Evaluation of Bone Conduction and Active-Noise-Cancellation Headsets Based on Listening Tests in a Virtual Environment

György Wersényi

Abstract—**Alternative design headsets incorporating active-noise-cancellation or bone conduction were evaluated in listening tests in a virtual reality environment. Virtual sound sources in the horizontal plane had to be identified using stereo panning in the frontal hemisphere. In addition, transfer characteristics and damping effects were measured with a dummy-head. Results indicate that up to five source locations can be used in real applications with high accuracy in virtual scenarios, independent of the spectral content of the excitation signals. Furthermore, the use of noise cancellation in presence of 80 dB background noise does not improve performance. Commercially available bone conduction headsets can provide the same detection accuracy even if the subjective sound quality is lower.**

Index Terms—**Active Noise Cancellation, headphone, measurement, bone conduction, virtual reality**

I. INTRODUCTION

Localization is the procedure during which human subjects try to find sound sources [1], [2]. Directional hearing is based on interaural time and intensity differences between the signals on the eardrums, as well as on spectral filtering cues of the outer ears [3]. Furthermore, many other parameters influence the localization accuracy, i.e., spectral and temporal properties of the signal, experimental environment (real-life or virtual), playback devices, additional sound signal processing, presence of distractor and masking sounds, and experience of the subjects, etc. It is also an important parameter whether source locations must be ”pointed to” in a so-called absolute localization task, or subjects must identify source locations from a limited number of possibilities (e.g., a given number of directions or loudspeakers). Both methods can be applied in real and virtual environments, and the latter is usually a simplified task due to the limited possibilities. The term lateralization can be also used along with localization, if the effect of externalization plays a significant role [4].

The significance of virtual reality (VR) environments has increased with the introduction of modern playback and feedback devices for both desktop and for full immersive VR with headsets. Furthermore, virtual audio displays (VAD) play a significant role in various applications from assistive technologies to simulators and gaming, extending or replacing visual screens [5], [6]. It is often a requirement to have privacy and to restrict the virtual audio/visual experience to the user only. On the other hand, safety issues - especially in mobile applications - may require some contact with the environment regarding environmental sounds, such as traffic noise, alarm sounds, and speech communication. Traditional headphones covering the ears or plugged in the ear canal damp the outer world, although the hearing modality may be the only sense receiving information from the surrounding environment, if vision is focused elsewhere. VR environments and simulation tools exist not only to create 3D visual spaces, but they are also appropriate for audio scene rendering, usually over two-channel headsets. Various methods can assist the simulation of directional information of virtual sources. Virtual labs for scientific purposes offer solutions for individual needs for experiments [7]–[9].

The introduction of nontraditional headsets with different solutions to enhance sound quality, can also provide convenience services or extended safety also for VADs and VR scenarios. Bone conduction is a technique that uses the human skull to transmit acoustic vibrations while leaving the ear canals open [10], [11]. Another interesting and widespread method for high quality headphones is the presence of an active-noise-cancellation circuit to increase performance [12], [13]. Localization, however, may be influenced and is different in contrast to traditional headphones. The next subsections discuss these two techniques briefly.

A. Bone Conduction

Bone conduction (BC) headsets or ”bonephones” may be an alternative to traditional headphones and earphones if an open ear canal entrance is required during operation. The most important application areas include electronic travel aids (ETAs) and wearables for assistive technology (mostly for the visually impaired [14], [15]); outdoor sport activities (safety during running or biking in traffic); military or combat activities with sparsely occurring commands [16]; as well as any application where environmental sounds and noises can not be blocked [17]. Although VR applications usually try to block the outer world for full immersion, many applications can benefit from an open ear canal. In case of BC, sound quality is inferior to traditional headphones due to the indirect sound path through the skull instead of the direct sound path through the eardrum. A restricted frequency range is common, while directional information can be disturbed. Furthermore, transmission depends strongly on the transducer position and whether the ear canal is open or closed [18]–[20].

Széchenyi István University, Győr, Hungary (e-mail: wersenyi@sze.hu).

DOI: 10.36244/ICJ.2023.3.6

Higher sensitivity (lower hearing threshold), however, can occur during skin penetration [21]. A threshold in case of a plugged ear canal was found to be 10-20 dB lower than the unplugged threshold at low frequencies (200 Hz and below), beginning to converge with the unplugged threshold at medium frequencies (250 – 1000 Hz), and the two being equal above 2 kHz [22].

The mastoid would be a preferable transducer location relative to the forehead or temple because it contains the inner ear. It is relatively immune to the interference associated with muscle tissue operating the jaw, and it allows stereo presentation of sounds [23]. In a study, binaural hearing ability of normal hearing adults with BC and air conduction (earphones) was contrasted. BC was applied in the audiometric position on the mastoid [24]–[27]. Results confirmed that binaural hearing processing with bilateral BC stimulation is present. However, the binaural benefit was overall greater with air conduction stimulation. On the other hand, other studies reported no significant difference [16], [28], [29]. Nevertheless, current devices do not use the mastoid, but the jaw bone and the zygomatic (cheek) bones in front of the ear.

BC use is traditionally limited to monaural applications due to the high propagation speed of sound in the human skull. Spatial audio does not occur naturally through bone conduction, although interaural level and intensity differences can be simulated. It was shown that stereo bone conduction headsets can be used to provide a limited amount of interaural isolation in a dichotic speech perception task [30]. The results suggested that reliable spatial separation is possible with bone conduction headsets, but they probably cannot be used to lateralize signals to extreme left or right apparent locations. However, the degree of lateralization can be similar to that of produced by using headphones. Results from an empirical user study conducted to compare one BC device, headphones, and a speaker array showed that subjects performed the best by using physical speakers with stationary sounds. Nevertheless, there was no difference in accuracy between the speakers and the BC device (outperforming even standard headphones) for moving sounds [31]. Elevation cues can also be adjusted for 3D simulations using BC [32]. Monaural spectral cues are also responsible for front-back and up-down discrimination and externalization of sound sources, and problems occurring frequently during headphone playback. Elevation cues can be simulated for BC with different methods using Head-Related Transfer Functions (HRTF) or simple high-pass and low-pass filtering [15], [33], [34].

B. Active Noise Cancellation

Active-Noise-Cancellation (ANC) is a technique, where an incident, unwanted sound signal is "cancelled out" by inverting and adding it to the computer-generated signal based on interference. In case of a perfect destructive interference, the sum of the incoming and inverted signals equals zero. Although the theory is very simple, practical realization of ANC circuits in noise cancelling headphones face many problems [35]–[37]. Some of the ANC headsets offer not only cancelling, but also allow the user to hear environmental noises by reducing isolation or enabling communication with the environment by pushing a button (and avoiding taking off the headphone), i.e., via pass-thru, mic-thru or hear-thru functions. This is beneficial for augmented reality applications as well, where environmental sounds and computer-generated signals are mixed together.

ANC, in general, is supposed to increase damping, allowing for reduced loudness during playback and leveling the subjective quality of the playback. This reduction of disturbing noises may also result in better localization performance in listening tests.

This paper presents the results of a classic subjective listening test, using BC and ANC headsets in a virtual environment where virtual sound sources had to be identified using various excitation signals. The goal was to test whether these cost-effective, commercially available devices can be an alternative to traditional headphones in applications where directional information is important. Section 2 presents the measurement setup, including the virtual environment, the headsets and the experimental procedure. The subsequent section presents the results of the listening tests and the objective measurement of the technical parameters (transfer function and damping). Then, the results will be discussed and conclusions will be drawn together by highlighting future research directions.

II. MEASUREMENT SETUP

A. The virtual environment

The virtual environment was created by applying the MaxWhere platform. This multipurpose 3D collaborative environment is a versatile platform made up of various virtual spaces that can be used effectively in education, virtual laboratory tests, and even for testing memory capabilities [38], [39].

Upon initialization, the user could set personal data (name/ID, age, gender) and the simulation environment. The main settings included the number of virtual sound source directions (3, 5 or 7), and the type of the sound sample (impulse, 1 kHz sinus, white noise, and female speech). Furthermore, a checkbox was marked if the user was visually impaired, if the ear canal entrance was closed, and in case of a looped signal presentation. Currently, only sighted subjects participated, thus the first checkbox was always set to false. The ear canal was closed in case of regular headphones and it was left open in case of BC devices. By default, all sound samples were played back once without being looped.

Figure 1 shows the directions of the virtual sources. Figure 2 depicts the screenshot of the arrangements where subjects had to click. In case of three directions, the front 0°, the left $+90^\circ$ and the right -90° were set. Because left and right correspond to a radiating sound only from the left or the right speaker respectively, the actual vizualisation would require a more complicated way where the user had to turn. To avoid this (that would make the usability worse), the extreme left and right speakers were displayed relative to the listeners spot (marked red). The users did not have any problem to interpret the user interface.

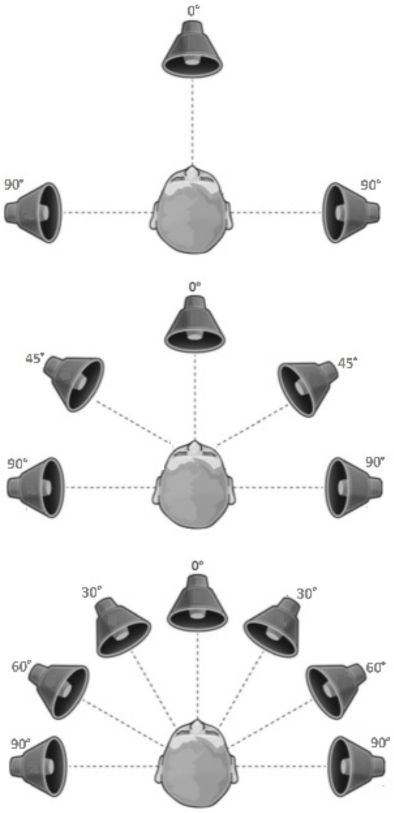


Fig. 1. Arrangement of the virtual speakers in three, five, and seven directions.

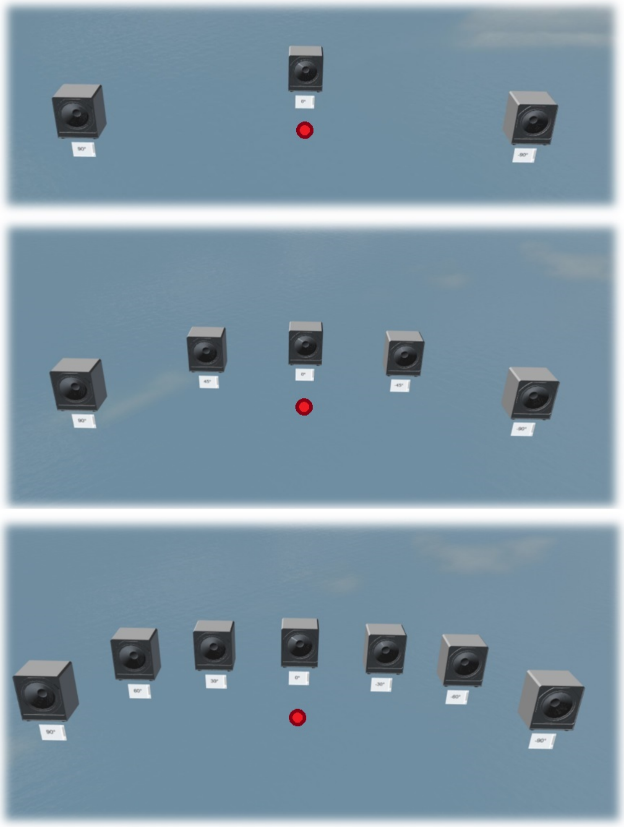
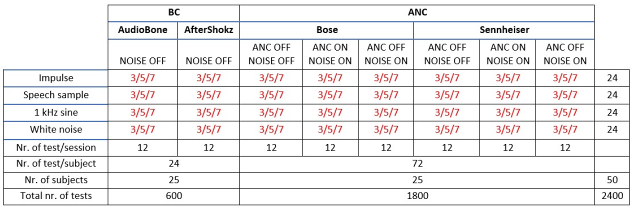


Fig. 2. Screenshots of the virtual environments. After playback of the sound, subjects identify the sound source by clicking on the loudspeaker icon.



	BC		ANC						
	AudioBone	AfterShokz	Bose			Sennheiser			
	NOISE OFF	NOISE OFF	ANC OFF NOISE OFF	ANC ON NOISE ON	ANC OFF NOISE ON	ANC OFF NOISE OFF	ANC ON NOISE ON	ANC OFF NOISE ON	
Impulse	3/5/7	3/5/7	3/5/7	3/5/7	3/5/7	3/5/7	3/5/7	3/5/7	24
Speech sample	3/5/7	3/5/7	3/5/7	3/5/7	3/5/7	3/5/7	3/5/7	3/5/7	24
1 kHz sine	3/5/7	3/5/7	3/5/7	3/5/7	3/5/7	3/5/7	3/5/7	3/5/7	24
White noise	3/5/7	3/5/7	3/5/7	3/5/7	3/5/7	3/5/7	3/5/7	3/5/7	24
Nr. of test/session	12	12	12	12	12	12	12	12	
Nr. of test/subject	24		72						
Nr. of subjects	25		25						50
Total nr. of tests	600		1800						2400

Fig. 3. Overview of the experiments. Red numbers indicate the number of sound directions.

In the case of five locations, two additional locations of 45° "halfway" left and right from the frontal direction were included. In the case of seven directions, instead of 45°, 30° and 60° were introduced. Directions were simulated by applying a simple two-channel stereo intensity panning between the left and the right channels. Although two-channel virtual reproduction of sound scapes often suffer from in-the-head localization (lack of externalization), complex full binauralization is not necessarily required for applications where usability does not rely on externalization.

The excitation signals are the following:

- computer-generated white noise sample of 1.5 sec with a flat spectrum between 20 Hz and 20 kHz;
- a female speech sample with a base frequency of 430 Hz and harmonic components up to 8.2 kHz (2 sec);
- 1 kHz sine, with damping of the higher harmonics more than 60 dB (1.5 sec);
- an impulse-like complex sound generated from a 258 Hz base frequency signal and high order odd harmonics up to 15 kHz (2 sec).

During a test session, all sound sources were activated three times in a randomized order (unknown to the user). Thus, in case of three directions, the users had to identify a source location 9 times, in case of five sources it was 15, and in case of seven locations the experiment lasted for 21 clicks. After completing the task, a log file in the JSON format was saved with the personal data, time and date (timestamp), headphone type, completion time in seconds and the total points obtained. The maximum points were 9, 15 and 21 respectively, and the result was also calculated in percentage. Furthermore, the number of correct and incorrect identifications for each location was recorded.

B. Experimental setup

Laboratory sessions were conducted in a silent, acoustically controlled room. The room had originally been designed for sound recordings with insulated walls and a reverberation time below 1,5 s. A loudspeaker was also installed in front of the listener behind a desk where the computer was placed.

The subject was sitting in a chair and he was controlling the experiment by mouse clicks on a standard laptop screen. In the case of a distractor sound, the loudspeaker radiated a previously pre-recorded traffic noise, having a SPL level of 80 dB at the listening point. An external distractor could be used to test masking effects and/or the effect of active noise cancellation.

Prior to the first session, the measurement procedure and the goal of the experiment were described to the subjects, which was followed by a brief introduction to the sound signals and to the user interface (how to click, how to get feedback etc.). During the experiment, subjects wore one of the headphones. Headsets were placed on the head and adjusted to fit the most comfortably. The users set the loudness to the most comfortable level, neither too loud, nor too quiet. Due to the relatively low sensitivity of the BC types, the maximum volume was set on the laptop to provide appropriate loudness levels. The vibrational actuators were positioned on the jawbones in front of the ear, instead of on the mastoid bones, similarly to [31]. Finally, personal data were entered for the JSON files.

One session was to test one of the headsets on all three scenarios with four signal types. This resulted in 12 test runs in total. The test always started with the impulse (3, 5 and 7 directions), followed by speech, sinus, and white noise. The total time needed for one session was about 30 minutes. The procedure was repeated with the other headset. In the case of BC devices, 25 subjects evaluated them in two sessions. In the case of ANC devices, a different set of 25 participants evaluated them in six sessions. All 50 participants had normal hearing self reportedly (no audiometric screening was applied). In the absence of distractor noise, the measurement was conducted with ANC off only. Figure 3 shows an overview of the sessions.

C. Devices

For the experiment, two types of BC and two types of ANC headsets were used as listed (Figure 4).

- Aftershokz Sportz M3 (wired) [34], [40];
- AudioBone Deluxe GDP 02 [31], [41];
- Sennheiser PXC 450 [42];
- Bose QC25 [43].

The AudioBone was already used in other experiments and the AfterShokz models are one of the most popular devices nowadays (usually for outdoor sport activities). Both are cost-efficient and they can be placed on the jawbone instead of behind the ears. ANC models are more widespread and commercially available from different vendors in various forms (supraaural, in-ear and buds).

Transfer characteristics of the ANC devices were measured following the standard protocol for headphone measurements [44], [45]. In a semi-anechoic room, the Brüel and Kjaer Type 5128 head and torso simulator was connected to the PULSE data acquisition system. The same headphone was placed and replaced 10 times for averaging, and two-channel measurements using white noise excitation, which were performed on ments using white noise excitation, which were performed on the left and the right ears, respectively. Test were performed both with and without ANC [46].

Fig. 4. Four headsets used in the experiments. AfterShokz (top left), AudioBone (top right), Bose QC25 (bottom left), and Sennheiser PXC 450 (bottom right).

III. RESULTS

A. Technical parameters

In the case of BC devices, the conventional transfer function measurements designed for headphones could not be used. The output signal is not the sound pressure on the eardrums, but the vibrations on the skull and/or inside the head (in the middle and inner ear). The sensation is very complex based on various transmission paths of vibrations with an airborne sound through the ear canal being also present. There is no standard method to determine the transfer characteristics. Although there is a measurement equipment called artificial mastoid for the calibration of audiometric bone vibrators, it was designed in a very simplified form to model the human receiver and for an excitation point on the mastoid, and not on the jawbone [47]–[49]. Measurements were conducted the same way as described above, also with the BC devices to check the effects of airborne sound transmission. Please note that this is only part of the operation (often regarded as malfunction rather than a feature), as the main transmission should be via the skull.

Another technical parameter that can easily be determined by utilizing the setup and the equipment described above, is damping in dB over the entire frequency range in case of any headphone. This can simply be done by measuring the transmission without the headphone, followed by the same procedure with the headphones on. The reference signal in this case was the transfer characteristics of the dummy-head from the free-field to the eardrum, also called a Head-Related Transfer Function (HRTF) [50], [51]. For the measurement, a loudspeaker was set up in front of the dummy-head (frontal direction). Although HRTFs vary with direction, we only used the frontal HRTFs for damping measurements. The quotient of HRTFs with and without headphones resulted in the

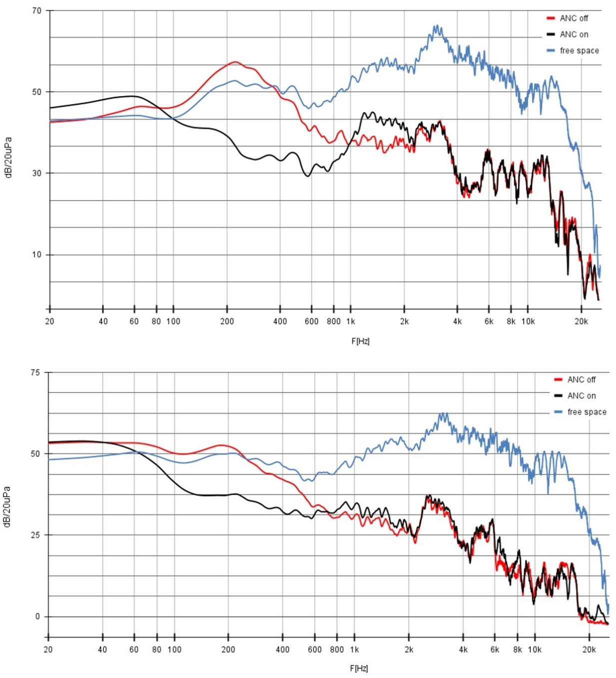


Fig. 5. Damping of the ANC devices: Bose (top) and Sennheiser (bottom).

damping characteristics of the headphone (eliminating other environmental influences, i.e., possible reflections, transfer characteristics of the loudspeaker, etc.). In the case of BC devices, damping is not measured as they do not cover the outer ears. Figure 5 shows the results for the ANC types. The blue lines are the free-field HRTFs from the frontal direction, the black and red lines represent the transmission if the headphones are on the ears with and without ANC. The difference in dB gives the damping in frequency. The next sections discuss the results of the listening test.

TABLE I
MEAN SCORES OUT OF 15 (5 DIRECTIONS)

	AVG	ANOVA
IMPULSE	10.24	F=1.87; Fcrit=2.70; p=0.14
SPEECH	11.92	
1KHZ SIN	11.36	
WNOISE	11.48	

B. BC

25 participants (mean age 38.2±15.1 years, 12 males and 13 females) took part in the tests. The statistical evaluation of the comparison of the two devices showed no significant difference between the means. Thus, the results were combined for the purpose of simplification.

Identification of three virtual directions showed 100% accuracy (9/9 points), only one subject clicked false once accidentally. This result for three directions is valid for both BC headsets. Further evaluation was carried out by applying five and seven directions based on ANOVA. In the case of five directions, there was no significant difference between the means of the four signals (F=1.87; Fcrit=2.70; p=0.14). The same is true for 7 directions (F=0.87; Fcrit=2.70; p=0.46). Tables I and II show the mean scores for all signal types.

TABLE II
MEAN SCORES OUT OF 21 (7 DIRECTIONS)

	AVG	ANOVA
IMPULSE	11.12	F=0.87; Fcrit=2.70; p=0.46
SPEECH	12.64	
1KHZ SIN	12.20	
WNOISE	12.12	

C. ANC

There were 25 participants (mean age 42.1±18.2 years, 13 males and 12 females) in the experiment. The statistical evaluation of the comparison of the two devices showed no significant difference between the means. Therefore, the results were combined for the purpose of simplification.

The evaluation of three virtual directions was completely identical to the results obtained with the use of BC. All subjects could detect all three directions with 100% accuracy (9/9 points), only one unintentional false click was recorded. This result for three directions is valid for both headsets with and without ANC activation and distractor noise.

Further evaluation was made using five and seven directions based on ANOVA. In the case of five and seven directions, headsets were used in three different scenarios such as:

- ANC ON - NOISE ON;
- ANC OFF - NOISE OFF;
- ANC OFF - NOISE ON.

The NOISE ON case refers to an active distractor sound.

TABLE III
MEAN SCORES OUT OF 15 (5 DIRECTIONS)

IMPULSE	ANC ON	ANC OFF	p-value
NOISE ON	10.6	10.4	0.864
NOISE OFF		10.5	
SPEECH	ANC ON	ANC OFF	p-value
NOISE ON	12.7	11.6	0.913
NOISE OFF		11.5	
1KHZ SIN	ANC ON	ANC OFF	p-value
NOISE ON	12.1	12.1	0.769
NOISE OFF		10.8	
WNOISE	ANC ON	ANC OFF	p-value
NOISE ON	13.1	12.7	0.448
NOISE OFF		11.1	

TABLE IV
MEAN SCORES OUT OF 21 (7 DIRECTIONS)

IMPULSE	ANC ON	ANC OFF	p-value
NOISE ON	9.7	9.1	0.740
NOISE OFF		9.4	
SPEECH	ANC ON	ANC OFF	p-value
NOISE ON	12.4	12.4	0.998
NOISE OFF		11.8	
1KHZ SIN	ANC ON	ANC OFF	p-value
NOISE ON	13.7	12.3	0.211
NOISE OFF		10.5	
WNOISE	ANC ON	ANC OFF	p-value
NOISE ON	14.0	12.3	0.158
NOISE OFF		12.5	

Tables III and IV present the results for five and seven directions, respectively. The mean scores are shown for the three scenarios and for all excitation signals. The p-values indicate whether there is a significant difference between ANC ON - NOISE ON and ANC OFF - NOISE ON situation.

IV. DISCUSSION

A. Objective evaluation

Transfer functions of the ANC devices showed typical frequency plots of high-quality headphones with the HRTF attenuation around 3 and 7 kHz, as expected. Furthermore, the effect of the ANC circuit is clearly visible causing an almost frequency independent wide band attenuation of about 10-12 dB in the case of Bose, and a negligible attenuation in the case of Sennheiser. The technical parameters of the tested devices are very similar; selection for an application can be based on subjective preference, pricing, or ergonomic considerations.

Former analyses of the transmission characteristics showed uneven frequency response of BC headphones, compared to conventional headphones or speakers [52]. Placement of the transducer on the skull is critical, especially, if a real human head is used that moves during operation. The variability in the measurements supports the subjective findings about the sound quality to be influenced by the placement. There is a significant airborne sound transmission from about 200-400 Hz up to 10 kHz in case of an open ear canal.

Increased damping can be observed between 100 Hz and 1000 Hz, using the ANC mode by both ANC devices, up to 20 dB. Above 1 kHz, there is no significant difference with or without ANC; however, the headphone damping becomes significant. We can conclude that the missing natural damping of the headphone below 1 kHz can be extended to about 100 Hz by activating ANC.

For BC devices, damping of airborne sound could be assured if the ear canal entrances were plugged. However, the main goal of this equipment is to leave the ear canal open as the primary operation. In a former measurement, the lowest threshold (i.e., maximum sensitivity) in both open and plugged conditions occurred around 1170 Hz-1370 Hz and within the frequency range of speech. At frequencies below or above these frequency values, greater intensity was required to detect the sound output by the bonephones [23]. Plugging the ears would lower the threshold up to about 2000 Hz.

B. Subjective evaluation

The main goals of the experiment were to compare four individual signal types in three different virtual speaker setups, testing the effect of ANC if a distractor sound is present, and to decide whether ANC and BC headsets have benefits in contrast to regular headphones.

The AudioBone headset was also used in a listening test where standard speaker setups were contrasted with BC audio. A speaker constellation of 90°, 45° and 0° was installed, and stationary and moving sources were emulated. Results showed the best accuracy with a physical speaker array and stationary sound, but there was no difference between the speaker array and the bone conduction device for sounds that were moving [31]. Another study used the Aftershokz Sportz3 [34]. Here the goal was to introduce vertical localization cues to BC playback. It was reported that this could result in decreased localization accuracy in the horizontal plane. Furthermore, there was also a significant ”compression” in the area directly in front of the observer in elevation. Our measurements indicate a similar effect horizontally in the case of seven directions, as subjects reported 30°-60° separation to be more difficult than separating 60° and 90°.

Employing the BC devices, participants reported upon their subjective feelings related to the signals. Impulse and noise were the most annoying and the most difficult ones to localize, and speech was the best/easiest. Indeed, results with impulse were behind the results of the others. On the other hand, the subjective feelings on white noise were not supported by the results. All but one participant mentioned this signal to be the most difficult one to localize; nevertheless, their performance was better than with other signals. As expected, it was a common observation that having seven instead of five directions increases difficulty. Seven directions were reported to be inappropriate and ”too many”, while three directions were ”too easy”. We can hypothesize that different results would be measured if the three directions were not 90° apart, but closer to one another. Subjects seemed to be connecting the unpleasant sounds to the difficulty of localization (i.e., noise and impulse are disturbing, therefore, more challenging to localize in contrast to speech). They were, however, not related. It was also often mentioned that it is harder to distinguish between 30° and 60° than between 60° and 90° in the 7-direction scenario. Furthermore, results may be biased by the actual order of the randomized signal presentation; directions can be compared more reliable if they are neighbors. For example, directions 30° and 60° from the same side can be easier separated if they are presented after one another.

In case of BC, the placement and adjustment of the transducer on the jawbone affect the observed sound quality. Displacement during the listening test due to small head movements, swallowing, moving the jaws can result in a drop of sound quality, and losing the symmetry between the left- and the right-hand sides. It is common that subjects frequently replace and adjust the position of the transducers. Furthermore, the AfterShokz has a belt that cannot be adjusted to the head size and the fixed transducer position is not optimal for all subjects.

The effects of learning and experience were not measured directly. It was, however, self-reported by the subjects that they had gotten better at the task after completing several sessions. This also had a motivational effect; some users tried to become better and ”break the record”. We could find evidence from the literature about the effect of training in listening tests. Although not tested directly, we can support the findings that the localization performance may be increased by repeated sessions [53]–[56].

Similarly, the experiment was not designed to test the dependence between the results and age. Elderly subjects (45+) were underrepresented in the sample. Although a statistical analysis showed no difference between the groups above and below 45, a correct sample of subjects with an equal number of participants in age groups may result in a different outcome.

There is no need to compare anything in the case of three directions, as subjects delivered a perfect accuracy in all cases. The mean scores and the corresponding ANOVA results for five and seven directions can be seen in Tables I and II. There is no significant difference among the four excitation signals.

Based on Table V, the average scores vary from 68.3% to 79.5% for five directions and from 52.9% to 60.2% for seven directions, respectively. As expected, there is a significant difference (F=39.35; Fcrit=5.98; p=0.0007).

TABLE V
COMPARISON OF AVERAGE SCORES FOR EACH SIGNAL (BC TYPES ONLY FOR 5 AND 7 DIRECTIONS)

	5 directions AVG SCORE OUT OF 15	%	7 directions AVG SCORE OUT Of 21	%
IMPULSE	10.24	68.3	11.12	52.9
SPEECH	11.92	79.5	12.64	60.2
1 KHZ SIN	11.36	75.7	12.20	58.1
WNOISE	11.48	76.5	12.12	57.7

For the ANC headsets, the mean scores and the corresponding ANOVA results for five and seven directions are depicted in Tables III and IV. There is no significant difference among the four excitation signals with and without ANC and external noise.

Based on Table VI the average scores vary from 65.6% to 76% for five directions and from 45.6% to 59.2% for seven directions, respectively. As expected, there is a significant difference (F=24.42; Fcrit=5.98; p=0.0026).

TABLE VI
COMPARISON OF AVERAGE SCORES FOR EACH SIGNAL (ANC TYPES ONLY FOR 5 AND 7 DIRECTIONS)

	5 directions AVG SCORE OUT OF 15	%	7 directions AVG SCORE OUT Of 21	%
IMPULSE	9.84	65.6	9.58	45.6
SPEECH	10.92	72.8	11.87	56.5
1 KHZ SIN	11.28	75.2	11.28	53.7
WNOISE	11.40	76.0	12.43	59.2

One may conclude that the impulse signal is the most difficult one to use, even though the difference is statistically not significant at the 0.05 level. Indeed, it may be related to the fact that the first signal was always the impulse; therefore, the subjects' performance improved based on user experience.

As expected, increasing the number of source locations from five to seven drops drastically the number of correct scores. An exhaustive paired statistical analysis across all measurements for all four signals between five and seven directions in the ANC ON-NOISE OFF (p=0.009, p=0.033, p=0.01, p=0.035); ANC ON-NOISE ON (p=0.035, p=0.003, p=0.045, p=0.011); and ANC OFF-NOISE ON (p=0.002, p=0.004, p=0.003, p=0.001) scenarios revealed statistically significant differences in each case. For the drastic drop of correct judgements (from 14 to 26%) is the introduction of two more directions responsible only, while the type of the excitation signal, ANC activation and the presence or absence of the external distractor sound are insignificant.

Comparing BC and ANC headsets is an interesting question. The combined results of the two BC and ANC devices respectively, showed no significant difference neither in case of five directions (F=1.44; Fcrit=3.49; p=0.28) nor in case of seven directions (F=0.67; Fcrit=3.49; p=0.59). Looking at the means over all signal types, this result could be expected. Average scores vary from 10.4 to 13.1 out of 15; and from 9.1 to 14 out of 21. We concluded that in this experiment both BC devices and both ANC headsets perform equally. Furthermore, if we consider ANC OFF-NOISE OFF situation as a reference traditional headphone case (all ANC devices act as a regular headphone if the battery is low or ANC is off), paired t-tests support that standard headphone modes do not outperform BC devices and will not be outperformed by an activated ANC mode.

In this experiment, no external distractor sounds were used in the case of BC headsets. The ANC effect was tested with simulated traffic noise. The additional damping of the ANC circuit did not affect the localization performance at all. We may expect different outcomes in case of higher SPL of the distractor sound. The applied 80 dB SPL was reduced by the damping effect to about 20 dB; thus, masking effect of the noise was not detected. The masking effect of the distractor sounds is a widely investigated area [57]–[61]. In a reverse condition where the distractor sound came from the BC set, listeners had to localize real sound sources (loudspeakers) located around them, with and without distractor sounds. Participants had greater localization error in the distractor-present conditions, especially in case of multiple distractors, regardless of whether or not participants ignored distractors [62].

Although not included in the final experiment, other models were also used in preliminary tests during the design-phase of the experiment. The Sony WH-1000XM3 model is a supraaural ANC headset, the Bose QC20 is an in-ear phone version of an ANC type device, and we also used the wireless AfterShokz BC model. Each device seemed to be applicable, they differed generally in comfort, size, weight and ergonomics.

Note that the task in this experiment was to detect and identify a limited number of possible sound sources. This is different than a so-called absolute localization task where the possible sound direction is ”continuous” in space. It is easier to select a sound source from three or five discrete positions than pointing to one perceived location. If having a limited, discrete number of potential source locations, subjects often do not localize at all. Nevertheless, they make their judgements based on the available information, which - in this case - was the visual representations of virtual loudspeakers. Subjects were not asked about externalization directly, but they did not mention any in-the-head localization problems during their selections. Even if they had it, they were not aware of it and not influenced by this phenomenon, neither with ANC nor with BC devices. Applying complex binaural rendering with HRTFs, individual parameter settings, equalization of headphones, etc., is not needed for simple applications. Using binaural spatialization in the horizontal plane with a BC headset, the minimum discernable angular difference between two successively presented sound sources was found to be around 10° and above [34]. However, accuracy was poorer than with headphone based reproduction. Localization errors between 30°–35° were measured in front, back, and sides. HRTFs were applied also in an experiment targeting the usability of BC headsets for VR applications. Azimuth accuracy was about the same as it was for a regular headphone, while for the same result in elevation accuracy, high frequency components were needed [63]. Both experiments applied binaural spatialization to determine the absolute localization accuracy.

V. CONCLUSIONS

A virtual listening test was installed in a VR scenario using three, five and seven source directions and four different excitation signals. 25 untrained subjects evaluated two bone conduction and another 25 participants evaluated two ANC-equipped headphones. The results showed that there is no statistically significant difference between the test signals. Furthermore, applying an 80 dB external distractor sound did not influence localization judgements with or without active noise cancellation. Applying stereo panning, three directions (F/L/R) can be used any time without errors. On the other hand, there is a significant difference between the mean scores of the 5-direction and the 7-direction scenarios in all cases. Five directions (in 45-degree spacing in the frontal hemisphere) can be used with 65-75% accuracy and it is recommended for real applications where errors are not critical. Based on the measurement results and the participants' subjective evaluations, seven directions are too demanding and inaccurate for applications. All ANC and BC devices performed equally in accuracy in a given task.

Future directions for research include recruiting visually impaired users, using different external or internal distractor sounds for BC and ANC, testing looped playback and a closed ear canal. Furthermore, rearranged virtual speaker setups (3-5 directions), testing in real life scenarios, and modification of the virtual environment to an explorable dynamic 3D scenario could provide additional information about the usability of the devices.

REFERENCES

[1] J. C. Middlebrooks and D. M. Green, "Sound localization by human listeners," *Annu. Rev. Psychol*, vol. 42, pp. 135–59, 1991. **DOI**: 10.1146/annurev.ps.42.020191.001031

[2] M. Vorländer, *Auralization*. Heidelberg: Springer, 2020.

[3] W. M. Hartmann, *Localization and Lateralization of Sound*. Cham: Springer International Publishing, 2021, pp. 9–45.

[4] G. Plenge, "On the differences between localization and lateralization," *The Journal of the Acoustical Society of America*, vol. 56, no. 3, pp. 944–951, 1974. **DOI**: 10.1121/1.1903353

[5] B. N. Walker and J. Lindsay, "Using virtual environments to prototype auditory navigation displays," *Ass. Technology*, vol. 17, no. 1, pp. 72–81, 2005. **DOI**: 10.1080/10400435.2005.10132097

[6] B. Kapralos, M. Jenkin, and E. Milios, "Virtual audio systems," *Presence: Teleoperators and Virtual Environments*, vol. 17, no. 6, pp. 527–549, 2008. **DOI**: 10.1162/pres.17.6.527

[7] T. Kuppanda, N. Degara, D. Worrall, B. Thoshkahna, and M. Müller, "Virtual reality platform for sonification evaluation," in *Proceedings of ICAD'15*, Graz, Austria. GeorgiaTech, 2015, pp. 117–124.

[8] G. C. Stecker, "Using virtual reality to assess auditory performance," *The Hearing Journal*, vol. 72, no. 6, p. 20, 2019. **DOI**: 10.1097/01.hj.0000558464.75151.52

[9] V. Hohmann, R. Paluch, M. Krueger, M. Meis, and G. Grimm, "The virtual reality lab: realization and application of virtual sound environments," *Ear and Hearing*, vol. 41, no. Suppl 1, p. 31S, 2020. **DOI**: 10.1097/AUD.0000000000000945

[10] J. Tonndorf, "Bone conduction." Berlin: Springer, 1976, pp. 37–84.

[11] S. Stenfelt, "Acoustic and physiologic aspects of bone conduction hearing," *Implantable bone conduction hearing aids*, vol. 71, pp. 10–21, 2011. **DOI**: 10.1159/000323574

[12] C. N. Hansen, *Understanding active noise cancellation*. Oxford: CRC Press, 1999.

[13] C.-Y. Chang and S.-T. Li, "Active noise control in headsets by using a low-cost microcontroller," *IEEE Transactions on industrial electronics*, vol. 58, no. 5, pp. 1936–1942, 2010. **DOI**: 10.1109/TIE.2010.2058071

[14] S. Spagnol, G. Wersényi, M. Bujacz, O. Bălan, M. Herrera Martínez, A. Moldoveanu, and R. Unnthorsson, "Current use and future perspectives of spatial audio technologies in electronic travel aids," *Wireless Communications and Mobile Computing*, vol. 2018, 2018. **DOI**: 10.1155/2018/3918284

[15] J. C. Lock, I. D. Gilchrist, G. Cielniak, and N. Bellotto, "Bone-conduction audio interface to guide people with visual impairments," in *International Conference on Smart City and Informatization*. Springer, 2019., pp. 542–553, **DOI**: 10.1007/978-981-15-1301-5-43.

[16] J. A. MacDonald, P. P. Henry, and T. R. Letowski, "Spatial audio through a bone conduction interface: Audición espacial a través de una interfase de conducción ósea," *International journal of audiology*, vol. 45, no. 10, pp. 595–599, 2006. **DOI**: 10.1080/14992020600876519

[17] B. F. Katz and L. Picinali, "Spatial audio applied to research with the blind," *Advances in sound localization*, pp. 225–250, 2011. **DOI**: 10.5772/15206

[18] G. A. Studebaker, "Placement of vibrator in bone-conduction testing," *Journal of speech and hearing research*, vol. 5, no. 4, pp. 321–331, 1962. **DOI**: 10.1044/jshr.0504.321

[19] M. McBride, T. Letowski, and P. Tran, "Bone conduction reception: Head sensitivity mapping," *Ergonomics*, vol. 51, no. 5, pp. 702–718, 2008. **DOI**: 10.1080/00140130701747509

[20] R. H. Margolis, R. H. Eikelboom, C. Johnson, S. M. Ginter, D. W. Swanepoel, and B. C. Moore, "False air-bone gaps at 4 khz in listeners with normal hearing and sensorineural hearing loss," *International journal of audiology*, vol. 52, no. 8, pp. 526–532, 2013. **DOI**: 10.3109/14992027.2013.792437

[21] B. Håkansson, A. Tjellström, and U. Rosenhall, "Hearing thresholds with direct bone conduction versus conventional bone conduction," *Scandinavian audiology*, vol. 13, no. 1, pp. 3–13, 1984. **DOI**: 10.3109/01050398409076252

[22] N. A. Watson, "Limits of audition for bone conduction," *The Journal of the Acoustical Society of America*, vol. 9, no. 4, pp. 294–300, 1938. **DOI**: 10.1121/1.1915936

[23] B. N. Walker and R. M. Stanley, "Thresholds of audibility for bone-conduction headsets," in *Proceedings of ICAD'05, Limerick, Ireland. Georgia Institute of Technology*, 2005, pp. 218–222.

[24] S. Stenfelt, "Transcranial attenuation of bone-conducted sound when stimulation is at the mastoid and at the bone conduction hearing aid position," *Otology & neurotology*, vol. 33, no. 2, pp. 105–114, 2012. **DOI**: 10.1097/MAO.0b013e31823e28ab

[25] S. Stenfelt and M. Zeitooni, "Binaural hearing ability with mastoid applied bilateral bone conduction stimulation in normal hearing subjects," *The Journal of the Acoustical Society of America*, vol. 134, no. 1, pp. 481–493, 2013. **DOI**: 10.1121/1.4807637

[26] M. Zeitooni, E. Mäki-Torkko, and S. Stenfelt, "Binaural hearing ability with bilateral bone conduction stimulation in subjects with normal hearing: Implications for bone conduction hearing aids," *Ear and hearing*, vol. 37, no. 6, pp. 690–702, 2016. **DOI**: 10.1097/AUD.0000000000000336

[27] D. Schonstein, L. Ferré, and B. F. Katz, "Comparison of headphones and equalization for virtual auditory source localization," *Journal of the Acoustical Society of America*, vol. 123, no. 5, p. 3724, 2008. **DOI**: 10.1121/1.2935199

[28] A. F. Jahn and J. Tonndorf, "Lateralization of bone-conducted sounds," *American journal of otolaryngology*, vol. 3, no. 2, pp. 133–140, 1982. DOI: 10.1016/S0196-0709(82)80044-6

[29] K. Kaga, M. Setou, and M. Nakamura, "Bone-conducted sound lateralization of interaural time difference and interaural intensity difference in children and a young adult with bilateral microtia and atresia of the ears," *Acta oto-laryngologica*, vol. 121, no. 2, pp. 274–277, 2001. DOI: 10.1080/000164801300043820

[30] B. N. Walker, R. M. Stanley, N. Iyer, B. D. Simpson, and D. S. Brungart, "Evaluation of bone-conduction headsets for use in multitalker communication environments," in *Proceedings of the human factors and ergonomics society annual meeting*, vol. 49, no. 17. SAGE Publications Sage CA: Los Angeles, CA, 2005. pp. 1615–1619, DOI: 10.1177/154193120504901725.

[31] R. W. Lindeman, H. Noma, and P. G. De Barros, "Hear-through and mic-through augmented reality: Using bone conduction to display spatialized audio," in *2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality*. IEEE, 2007. pp. 173–176, DOI: 10.1109/ISMAR.2007.4538843.

[32] R. M. Stanley, "Measurement and validation of bone-conduction adjustment functions in virtual 3d audio displays," Ph.D. dissertation, Georgia Institute of Technology, 2009.

[33] G. Wersényi, "Localization in a head-related transfer function-based virtual audio synthesis using additional high-pass and low-pass filtering of sound sources," *Acoustical science and technology*, vol. 28, no. 4, pp. 244–250, 2007. DOI: 10.1250/ast.28.244

[34] A. Barde, G. Lee, W. S. Helton, and M. Billinghurst, "Binaural spatialisation over a bone conduction headset: Elevation perception," in *Proceedings of ICAD'16*, Canberra, Australia. International Community on Auditory Display, 2016, pp. 1–4.

[35] S. M. Kuo, K. Kuo, and W. S. Gan, "Active noise control: Open problems and challenges," in *The 2010 International conference on green circuits and systems*. IEEE, 2010., pp. 164–169, DOI: 10.1109/ICGCS.2010.5543076.

[36] M. Guldenschuh, A. Sontacchi, M. Perkmann, and M. Opitz, "Assessment of active noise cancelling headphones," in *2012 IEEE Second International Conference on Consumer Electronics-Berlin (ICCE-Berlin)*. IEEE, 2012., pp. 299–303, DOI: 10.1109/ICCE-Berlin.2012.6336504.

[37] S. Liebich, J. Fabry, P. Jax, and P. Vary, "Signal processing challenges for active noise cancellation headphones," in *Speech Communication; 13th ITG-Symposium*. VDE, 2018, pp. 1–5.

[38] "MaxWhere 3D Platform," https://www.maxwhere.com//, 2022.

[39] B. Berki, "Better memory performance for images in maxwhere 3d vr space than in website," in *2018 9th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*, 2018., pp. 281–284, DOI: 10.1109/CogInfoCom.2018.8639956.

[40] "AfterShokz," https://shokz.com///, 2022.

[41] "AudioBone," https://www.goldendance.co.jp/English/product/p-ab01.html//, 2022.

[42] "Sennheiser," https://en-au.sennheiser.com/pxc450//, 2022.

[43] "Bose QuietComfort 25," {https://www.bose.com//}, 2022.

[44] H. Møller, D. Hammershøi, C. B. Jensen, and M. F. Sørensen, "Transfer characteristics of headphones measured on human ears," *Journal of the Audio Engineering Society*, vol. 43, no. 4, pp. 203–217, 1995.

[45] A. Kulkarni and H. S. Colburn, "Variability in the characterization of the headphone transfer-function," *The Journal of the Acoustical Society of America*, vol. 107, no. 2, pp. 1071–1074, 2000. DOI: 10.1121/1.428571

[46] G. Wersényi, "Comparison of Transfer Functions of Open Ear Canal Headsets Measured on a Dummy-Head and a Human Head," in *Forum Acusticum 23*. EAA, 2023, p. 5.

[47] L. A. Wilber and V. Goodhill, "Real ear versus artificial mastoid methods of calibration of bone-conduction vibrators," *Journal of Speech and Hearing Research*, vol. 10, no. 3, pp. 405–416, 1967. DOI: 10.1044/jshr.1003.405

[48] R. H. Margolis and S. M. Stiepan, "Acoustic method for calibration of audiometric bone vibrators," *The Journal of the Acoustical Society of America*, vol. 131, no. 2, pp. 1221–1225, 2012. DOI: 10.1121/1.3675007

[49] B. Karaböce, "Force sensitivity determination of artificial mastoid used for audiometric bone-conduction measurements," in *2020 IEEE International Symposium on Medical Measurements and Applications (MeMeA)*. IEEE, 2020., pp. 1–5. DOI: 10.1109/MeMeA49120.2020.9137333.

[50] S. Li and J. Peissig, "Measurement of head-related transfer functions: A review," *Applied Sciences*, vol. 10, no. 14, p. 5014, 2020. DOI: 10.3390/app10145014

[51] C. I. Cheng and G. H. Wakefield, "Introduction to head-related transfer functions (hrtfs): Representations of hrtfs in time, frequency, and space," in *AEA Convention 107*. Audio Engineering Society, 1999.

[52] T. M. Voong and M. Oehler, "Auditory spatial perception using bone conduction headphones along with fitted head related transfer functions," in *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 2019., pp. 1211–1212, DOI: 10.1109/VR.2019.8798218.

[53] M. Ohuchi, Y. Iwaya, Y. Suzuki, and T. Munekata, "Training effect of a virtual auditory game on sound localization ability of the visually impaired," in *Proc. of ICAD'05*, Limerick, Ireland, 2005, pp. 283–286.

[54] P. Majdak, M. J. Goupell, and B. Laback, "3-d localization of virtual sound sources: Effects of visual environment, pointing method, and training," *Attention, perception, & psychophysics*, vol. 72, no. 2, pp. 454–469, 2010. DOI: 10.3758/APP.72.2.454

[55] M. A. Steadman, C. Kim, J.-H. Lestang, D. F. Goodman, and L. Picinali, "Short-term effects of sound localization training in virtual reality," *Scientific Reports*, vol. 9, no. 1, pp. 1–17, 2019.

[56] S. Bech, "Selection and training of subjects for listening tests on sound-reproducing equipment," *Journal of the Audio Engineering Society*, vol. 40, no. 7/8, pp. 590–610, 1992.

[57] M. D. Good, R. H. Gilkey, and J. M. Ball, "The relation between detection in noise and localization in noise in the free field," *Binaural and spatial hearing in real and virtual environments*, pp. 349–376, 1997.

[58] E. A. Macpherson and J. C. Middlebrooks, "Localization of brief sounds: effects of level and background noise," *The Journal of the Acoustical Society of America*, vol. 108, no. 4, pp. 1834–1849, 2000. DOI: 10.1121/1.1310196

[59] D. S. Brungart, B. D. Simpson, and A. J. Kordik, "Localization in the presence of multiple simultaneous sounds," *Acta Acustica united with Acustica*, vol. 91, no. 3, pp. 471–479, 2005.

[60] M. F. Mueller, A. Kegel, S. M. Schimmel, N. Dillier, and M. Hofbauer, "Localization of virtual sound sources with bilateral hearing aids in real- istic acoustical scenes," *The Journal of the Acoustical Society of America*, vol. 131, no. 6, pp. 4732–4742, 2012. DOI: 10.1121/1.4705292

[61] K. Lopatka, J. Kotus, and A. Czyzewski, "Detection, classification and localization of acoustic events in the presence of background noise for acoustic surveillance of hazardous situations," *Multimedia Tools and Applications*, vol. 75, no. 17, pp. 10 407–10 439, 2016. DOI: 10.1007/s11042-015-3105-4

[62] K. R. May and B. N. Walker, "The effects of distractor sounds presented through bone conduction headphones on the localization of critical environmental sounds," *Applied ergonomics*, vol. 61, pp. 144–158, 2017. DOI: 10.1016/j.apergo.2017.01.009

[63] M. V. Tray, M. do VM da Costa, M. Regener, C. Reuter, and M. Oehler, "Suitability of bone conduction headphones for virtual acoustic environments," in *AES 2022 Int. Audio for Virtual and Augmented Reality Conference*. Audio Engineering Society, 2022.

György Wersényi was born in 1975 in Győr, Hungary. He received his MSc degree in electrical engineering from the Technical University of Budapest in 1998 and PhD degree from the Brandenburg Technical University in Cottbus, Germany. Since 2002 he has been member of the Department of Telecommunications at the Széchenyi István University in Győr. From 2020 to 2022 he was the dean of Faculty of Mechanical Engineering, Informatics and Electrical Engineering, as well as the scientific president of the Digital Development Center at the university. Currently, he is a full professor, member of the European Acoustics Association (EAA) and the Audio Engineering Society (AES). His research focus is on acoustic measurements, virtual and augmented reality solutions, sonification, cognitive infocommunications, and assistive technologies.