The 11th International Scientific Conference
eLearning and software for Education
Bucharest, April 23-24, 2015
10.12753/2066-026X-15-079

COMPARATIVE RESEARCH ON SOUND LOCALIZATION ACCURACY IN THE FREE-FIELD AND VIRTUAL AUDITORY DISPLAYS

Oana BĂLAN, Alin MOLDOVEANU, Alexandru BUTEAN, Florica MOLDOVEANU
Faculty of Automatic Control and Computer Science, University POLITEHNICA of Bucharest, Splaiul Independenţei 313, Bucharest, Romania.
oanab_2005@yahoo.com, alin.moldoveanu@cs.pub.ro, florica.moldoveanu@cs.pub.ro, alexandru@butean.com

Ionuţ NEGOI
General Surgery Department, Emergency Hospital of Bucharest, University of Medicine and Pharmacy Carol Davila Bucharest, Romania
negoiionut@gmail.com

Abstract: *The following paper aims to present a comparative study on the audio localization accuracy (directional judgment, absolute spatial perception and rate of front-back confusions- situation in which the listener perceives the sound coming from the front as coming from the rear and vice-versa) in both free-field and virtual sound source conditions. Sound localization experiments in the free-field rely on the use of loudspeakers for delivering the auditory information to the listener. On the other hand, virtual auditory displays are based on 3D sounds (resulting from the filtering of a particular sound with the Head Related Transfer Function (HRTF) corresponding to the direction of the sound source in space) that are rendered to the listener through a pair of stereophonic headphones. 3D sounds are used in a wide range of applications, as they can simulate the perception of an external sound source in real-world hearing conditions and generally increase situational awareness. Nonetheless, they can introduce several localization errors (caused primarily by the use of non-individualized Head Related Transfer Functions), such as poor performance in the median plane (for vertical localization) and an increase in the rate of front-back confusions, especially for the directions of 0 degrees (to the front) and 180 degrees (to the rear). As a result, we intend to include in our research a comprehensive psychophysical evaluation, interpretation and analysis of the accuracy of free-field and headphones-presented stimuli localization, in order to bring to light the auditory particularities that differentiate sound localization performance under the two presented conditions.*

Keywords: *Sound localization; 3D sound; virtual auditory display; free-field sound localization; front-back confusions*

I. INTRODUCTION

Auditory spatial processing is a subject of considerable interest, mainly because of its relevance to spatial orientation, navigation and mental representation of the environment [1].

The human auditory system is truly remarkable, not only because of its structural and functional complexity, but also for its extraordinary capacity of dealing with various situations, i.e. resolving sound localization problems, disambiguating and separating simultaneous audio signals and decoding the sound based on its physical characteristics - amplitude, frequency, timbre etc. In a 3D environment, the spatial audio information can be represented as the combination of a wide range of cues (varying in intensity and temporal spectrum) that are distributed across the audible frequency bands. If we consider them separately, these cues are spatially ambiguous and impossible to discriminate within a single frequency band. Nonetheless, the auditory system is able to associate and combine them across frequencies, in order to integrate an exhaustive acoustic spatial image [2].

The act of sound localization represents the estimation of the current position of a sound source in space and is defined by a certain amount of implicit uncertainty and functional bias caused by the means of interaction with the system, the methods of directional pointing (head, hand, virtual pointing) and by other errors that lead to a degree of estimation uncertainty [3].

Localization performance in the free-field and in virtual auditory environments has been an intensive subject of study for many researchers in the last decades. In the free-field, the listeners are presented real audio signals delivered through loudspeakers. In these conditions, their ears are not obstructed or inhibited by any means and they can rely on both the binaural audio cues (for horizontal localization) and the monaural cues (which are highly dependent on the anatomical features, size, shape and orientation of the pinna – the external ear) for localization in the vertical plane and front-back directional discrimination.

On the other hand, virtual auditory environments are used to test and evaluate the sound localization accuracy and spatial auditory perception of virtual sound sources synthesized in binaural sound reproduction systems [4]. In the virtual auditory environments, the listeners are offered the same perception of sound as in real-world conditions. This is possible by using 3D sounds (synthesized audio signals which give the sensation of being present at the listening position) that are necessarily rendered through stereophonic headphones. In many sound localization experiments, the subjects are required to use a head tracking device that enables the audio signals delivered to the ears to change according to the head movements. In this way, the perception of directional sound remains unaltered and the audio localization accuracy increases, offering an efficient means of resolving front-back localization misjudgements.

The primary goal of this paper is to study and review the localization performance (sound source localization error and front-back confusion rate) in both the free-field and in virtual auditory displays. In this way, we intend to make inferences about the processes involved in sound localization, to investigate the most effective directional pointing methods and sonification techniques and to assess the impact of auditory cues on human spatial perception.

II. SPATIAL SOUND PERCEPTION IN THE FREE-FIELD AND VIRTUAL AUDITORY ENVIRONMENTS

The free-field simulation involves the presence of several loudspeakers (disposed in an array or following an arrangement rule) around the listener, in both the horizontal and the vertical planes. More often, the localization experiment takes place in complete darkness, to prevent the subjects from seeing the active loudspeaker. In a virtual auditory display, the sound allows a naturalistic and dynamic representation of the scene objects from all the possible directions in space and improves the sense of presence, immersion and realism [5] [6].

Virtual auditory displays are useful for a wide range of applications, such as air traffic control displays, teleconference environments, assistive devices for the visually impaired and audio games [7] [8] [9]. In a virtual acoustic application, the 3D binaural sound is delivered through stereophonic headphones. Headphones audio rendering is effective, as it allows complete control over the sound stimuli presented at the ears. In addition to this, a dynamic and interactive acoustic environment can be simulated by employing a head-tracking device that enables the user to move his head and sequentially hear a change in the perceived sound that corresponds to the direction of the head displacement [10].

III. SOUND LOCALIZATION EXPERIMENTS IN THE FREE-FIELD AND VIRTUAL AUDITORY ENVIRONMENTS

This section presents some of the most significant sound localization experiments performed in the free-field and in virtual auditory environments, together with their stimuli synthesis technique, training and test procedure and most relevant results in respect with localization in the horizontal and vertical planes, degree of sound externalization, system latency and front-back confusion error rate (Table 1).

Auditory localization experiment of nearby broadband sources

The experiment presented in [11] evaluates the proximal-region localization performance of four subjects. The source was located randomly within 1 m distance from the subject's head and the user was required to indicate the perceived direction with an electromagnetic position sensor. The experiment demonstrated that the azimuth error increased slightly as the sound source approached the head, although elevation accuracy did not vary with distance.

Auditory localization experiment with additional visual cues in the horizontal and vertical planes

The experiment described in [1] compares the free-field localization ability of eight sighted individuals before and after being blindfolded, in both the horizontal and vertical planes, by using two pointing methods: head and hand pointing. The results show that all the three parameters influence the localization accuracy, especially in the case of the blindfolding condition. Blindfolding conducted to an increase in the azimuth localization error for the head-pointing method, as this localization modality relies more on visual than on proprioceptive cues in order to offer an accurate directional response.

Sound source localization experiment using filtered noises

The experiment undergone in [12] presents several measurements of the localization accuracy of 45 listeners, using filtered noise bursts. The results showed that the subjects with normal hearing had a constant localization performance across all the experimental phases and responded with reduced uncertainty and localization bias, recording a root-mean-square error of 6.2 degrees and an overall standard deviation from the mean of 1.79 degrees.

Localization performance experiment of real and virtual 3D sound for use in fighter aircraft

The experiment described in [4] evaluates the localization performance of real and virtual sound sources using a 3D sound system designed for fighter aircraft communication (with short and long duration stimuli) and a head-tracking device. The results showed that the localization accuracy was better for long-duration stimuli in the free-field (for the sources located in the horizontal plane) than for virtual sound sources with variable elevation. Moreover, head movements played an important role in reducing the rate of reversal errors.

Sound localization experiment on the nature and distribution of errors by human listeners

The experiment described in [13] measured the ability of human listeners to localize a short noise burst by pointing the nose towards the perceived sound source direction. In order to familiarize the subjects with the requirements of the task, the tests were preceded by a closed-loop training session which consisted of instant user feedback about the location of the sound source.

Sound localization experiment with visual feedback training

The experiment presented in [14] proposes a sound localization experimental procedure based on perceptual feedback training for learning the correct sound source position. The results showed that the training approach significantly improved localization accuracy and reduced the reversal error rate. In addition to this, the improvements still persisted a few days after the training procedure, supporting the idea that learning, training and perceptual visual-auditory adaptation offer an efficient means to deal with the localization inconsistencies provided by the use of non-individualized HRTFs.

A psychophysical validation of headphone simulation of the free-field listening

Wightman and Kistler [15] proposed a comparative analysis of the localization accuracy for free-field and headphones listening conditions, using wideband noise bursts stimuli. For the headphones condition, the 3D sound was synthesized using the listeners' own HRTFs. The results of the experiment demonstrated that azimuth localization accuracy is comparable in both conditions, although there has been recorded an increase in the elevation error and front-back confusion rate when virtual 3D sounds have been employed.

A two-dimensional sound localization test by human listeners

The experiment presented in [2] studied the ability of subjects to localize broadband, brief sounds (150 ms) and continuous sounds in the free-field, by turning the face towards the perceived direction in both the horizontal and the vertical planes. The results showed that localization is better in the horizontal plane, especially in the front (with errors of 2-3 degrees) but worse in the lateral directions (maximum errors of 20 degrees). This is explained by the fact that the areas of maximum sound pressure for any given frequency are located in front. In addition to this, the ear is more sensitive to high frequencies; thus, the sounds arriving from behind are more attenuated (due to the shadowing effect of the head, they tend to lose their high-frequency components) compared to the frontal audio signals. The rate of change of the direction-dependent ILD (Interaural Level Difference) decreases with increasing the lateral angle. As the just noticeable difference (JND) cue is directly proportional with the value of the ILD, it becomes more difficult for the listener to decode and discriminate the sounds located on the side and in the rear hemifield, as they have a reduced ILD changing rate and a higher degree of localization uncertainty [2].

Localization experiment with non-individualized head-related transfer functions

In the experiment described in [16], several subjects were asked to indicate the perceived direction of both horizontal and vertical sound sources under the free-field and headphones conditions, by listening to short bursts of broadband noises that were convoluted with the HRTFs of a representative listener from Wightman and Kistler's experiment [15]. The subjects showed comparable horizontal localization results under both conditions. Nonetheless, there has been recorded a lower localization accuracy in the vertical plane and an increased rate of front-back confusions.

Experiment on the effects of increasing system latency for virtual sound source localization

The study introduced in [17] evaluated the sound localization accuracy of 5 listeners who were the subject of an experimental procedure which involved the directional estimation of 12 sound sources in a virtual auditory environment (using individualized HRTFs). The listeners perceived a variable sound latency that lasted from the moment of the head movement to the moment of the corresponding change in the sound delivered over the headphones. The results of the experiment showed that the minimum perceived delay is 250 ms. Nonetheless, the sound latency did not affect localization accuracy (the subjects were able to ignore even the largest delays of 500 ms).

Table 1. Comparative analysis of sound localization in the free-field and in virtual auditory displays

Experiment	Experimental procedure	Results
Auditory localization experiment of nearby broadband sources [11] **Free-field**	**Subjects:** 4 male subjects, aged 20-25; **Stimulus:** 5 different 150-ms pulses of white Gaussian noise, separated by 30-ms intervals of silence; Frequency: 200 Hz-15 kHz; Maximum amplitude: 59 dB SPL at 1m. **Procedure:** The control computer read 3 numbers, ranging from 1 to 6 (for azimuth, elevation and distance: 1 - 0 degrees azimuth, 6 - 180 degrees azimuth, 1 - +90 degrees elevation, 6 - -90 degrees elevation, 1 - 10-15 cm distance, 6 - 1 m distance). The listener kept his head fixed and the eyes closed during the test. **Training sessions** prior to data collection; the subjects were not given any feedback about their responses. **Test:** 2 h sessions consisting in 4-5 blocks of 100 trials (each trial took 2 s), separated by short breaks. Each subject had undergone 4 or 5 2h sessions.	Overall angular error: 17 degrees. The largest error occurred at locations above and behind the subject, especially at close distance (mean error 27 degrees), while the smallest errors occurred at distant locations, in front and to the side. Azimuth error increased as the sound source approached the head, but elevation performance is independent of distance; Localization performance depends on azimuth: it is higher in the lateral regions of the head than in the median plane; The elevation error is lowest for lateral sources and greatest for locations behind the listener. Front-back confusions occurred in 10% of the trials; The rate of front-back confusions increases slightly at close distances. Distance perception is most accurate for lateral sources and least accurate near the median plane. Also, distance perception is better in the proximal region of the head (<1 m to the head).

Experiment	Experimental procedure	Results
Auditory localization experiment with additional visual cues in the horizontal and vertical planes [1] **Free-field**	**Subjects:** 8 sighted subjects with and without a blindfold; **Stimuli:** Broadband noise (100-ms pink noise bursts at 60 dB SPL), delivered via 2 loudspeaker arrays: a horizontal array with 25 loudspeakers (from -90 to 90 degrees, 7.5 degrees interval) and a vertical array with 16 loudspeakers (from -45 to 67.5 degrees). **Procedure:** The following conditions were compared: blindfold vs. non-blindfold, the horizontal vs. the vertical plane and two pointing methods: hand vs. head. **Training:** The users performed practice trials until they got accommodated with the apparatus (10-15 trials). **Test:** Each subject was tested in two separate 1 h long sessions that were scheduled 1 week apart. Each subject performed the tasks under the all 8 possible conditions (blindfold vs. non-blindfold, horizontal vs. vertical localization and hand vs. head pointing method – using either a laser pointer mounted onto the subject's head or a hand-held laser pointer).	Proprioceptive cues are sufficient for accurate hand pointing, while head pointing depends on the visual feedback. Localization performance in the horizontal plane is better for the hand-pointing than for head-pointing modality, while the head-pointing method offered better results in the vertical plane. Blindfolding significantly increased the absolute localization error for both the horizontal and the vertical plane and for both pointing conditions. Average localization error in the horizontal plane using the head pointing method is 6.36 degrees.
Sound source localization experiment using filtered noises [12] **Free-field**	**Subjects:** 45 listeners; **Stimuli:** 200 ms noise bursts (filtered using Butterworth bandpass filters with cutoffs of 125 to 500 Hz - lowpass, 1500 to 6000 Hz - high-pass and 125 to 6000 Hz - broadband), 65 dB, delivered via 13 loudspeakers arranged in the front hemifield; **Procedure:** The listeners were required to keep their head fixed. After the presentation of a noise burst, they entered a number on the keypad (between 1 and 13), indicating the number of the loudspeaker which produced the sound. **Test:** Blocks of 33 trials (11 loudspeakers x 3 filtering conditions), 4 blocks for each listener, resulting in 132 trials per listener.	The subjects performed with high reliability and repeatability, with a root-mean-square error (rms) of 6.2 degrees and a standard deviation of 1.79 degrees. The rms performance ranges between 2.6 and 9.8 degrees. The rms is smaller in the broadband condition than in the low-pass and high-pass condition. Filtering the sound does not affect sound localization performance. Free-field localization for tonal stimuli is poor compared to broadband stimuli (such as clicks and noises) and sound source localization acuity is better for low-frequency tones as compared to high-frequency tones.
Localization performance experiment of real and virtual 3D sound for use in fighter aircraft [4] **Free-field** **Virtual auditory display**	**Subjects:** 26 listeners. **Stimuli:** White noise bursts of 250 ms and 2 s, 75dB SPL. **Procedure:** Both virtual and free-field testing conditions. 15 values for azimuth and 9 values for elevation. Pointing device: a toy gun. The listeners were not allowed to view the loudspeaker setup. **Training:** Both headphones and loudspeaker familiarization using long bursts of 2 s, 16 directions, 3 repetitions. **Test**: Both virtual and real sound sources stimuli sequences (16 directions, each sound direction was repeated 3 times) for short and long duration stimuli (250 ms and 2 s).	Uncertainty of 10-14 degrees for azimuth and 12-24 degrees for elevation. The standard deviation error for azimuth using real sources and 2 s stimuli is 4.9 degrees and for elevation is 6.2 degrees. Virtual sound sources of 2 s yielded a standard deviation of 7.3 degrees for azimuth and 12.1 degrees for elevation (for 250-ms stimuli, it increases to 21.4 degrees for azimuth and 19.2 for elevation). Rate of front-back confusions: 4.2% (2 s stimuli), 9.1% (250 ms stimuli) for real sources and 5.1% (2 s stimuli) and 21.3% (250 ms stimuli) in the case of virtual sources.
Sound localization experiment on the nature and distribution of errors by human listeners [13] **Free-field**	**Subjects:** 19 subjects. **Stimuli:** Broadband short noise stimuli at 70 dB SPL. **Procedure:** The loudspeaker system was positioned at a radius of 1 m from the center of the listener's head. The listener indicated the perceived location by pointing his nose towards the sound source. **Training:** The listener was required to face his head towards the perceived location of the noise burst (the open-loop estimate of location) in complete darkness. Afterwards, a small light-emitting diode on the speaker was activated and the subject was allowed to readjust his head (closed-loop visual condition). Consequently, the listener passed through a closed-loop audio component where the sound stimulus was played repetitively,	Localization errors for azimuth locations in the front: 1.3 - 4.4 degrees. Average spatial misjudgments: 3 degrees in azimuth and 4 degrees in elevation. Broadband stimuli offer the best localization accuracy. The rate of front-back confusions: 3.2%. There were not recorded any up-down confusions.

Experiment	Experimental procedure	Results
	allowing him to adjust his head orientation towards the source. Each training block consisted of 36 locations. 10 subjects were trained on 9 blocks and the other 10 were trained on 4 or 5 blocks. **Test:** 4-6 blocks of 76 target locations for each subject.	
Sound localization experiment with visual feedback training [14] **Virtual auditory display**	**Subjects:** 6 subjects. **Stimuli:** 100 ms Gaussian noise bursts presented over a 3D headphone-based virtual auditory display. **Procedure:** The spatial positions included a 360 degrees azimuth range and a +/- 40 degrees elevation interval. Two types of visual stimuli were used: a head orientation “cross—hair” that was presented through a HMD and a second visual stimulus that indicated the correct source location. **Pre-test phase:** the sound localization accuracy was tested, without giving feedback about the correct sound source location. 144 spatial positions were tested. **The training phase:** Similar to the pre-test phase, except that the listener was given visual feedback about the correct sound source location. After the listener indicated his perceived location, both visual and audio feedback was presented. There have been used 24 spatial positions, repeated for 3 times, resulting in 72 trials per block, presented in random order. **Post-test phase:** identical to the pre-test phase, but conducted 4 days after the experiment.	The amount of localization improvement after training varied among listeners, leading to similar results as those of Wightman and Kistler [15] (between 12 and 24 degrees in virtual auditory displays). There has been recorded a reduction in the front-back confusion rate as a result of training (from 40 degrees to less than 25 degrees of error). A lasting effect of the perceptual training was observed: 4 days between the training and the post-test session did not affect accuracy in the last.
A psychophysical validation of headphone simulation of the free-field listening [3] **Free-field** **Virtual auditory display**	**Subjects:** 8 subjects. **Stimuli:** 250-ms bursts of Gaussian noise with 300 ms of silence between the bursts (70 dB SPL). **Procedure:** The subject indicated the apparent spatial position of a sound source by calling out numerical estimates of the perceived azimuth and elevation. **Training:** 10 h of practice in the free-field condition. **Test:** Free-field & headphones condition 36 different positions, covering a 360-degree azimuth range and elevations between 36 degrees below the horizontal plane and 54 degrees above it. The subjects were blindfolded. Each subject completed 6 runs in the free-field condition, 10 runs in the headphones condition, again 6 runs in the free-field condition and another 6 runs in the headphone condition to assess the effects of learning.	Sound perception is best on the side (laterally), slightly poor in the front and poorest at high elevations and in the back. The headphones condition provided similar localization results as in the free-field condition. There is a clear increase (approximately double) in the frequency of front-back confusions in the headphones condition. The difference between the free-field and the headphones condition lies in the elevation component of the response (elevation localization is slightly poorer in the headphones condition).
A two-dimensional sound localization test by human listeners [2] **Free-field**	**Subjects:** 6 subjects, aged 24-34 years. **Stimuli:** Broadband noise. **Procedure:** The subject was required to indicate the perceived direction by turning his face towards the apparent location of the source. The localization experiments consisted of blocks of open-loop (brief stimuli, 150 ms long that finished before the listener gave the response) and closed-loop (continuous stimuli, where the end of the stimuli was determined by the listener’s response). **Training:** 10-20 training sessions per listener. In a training trial, the subjects were given instant feedback about the correct location of the source through a light emitting diode situated on the top of the active loudspeaker. **Test:** The sound sources ranged in azimuth from -170 to +170 degrees and in elevation from -45 to +55 degrees, with a 10 degrees increment. A block of trials, consisted of 249 sound directions (3 test sessions of 83 trials each). The subjects have undergone 5 blocks of trials.	Localization accuracy is better in the horizontal plane than in the vertical plane (2-3 degrees mean error). The best localization performance was recorded in the frontal midline (<5 degrees) Vertical localization was better for the sounds located laterally. Vertical localization in the front was better for the sources situated near the horizontal plane; the localization errors increased at higher and lower elevation angles along the median plane. Error localization was lower under the closed-loop condition, especially in the frontal horizontal plane. The front-back confusion rate was higher under the open-loop condition (6%).

Experiment	Experimental procedure	Results
Localization experiment with non-individualized head-related transfer functions [16] **Free-field Virtual auditory display**	**Subjects:** 16 subjects (2 male and 14 female). **Stimuli:** A train of 8, 250-ms bursts of Gaussian noise with 300 ms of silence between the bursts. **Procedure:** In both the free-field and headphones conditions, the listeners pointed the apparent sound source direction by calling out the angles of perceived azimuth and elevation. The subjects did not receive any feedback about the accuracy of their responses. **Training:** A 15-min training session with verbal instructions concerning the required tasks and a practice session, consisting of a single block of trials, in the free-field condition. **Test:** 18 alternating (between the free-field and the virtual free-field condition) blocks of trials (extended over a period of 3 days). Each block of trials consisted of 24 sound source positions, randomly distributed in the horizontal and vertical plane, for both the free-field and headphones conditions.	All the listeners recorded accurate localization judgments for both conditions in the horizontal plane. The mean rate of front-back confusions in the free-field condition is 6.5% (ranging from 2% to 10%). The mean rate of front-back confusions under the headphones condition is 32% (ranging from 20% to 43%). The binaural cues are effectively synthesized under headphones condition, whereas the spectral cues that are in charge of vertical localization and front-back disambiguation are severely damaged during the non-individualized HRTF filtering process.
Experiment on the effects of increasing system latency for virtual sound source localization [17] **Virtual auditory display**	**Subjects:** 5 subjects (3 male, 2 female), aged 16-24. **Stimuli:** Broadband Gaussian noise, 8 seconds duration with latency values of 33.8, 100.4, 250.4 and 500.3 ms. **Procedure:** The listeners indicated the perceived azimuth, elevation and distance using a graphical interface. Training: The training session included 2 practice block, using the lowest latency of 33.8 ms. **Test:** 12 sound locations; 10 test blocks made out of 24 localization trials.	The front-back confusion rate: 5.2% - 8.8%. The front-back confusion rate increased with increasing latency. Azimuth error angles: 26.2 – 36.3 degrees and tend to increase with increasing sound latency. The minimum perceived latency is of 250 ms. The minimum latency that is required to affect localization accuracy is of 500 ms.

IV. DISCUSSION

The experiments presented in this paper show that the sound localization accuracy in the free-field is comparable with the localization performance in virtual auditory environments, especially for the sources situated in the horizontal plane. The use of non-individualized HRTFs in virtual auditory displays conducts to an increase in the rate of reversal errors (front-back and up-down confusions) and to a higher localization uncertainty in the vertical plane. The listeners were able to acquire significant directional information from the binaural cues, which were accurately synthesized in virtual auditory displays, even for non-individualized HRTFs. Nevertheless, the high incidence of front-back confusions still remains a problematic issue. The high localization error rate is determined by the static nature of the stimuli, the lack of visual cues (it has been demonstrated that vision plays a fundamental role in auditory discrimination), the use of distorting sound spectrum characteristics induced by the 3D audio synthesis process, the lack of sound externalization (in-the-head localization) and the uncertainty caused by the cone of confusion. However, there are significant studies which demonstrated that sound localization accuracy under headphones condition can be substantially improved as a result of training, perceptual learning and auditory adaptation [18] [19] [20].

In order to minimize the rate of reversal errors, a virtual auditory display should incorporate dynamic cues (a head motion tracking method), visual cues and other sonification techniques inspired from the acoustic model of enclosed spaces (reverberations, distance cues, externalization or the ratio of direct to reflected sound energy). Additionally, it should employ broadband stimuli [13] (which have a larger range of frequency variation), with minimum system latency between the moment of the head movement to the moment of the perceived sound change in the headphones [17]. In what concerns the type of experimental localization procedure, the closed-loop condition [2] conducted to the best localization performance, as it enabled the subjects to continuously listen to the sound and to move their heads freely, in order to accurately calibrate their direction towards the sound source and to make the best localization judgment.

V. CONCLUSIONS

This paper presented a comparative study of the sound localization accuracy in the free-field and in virtual auditory displays. As discussed previously, both the free-field and the virtual auditory displays offer a comparable localization performance for azimuth judgments. Nonetheless, the rate of front-back confusions is significantly higher for the headphones condition, especially if the 3D sound is synthesized with generic HRTFs. The results of this study highlight the importance of virtual auditory displays, as a modality of simulating the perception of sound in the virtual environment. Consequently, the design of a virtual auditory display should meet some conceptual and methodological requirements in regard with the means of reducing the rate of reversal errors, the vertical localization misjudgments and the lack of sound externalization, in order to accurately deliver to the listener a spatial auditory perception as natural as possible.

Acknowledgements

The work has been funded by the Sectoral Operational Programme Human Resources Development 2007-2013 of the Ministry of European Funds through the Financial Agreement POSDRU/159/1.5/S/132395 and POSDRU/159/1.5/S/134398. This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 643636 "Sound of Vision.

Reference Text and Citations

[1] Tabry, V., Zatorre, R.J., Voss, P., 2013. The influence of vision on sound localization abilities in both the horizontal and vertical planes, Front Psychol. 2013 Dec 12;4:932.

[2] Makous, J.C., Middlebrooks, J.C., 1990. Two-dimensional sound localization by human listeners, J Acoust Soc Am. 1990 May; 87(5):2188-200.

[3] Letowski, T., Letowski, S., 2011. Localization Error: Accuracy and Precision of Auditory Localization (Chapter 4), "Advances in Sound Localization", book edited by Pawel Strumillo, ISBN 978-953-307-224-1, published: April 11, 2011

[4] Padersen, J.A., Jorgensen, T., 2005. Localization Performance of Real and Virtual Sound Sources, RTO MEETING PROCEEDINGS NORTH ATLANTIC TREATY ORGANIZATION RTO MP HFM -CD-ROM EDITION-, 123; P29 , 2005

[5] Wersényi, GY., 2002. HRTFs in Human Localization: Measurement, Spectral Evaluation and Practical Use in Virtaul Audio Environment. Ph.D. doctoral thesis, BTU Cottbus, 2002.

[6] Wersényi, GY., 2005. What Virtual Audio Synthesis Could Do for Visually Disabled Humans in the New Era? Proceedings of the 12th AES Regional Convention Tokyo, June 12-14, 2005. /180-183/.

[7] Bălan, O., Moldoveanu, A., Moldoveanu, F., Dascalu, M.I., 2014. Navigational 3D Audio-Based Game- Training Towards Rich Auditory Spatial Representation of the Environment, Proceedings of the 18th International Conference on System Theory, Control and Computing, Sinaia, Romania, October 17-19, 2014, pp. 688-693.

[8] Wersényi, GY., 2007. Virtual simulation and the binaural technology. Híradástechnika, Vol. LXII, Nr.2, 2007 February. /25-32/

[9] Wersényi, GY., 2008. Computer aided software application for soundfiled optimalization. Journal of Telecommunications, Vol. LXIII, Nr.5, 2008. /35-44/

[10] Wenzel, E.M., Wightman, F.L., Kistler, D.J., 1991. Localization with non-individualized virtual acoustic display cues. In Proceedings of CHI. 1991, 351-359.

[11] Brungart, D.S., Durlach, N.I., Rabinowitz, W.M., 1999. Auditory localization of nearby sources. II. Localization of a broadband source, J Acoust Soc Am. 1999 Oct; 106(4 Pt 1):1956-68

[12] Yost, W.A., Loiselle, L., Dorman, M., Burns, J., Brown, C.A., 2013. Sound source localization of filtered noises by listeners with normal hearing: a statistical analysis, J Acoust Soc Am. 2013 May;133(5):2876-82.

[13] Carlile, S., Leong, P., Hyams, S., 1997. The nature and distribution of errors in sound localization by human listeners, Hear Res 1997 Dec;114(1-2):179-96

[14] Zahorik, P., Tam, C., Wang, K., Bangayan, P., Sundareswaran, V., 2001. Localization accuracy in 3D sound displays: The role of visual-feedback training. In P. N. Rose (Ed.), Proceedings of the Advanced Displays Consortium: ARL's 5th Federated Laboratory Annual Symposium, College Park, MD. pp. 17-22

[15] Wightman, F.L., Kistler, D.J., 1989. Headphone simulation of free-field listening. II: Psychophysical validation, J Acoust Soc Am. 1989 Feb;85(2):868-78.

[16] Wenzel, E.M., Arruda, M., Kistler, D.J., Wightman, F.L., 1993. Localization using nonindividualized head-related transfer functions, J Acoust Soc Am. 1993 Jul; 94(1):111-23.

[17] Wenzel, E.M., 2001. Effect of increasing system latency on localization of virtual sounds with short and long duration, in Proceedings of the 2001 International Conference on Auditory Display, Espoo, Finland, July 29-August 1, 2001,pp. 185–190.
[18] Shinn-Cunningham, B. G., Streeter, T., Gyss, J.-F., 2005. Perceptual plasticity in spatial auditory displays, TAP 2 (4), 418-425.
[19] Honda, A., Shibata, H., Hidaka, S., Gyoba, J., Iwaya, Y., Suzuki, Y., 2013. Effects of head movement and proprioceptive feedback in training of sound localization, Iperception. 2013 Jun 3;4(4):253-64.
[20] Parseihian, G., Katz, B.F.G., 2012. Rapid Head-Related Transfer Function adaptation using a virtual auditory environment, J. Acoust. Soc. Am., Volume 131, Issue 4, 2012.