0.98	0.47	0.21
125	0.98	0.47	0.21
250	0.97	0.47	0.21
500	0.97	0.47	0.21
1000	0.96	0.47	0.21
2000	0.93	0.46	0.21
4000	0.83	0.44	0.20
8000	0.60	0.36	0.19

Finally, room 4 is a large room obtained from room 1 by doubling all its dimensions, as shown in Figure 5, yielding the final room dimensions of 24 m x 14 m x 6 m and the volume of 2016 m^3. The reverberation times are shown in Table 4.

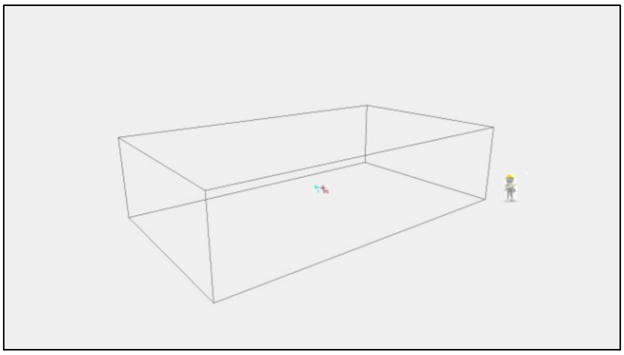

Figure 5. Room 4: large room, dimensions 24 m x 14 m x 6 m, volume 2016 m^3

Table 4. Reverberation times in room 4 vs. absorption coefficients of relevant surfaces

RT_{60} (s)	α ( )		
f (Hz)	0.1	0.2	0.4
63	3.70	1.77	0.99
125	3.67	1.77	1.15
250	3.60	1.75	0.95
500	3.50	1.73	0.85
1000	3.35	1.70	0.79
2000	3.02	1.60	0.75
4000	2.20	1.34	0.69
8000	1.09	0.83	0.52

3. Results and discussion

The initial analysis of the results was performed using the X^2-test. The null hypothesis used in the tests is that the listeners can neither recognize the positions in the room nor differentiate between various room sizes based only on acoustic response of the room. In other words, in the self-localization experiment, all six permutations obtainable from position markers A, B and C (e.g. ABC, ACB, BAC,...) are equally probable. Analogously for the room size assessment experiment, all twenty four permutations obtainable from room markers 1, 2, 3 and 4 (e.g. 1234, 2134, 4321, 3412, ...) are equally probable as well. The summarized results of this analysis are shown in Tables 5 and 6.

Table 5. The results of the statistical analysis for the self-localization test

Hand claps		Scattering coefficient ( )		
	$df = 5$	all 0.05	ceiling 0.9	wall 0.9
Absorption coefficient ( )	0.1	$\chi^2 =$ 23.18	$\chi^2 =$ 6.94	$\chi^2 =$ 4.12
		$p <$ 0.001	$p =$ 0.225	$p =$ 0.533
	0.2	$\chi^2 =$ 0.94	$\chi^2 =$ 3.41	$\chi^2 =$ 4.12
		$p =$ 0.967	$p =$ 0.637	$p =$ 0.533
	0.4	$\chi^2 =$ 7.65	$\chi^2 =$ 2.35	$\chi^2 =$ 0.24
		$p =$ 0.177	$p =$ 0.798	$p =$ 0.999
Footsteps		Scattering coefficient ( )		
	$df = 5$	all 0.05	ceiling 0.9	wall 0.9
Absorption coefficient ( )	0.1	$\chi^2 =$ 10.18	$\chi^2 =$ 4.88	$\chi^2 =$ 0.65
		$p =$ 0.070	$p =$ 0.430	$p =$ 0.986
	0.2	$\chi^2 =$ 20.41	$\chi^2 =$ 8.06	$\chi^2 =$ 10.53
		$p =$ 0.001	$p =$ 0.153	$p =$ 0.062
	0.4	$\chi^2 =$ 2.41	$\chi^2 =$ 3.82	$\chi^2 =$ 1.00
		$p =$ 0.790	$p =$ 0.575	$p =$ 0.963

Table 6. The results of the statistical analysis for the room size assessment test

Hand claps		Scattering coefficient ( )					
$df = 23$		all 0.05		ceiling 0.9		wall 0.9	
Absorption coefficient ( )	0.1	$\chi^2 =$	197.33	$\chi^2 =$	228.00	$\chi^2 =$	162.67
		$p <$	0.001	$p <$	0.001	$p <$	0.001
	0.2	$\chi^2 =$	32.00	$\chi^2 =$	98.67	$\chi^2 =$	209.33
		$p =$	0.100	$p <$	0.001	$p <$	0.001
	0.4	$\chi^2 =$	120.00	$\chi^2 =$	102.67	$\chi^2 =$	73.33
		$p <$	0.001	$p <$	0.001	$p <$	0.001
Footsteps		Scattering coefficient ( )					
$df = 23$		all 0.05		ceiling 0.9		wall 0.9	
Absorption coefficient ( )	0.1	$\chi^2 =$	119.29	$\chi^2 =$	163.17	$\chi^2 =$	179.63
		$p <$	0.001	$p <$	0.001	$p <$	0.001
	0.2	$\chi^2 =$	142.60	$\chi^2 =$	141.23	$\chi^2 =$	130.26
		$p <$	0.001	$p <$	0.001	$p <$	0.001
	0.4	$\chi^2 =$	116.54	$\chi^2 =$	89.11	$\chi^2 =$	138.49
		$p <$	0.001	$p <$	0.001	$p <$	0.001

The X^2-test shows only the differences between the expected and the observed overall situation. It does not show specific information. In this case the test shows only that the listeners are able to differentiate between different conditions in a room, but it will not show whether their assessment is accurate or not. Furthermore, the test applied to room size assessment data does not meet all the requirements on the dataset. Therefore, additional data on the results of the listening experiments is given in graphic form in Figures 6 to 9.

The figures show the distribution of the percentage of answers given by the listeners in listening experiments for each acoustical treatment. The correct answers, i.e. the correct sequence of positions in the self-localization experiment and the correct sequence of rooms in the room size assessment experiment are marked with a red column. All other (incorrect) sequences are marked in black.

A fundamental difference can be observed between the results of the self-localization test and the room size assessment test, and this difference is also reflected in the results of the X^2-tests. In the self-localization test there were three positions in the room, and upon listening to a sequence of three recordings, the listeners could have given 3! = 6 different sequences of positions as their answer for a given situation. The charts shown in Figures 6 and 7 reveal that all six sequences were valid answers, in the opinion of the listeners. Moreover, apart from one or two cases, the percentage of answers does not deviate much from the expected null-hypothesis value of 16.67 %, suggesting that the listeners could not differentiate between the positions in the room and have given their answers almost randomly, as confirmed by the results of the X^2-tests. The correct sequence does not appear to be chosen more frequently than other sequences.

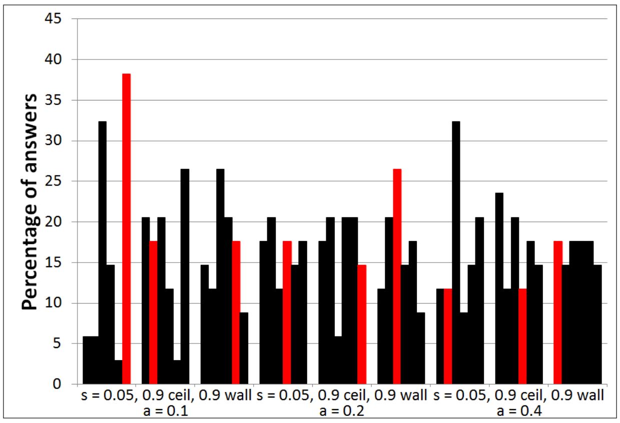

Figure 6. The self-localization experiment – hand claps as the source – expected value 16.67 %

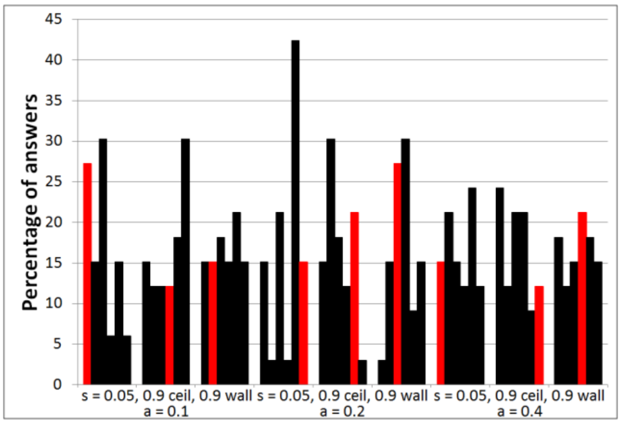

Figure 7. The self-localization experiment – footsteps as the source – expected value 16.67 %

In the room size assessment test there were four rooms and 4! = 24 possible ways to put them in order after listening to four recordings. Figures 8 and 9 reveal that not all 24 possibilities have been recognized by the listeners. Instead, their number for a given case was reduced to about one half. With hand claps as the source and low absorption in the room, the rate of correct assessment of room size starts at 40-50 percent and drops significantly with the increase of absorption in the room. With footsteps as the source, the correct room size assessment is more stable at a rate of about 25-30 percent. It is interesting to note that in many cases two or sometimes three sequences have been singled out as valid solutions, one of them being the correct sequence. The other sequence (or sequences) differs from the correct one merely in the sense that two rooms have been swapped. The room that got confused most often with other rooms is room 1, i.e. the middle-sized room, as expected, due to the fact that it has the same volume as room 2 (the hallway), and its volume puts it between rooms 3 and 4 (small and large).

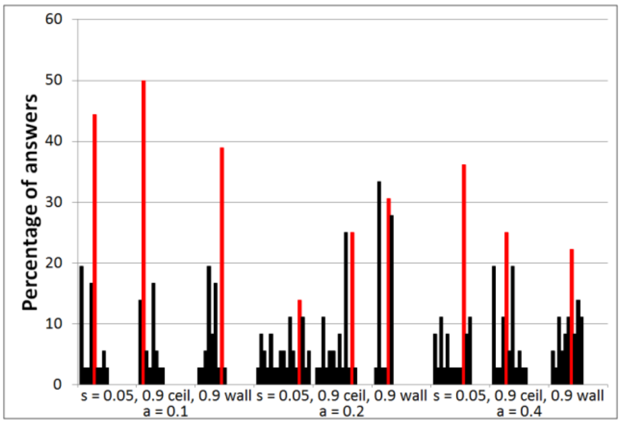

Figure 8. The room size assessment experiment – hand claps as the source – expected value 4.16 %

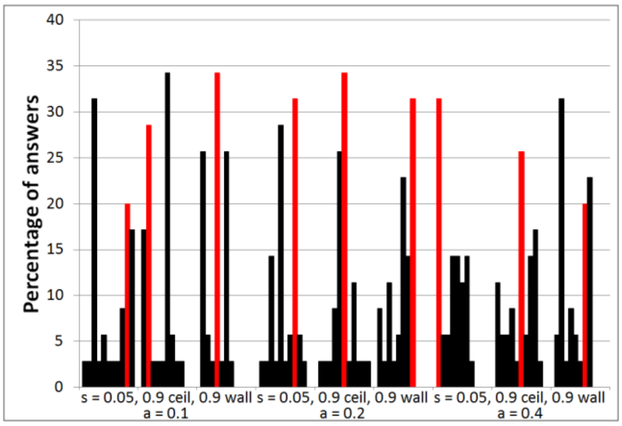

Figure 9. The room size assessment experiment – footsteps as the source – expected value 4.16 %

4. Conclusions

The results of the experiments described above show that the ability of self-localization in a room based on acoustical cues is not pronounced, while the performance in room size assessment is considerably better. In our opinion, the reason for this is the fact that over time people generally gain some experience in relating the acoustical response of a room with its size merely by using different spaces, public or other, in everyday life. On the other hand, self-localization in a room based on sound is not crucial for normally sighted persons, such as the ones that took part in these experiments, because they use their sight to do that. On the other hand, it is expected that visually impaired people will rely much more on sound and will have a better developed ability of sound-based self-localization, which remains a topic for verification in further research.

The questions that still remain are the influence of the laboratory space itself on the listeners, i.e. whether they are able to imagine themselves in a virtual environment created by sound or not. Additionally, in a real situation, hand claps are produced by people themselves, whereas in these experiments they were reproduced by the loudspeaker system, along with the response of a virtual room.

Acknowledgement

This work has been supported by the European Community Seventh Framework Programme under grant No. 285939 (ACROSS).

References

[1] J. M. Hull: On Sight and Insight. A Journey into the World of Blindness. Oxford SPCK, 1990.

[2] T. Kujala, K. Alho, M. Huotilainen: Electrophysiological evidence for cross-modal plasticity in humans with early- and late-onset blindness. Psychophysiology 34 (1997) 213-216.

[3] N. Lessard, M. Pare, F. Lepore: Early-blind human subjects localize sound sources better than sighted subjects. Nature 395 (1998) 278-280.

[4] P. Voss, F. Gougoux, R. J. Zatorre: Differential occipital responses in early- and late-blind individuals during a sound-source discrimination task. Neuroimage 40 (2008) 746-758.

[5] F. Gougoux, P. Belin, P. Voss: Voice perception in blind persons: A functional magnetic resonance imaging study. Neuropsychologia 47 (2009) 2967-2974.

[6] J. Herssens, L. Roelants, M. Rychtarikova, M. Heylighen: Listening in the absence of sight: The sound of inclusive design. Include 2011, London, 2011.

[7] V. Chmelik: Principles of inclusive design in architecture and room acoustics. doctoral thesis, Slovak university of technology, Bratislava, 2013.