AN OVERVIEW OF TOOLS, STANDARDS, AND SPATIAL AUDIO REPRODUCTION SYSTEMS FOR USE IN LISTENING TESTS

Dominik Kisić, Marko Horvat, Kristian Jambrošić

1 Introduction

Sound is a very multidimensional phenomenon. Some of its properties are purely physical and can be measured in the form of speed, air pressure amplitude, velocity or acceleration, the same as concepts like loudness, intensity, energy, pitch, timbre [IVANCEVIC(2007)]. Nonetheless, human perception of sound and influence of sound on people is a much more complex problem [BECH & ZACHAROV(2006)]. Unfortunately, it is still not possible to put a metering device directly on listeners and get a specific value out of it, although the popularity of research that includes measuring the EEG [NAWROCKA & HOLEWA(2014)], EKG [BHASKAR ET AL.(2012)] and other physiological signals responses on audio stimuli have been on the rise in the last a couple of decades.

The alternative way of assessing how listeners perceive audio is to ask them directly in order to quantify their experience. This is the most common form of perceptual evaluation that often takes the form of a formal listening test [BECH & ZACHAROV(2006)]. Although testing of audio quality and other properties of sound reproduction systems have existed in some form of listening tests ever since the first ”Mr. Watson, come here. I want to see you.” that was said at Graham Bell’s Lab, many improvements in the listening test methodology have been made over time. Since all listening tests have to take into consideration the uncertainty linked with the human factor, it is a field of science that is still being explored.

This article presents a partial overview of current listening test tools and standards and an explanation of the differences benefits and downsides, of Ambisonics and binaural formats for performing listening tests. Although Ambisonics can be decoded into a binaural format, the comparison considered in this article is mostly made between the basic reproduction systems of the two formats: multichannel loudspeaker system reproduction for Ambisonics and headphone reproduction for binaural.

Not many studies have explored the differences between reproduction systems for conducting listening tests. KOEHL ET AL.(2011) found that “*reproduction methods provided consistent similarity and preference judgments”*, while BRINKMAN ET AL.(2015) concluded that “*People could distinguish between mono, stereo, Dolby surround and 3D audio of a wasp”*, and that there were “*significant effects for audio techniques on people’s self-reported anxiety, presence, and spatial perception”*. They also found that there was “n*o difference in virtual world experience between stereo and 3D audio”*. These findings show that the listeners can notice the difference between different reproduction systems. On the other hand, that difference does not have a significant influence on the overall listening experience.

Nonetheless, the literature review showed no information about the sound reproduction system in research of subjective evaluation of sound insulation and acoustic comfort, which is one of the authors’ interests and has potential for scientific contribution.

2 Standards for performing the listening tests

In the field of audio, a number of standards or recommendations cover a wide range of topics, from measurement devices to perceptual evaluation methods for telecommunications or audio systems. They try to provide information on the best-agreed practice for performing listening tests. Standards do not always suggest the most advanced method, just the best-agreed method. They are always application-oriented and should not be used interchangeably. Although it is always a possibility to take a core idea of a standard and modify it for a specific need, at that point it cannot be called a standardized test anymore [BECH & ZACHAROV(2006)], [ZACHAROV & WICKELMAIER(2007)].

The standardization of listening tests has not yet gotten to the point where specific attributes of the test are stated. In case of a listening test where more advanced and complex methods need to be used, there are many other studies that could be referenced for their methodology (e.g. [DE MAN ET AL.(2016)]). Key organizations are ITU-R and ITU-T.

2.1 ITU-T telecommunication applications

They are focused on telecommunication applications, i.e. speech codecs, echo cancellation, etc. They are speech-oriented, Mean Opinion Score (MOS) based, a mostly narrowband (300 Hz – 3400 Hz), or wideband (100 Hz – 7000 Hz), the usual number of naïve assessors is from 12 up to 36. It covers a number of methods such as Absolute category rating (ACR), Comparison category rating (CCR), and Degradation category rating (DCR). The key standard is ITU-T P.800. [BECH & ZACHAROV(2006)]

2.2 ITU-R Radio communication section

Audio applications, e.g. audio codecs, basic audio quality (BAQ) based, full-band audio applications (20 Hz – 20000 Hz), a usual number of expert assessors is 20. Key standards are ITU-R BS.1116-1 and BS.1534-1. [BECH & ZACHAROV(2006)]

2.3 Current key standards and recommendations for performing listening tests

ITU-T Recommendation P.800: [ITU-T(1996)]

- Absolute category rating (ACR): single stimulus method, dependent variable (5–point categorical scale: listening quality, listening effort, loudness preference), Independent variables (system/codec, speech sample, talker gender, sentence, listening level), naive subjects (24–36), ANOVA based analysis.
- Comparison category rating (CCR): paired comparison, hidden reference, dependent variable (7–point categorical scale), independent variables (system/codec, speech sample, talker), naive subjects (24–36), ANOVA based analysis.
- Degradation category rating (DCR): fixed reference paired comparison, dependent variable (5–point degradation categorical scale), independent variables (system/codec, speech sample, talker, background), naive subjects (32), ANOVA based analysis.

ITU-R Recommendation BS.1116-1 ABC/HR: [ITU-R(1997)] Evaluation of small impairment (only), double-blind triple stimulus hidden reference, dependent variable (5–point continuous rating scale: basic audio quality, stereophonic image quality, front image quality, impression of surround quality), independent variables (system/codec, program, subject), expert assessors (20 with defined selection process), ANOVA based analysis. Listening room definition and loudspeaker setup definition.

ITU-R Recommendation BS.1534-1 MUSHRA: [ITU-R(2003)] Double-blind multi-stimulus with hidden reference and hidden anchors, dependent variable (0–100 continuous quality scale with 5 equal intervals, basic audio quality, stereophonic image quality, from image quality, impression of surround quality), independent variables (system/codec, program, subject), partially screen subjects (more than 20).

3 Listening test tools

Several tools for the design of listening tests have been developed over the years. *When performing listening tests, audio stimuli are to be presented to subjects and their responses are to be collected. While these steps can be performed manually, this is a highly complex, time-consuming, and very error-prone approach. Nowadays, computer-based systems are available to automate stimulus presentation and/or data collection, avoiding most of the limitations associated with a manual procedure. Such software tools are highly desirable in listening test work to lighten the burden on the experimenter, and to provide a better control over the experiment. This latter aspect leads to a reduction in experimental error, as well as providing robustness. Additionally, using a computer-based system allows for similar experiments to be perfectly duplicated or repeated at different locations or times.* [BECH & ZACHAROV(2006)]

Knowledge and use of these tools can save precious development time and money during the experimental design. Although some of them have templates for many listening tests, some of the tools can be modified to a specific need. Most popular and advanced listening test tools today are:

- HULTI-GEN [LEE(2015)]: Max/MSP based, very versatile tool
- WAET [JILLINGS ET AL.(2015)]: JavaScript browser based, very versatile tool
- WhisPER [CIBA ET AL.(2012)]: Matlab based
- APE [DE MAN & REISS(2014)]: Matlab based
- Scale [GINER(2013)]: Matlab based
- MUSHRAM [VINCENT ET AL.(2012)]: Matlab based
- BeaqlesJS [KRAFT & ZÖLZER(2014)]: JavaScript based
- STEP [LAB(2019)]: Windows based, ITU-R BS.1116-1, ITU-R BS.1534-1, ITU-T P.800 ACR
- webMUSHRA [SCHOEFFLER ET AL.(2018)]: web-based ITU-R BS.1534-1
- GuineaPig [HYNNINEN & ZACHAROV(2012)]

A side-by-side comparison of some of these listening test tools can be seen in Table 1.

4 Ambisonics systems

Ambisonics is a full-sphere surround sound format. In listening test applications, Ambisonics is usually used with a multichannel reproduction system that consists of loudspeakers placed around the listening sweet spot not only in the listening plane, but also above and below it. An example of a listening room enhanced with an Ambisonics system can be found in HORVAT ET AL.(2013), and another one can be seen in Figure 1.

Table 1. Comparison of the listening test tools. [JILLINGS ET AL.(2016)] [BECH & ZACHAROV(2006)]

Toolbox	**APE**	**Beaqles JS**	**Hulti-gen**	**MUSH RAM**	**Scale**	**WhisP ER**	**WAET**	**STEP**	**Guinea Pig**
Language	**Matlab**	**JS**	**Max**	**Matlab**	**Matlab**	**Matlab**	**JS**		**Linux**
Remote		**+**		**+**			**+**		
MUSHRA (ITU-R BS. 1534)		+	+	+			+	+	
APE	+						+		
Rank Scale				+			+		+
Likert Scale				+			+		
ABC/HR (ITU-R BS. 1116)				+			+	+	+
-50 to 50 Bipolar with Reference				+			+		
Absolute Category Rating Scale (ITU-T P.800)				+			+		+
Degradation Category Rating Scale (ITU-T P.800)				+			+		+
Comparison Category Rating Scale (ITU-T P.800)				+			+		+
9 Point Hedonic Category Rating Scale				+			+		
ITU-R 5 Continuous Scale				+			+		
Pairwise / AB Test				+			+		+
Multi – Attribute Ratings				+			+		
ABX Test		+		+			+	+	+
Adaptive Psychological methods						+			
Repetory Grid Technique						+			
Semantic Differential					+	+	+		
n-Alternative Forced Choice					+				

4.1 What is Ambisonics

Ambisonics is a method of codifying a sound field taking into account its directional properties. In traditional multichannel audio (e.g., stereo, 5.1 and 7.1 surround) each channel has the signal corresponding to a given loudspeaker. Instead, in Ambisonics each channel has information about certain physical properties of the acoustic field, such as the pressure or the acoustic velocity [ARTEAGA(2015)].

The fundamental theory of Ambisonics can be divided into a couple of basic principles:

At zeroth order: Ambisonics has information about the pressure field at the origin (the recording made with an omnidirectional microphone). The channel for the pressure field is conventionally called W.

At first order: Ambisonics adds information about the acoustic velocity at the origin (recording of three figure-of-eight microphones at the origin, along each of the three axes). These channels are called X, Y, and Z. Following the Euler equation, the velocity vector is proportional (up to some approximation) to the gradient of the pressure field along each one of the axis.

At second and higher orders: Ambisonics adds information about higher-order derivatives of the pressure field. [ARTEAGA(2015)]

An Ambisonics recording is usually recorded with a specialized 4-channel microphone that gives a first order A-format recording. That kind of microphone usually has four capsules placed in a tetrahedral configuration, e.g. RODE NT-SF1 microphone. For higher-order Ambisonics, recordings a microphone with more capsules is required. First-order Ambisonics A-format corresponds to the direct recordings of each of the four capsules of the microphone. Ambisonics recordings are never kept in the A-format, but are rather transformed into the B-format. The B-format also has four channels, but it corresponds to the omnidirectional information (W), and three directional channels where each channel holds the information about one of the axes (X, Y, Z). Although this gives the complete information about the spatial sound, before playback, the B-format has to be decoded into a set of signals that carry information for each of the loudspeaker channels of the reproduction system.

Figure 1. Example listening room enhanced with an Ambisonics multichannel system

4.2 In which fields and situations is it used?

Although Ambisonics had its beginnings with Michael Gerzon in the 1970's [GERZON(1992)] with the basic principles dating all the way back to 1930s [MALHAM(2019)], it was not met with commercial success at that time. During the 1990s the theory of higher-order Ambisonics (HOA) was founded, which brought new light onto the Ambisonics format, and it still remains a topic of research in the academic community today [STEIN & GOODWIN(2019)]. Lately, Ambisonics has found new applications with the increasing popularity of virtual reality [FARINA ET AL.(2018)], [SHERBOURNE(2017)] but also as a sound reproduction system for listening tests [CHMELIK ET AL.(2019)]. It was even used as a format of audio distribution for different venues and broadcasts in real-time [FRANK & SONTACCHI(2017)].

The growing internet community opened a new possibility to experience virtual reality (VR) enhanced with surround sound. Platforms such as Youtube [YOUTUBE(2019)], Google [GOOGLE(2019)], Facebook [FACEBOOK(2019)] started creating a framework for VR which would include and allow integration of spatial audio.

Furthermore, audio production companies started the development of tools for easier manipulation and transformation of spatial audio signals. These tools can be easily used with existing digital audio workstations (DAWs) with a similar workflow as traditional production of audio, which is supporting a smoother transition to novice technologies [Adobe(2019)], [STEINBERG(2019)], [COCKOS INC.(2019)], [KRONLACHNER(2019)], [BLUE RIPPLE SOUND LIMITED(2019)], [AUDIO EASE B.V.(2019)], [SPOOK(2019)], [DEARVR(2019)].

4.3 Advantages and drawbacks of Ambisonics

Advantages: Bigger potential number of participants at the same time. Real feel of the situation, i.e. people do not usually sit at home in an empty room and listen to sounds on headphones. Fully immersive surround sound which would mean that the sound is heard by the listener from every direction, 360 and also up and down directions. Perceived spatial localization properties can be improved by the rotation of the head [MCANALLY & MARTIN(2014)].

A relatively small number of audio channels is needed to describe complete surround sound spatial audio. [STEIN & GOODWIN(2019)]. Ambisonic formats are independent of the reproduction system; i.e. an Ambisonics signal can be decoded to any loudspeaker configuration or for binaural or transaural rendering [MCKEAG & MCGRATH(1996)], [HELLER ET AL.(2008)], [WIGGINS(2007)], [ENGEL ET AL.(2019)].

Drawbacks: downsides of Ambisonics would include complexity of equipment for reproduction system, larger file sizes, different conversions, specialized room, cost.

5 Binaural systems

Localization in binaural audio is achieved by perceiving the interaural time and level differences. Considering hardware requirements, binaural listening test are rather simple and just require headphones. The beginnings of binaural reproduction format go back to 1933, when one of the divisions of Bell laboratories demonstrated a dummy human head with microphones in the ears. The signals from these microphones were being played directly back into the listener’s ears using headphones. Around the same time, a Connecticut radio station broadcast a number of shows in binaural stereo, using two separate radio frequencies - the listener had to use two separate radios to feed two earphones. Although the idea of binaural recording is attractive, it turns out to have very variable effectiveness for different people, and to be unsuitable for playback through loudspeakers. For these reasons, it has remained rather a niche approach to recording for many years. [HODGES(2017)]. Nowadays research is still done considering binaural formats, e.g. in the field of Television and film. Lopez et. al.

[LOPEZ ET AL.(2016)] are developing ways of enhancing accessibility for TV and film for visually impaired.

5.1 HRTFs vs Mono or Stereo headphone signals, transaural

An important difference has to be made between simple stereo or mono headphone signal, and binaural signals. Stereo and mono headphone signals are made by just simple reproduction of audio signals over the headphone speakers, the same way as it would be reproduced over regular speakers. In the case of binaural reproduction, the signals either need to be modified before reproduction or they need to be recorded with a head and torso simulator i.e. a dummy head. An example of a dummy head can be seen in Figure 2.

Figure 2. Dummy head - Bruel and Kjaer head and torso simulator type 4128t.

Modifications of the regular audio signal for binaural reproduction are best made by convolving them with appropriate head-related transfer functions (HRTF). Head-related transfer functions are impulse responses recorded with a dummy head that has microphones mounted inside the ear canal and with the source signal coming from different directions. The set of HRTFs can be made by rotating the dummy head with 1 or similar steps in the horizontal plane, but also with a bit more complex rotation in the vertical direction. Transaural rendering can also be made by using basic binaural configuration but enhancing it with crosstalk cancellation so it can be played back over a pair of speakers. This technology is used in aixCAVE at RWTH Aachen University [WEFERS ET AL.(2015)].

5.2 Head tracking

Head tracking is a software application that monitors a user's head position and orientation. It's often used alongside face and eye tracking to help and improve human-computer *interaction (HCI). Head tracking is often used to simulate the experience of freely looking* *around in virtual (VR) or augmented reality (AR), allowing the user to experience an immersive and natural way to look around in virtual environments. There are a number of methods used for head tracking. Screen quality and head-tracking responsiveness are some of the most significant user experience differentiators between high-end headsets, like Oculus Rift, and low-end headsets and smartphone holding designs like Google Cardboard. Devices that use smartphones often rely on phone accelerometers and gyroscopes. High-end headsets have more accurate tracking with precise sensors, along with other systems including infrared LEDs, cameras and magnetometers. Because head tracking in AR or VR can simulate real-life experiences, it can fool the brain even better than standard viewing for more engaging and immersive user experience* [WHATIS.COM(2017)].

With head tracking information it is possible to choose an appropriate HRTF which would not rotate an entire sound field but would just allow the listener to move its head during a listening test, which would give a more realistic and immersive feeling to the listening experience.

5.3 Advantages and drawbacks of binaural systems

Since every experiment and application has its specific set of requirements, it has to be mentioned that some of pros and/or cons are situation-dependent, so this list gives generalized information.

Advantages: In its basic form setup, a binaural listening test is definitely not as budget-consuming as other audio reproduction options. In addition, the availability of the system is beneficial because headphones are a part of every household. Listening test rooms that use headphones are usually much less expensive than specialized multi-channel rooms. Also, in the case where there is no specific need for extremely controlled setup, the experimenters are able to distribute the listening test online to subjects all over the globe, and that way more participants can be reached in which case the data-set can give a better representation of a general population.

Drawbacks: The downsides of binaural would definitely include a limitation of a natural feeling concerning localization because it is either needed for the head of the participant to remain fixed or to use head tracking with much more complex set of HRTF. In the case where there is a movement of the head, a head-tracking enhancement of the system should be engaged. Headphone positioning variability presents a big problem considering repeatability of the experiments conducted with headphones and binaural [PAQUIER & KOEHL(2010)].

6 Conclusions and Further work

As can be seen, both Ambisonics and binaural approach in listening test design have many benefits, but also some downsides. In addition, listening test standards just give a general recommendation about the reproduction system without specifying or recommending situations in which it would be beneficial to use one or the other. Many tools for listening tests include some of those standards and usually provide a useful jump-start in the design of listening tests.

Future work will focus on the comparison of the reproduction system and further investigation of differences with focus on use in subjective evaluation of acoustic comfort. *Preliminary design of the listening tests has already started. The tool is being developed for* *listening tests that will allow sound samples to be easily reproduced over different sound reproduction systems. Principal block scheme of the system can be seen in Figure 3. Although this system is still in its early stage of development, it will be continuously improved through further use and implementation of individual listening tests. The addition of a real-time head tracking function to this system is something that will be added in the future.* [KISIC ET AL.(2019)] This research is primarily meant to be conducted in two similar listening rooms enhanced with multichannel reproduction systems with capabilites of Ambisonic playback, e.g. AuraLAB at the University of Zagreb [HORVAT ET AL.(2013)], or TGM – Fachbereich für Akustik und Bauphysik in Vienna (Figure 1).

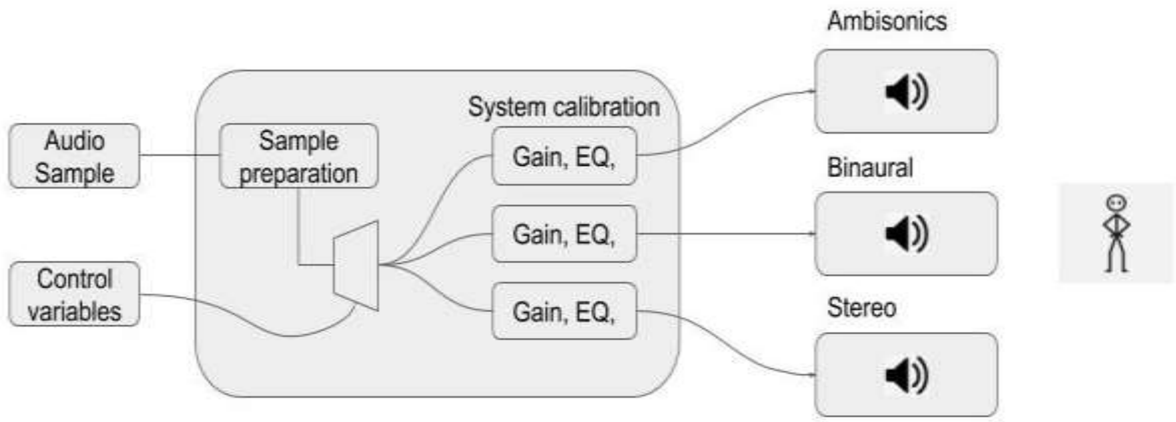

Figure 3. Principal block scheme of the system.

Research can also expand to investigate the position of the low-frequency subwoofer in the listening room [BECH & ZACHAROV(2006)], and also the possibility of expansion of binaural headphone reproduction with a subwoofer which could be heard with the vibrations throughout the listener's body.

Acknowledgements

This work has been supported by the European Commission project H2020-MSCA-RISE-2015: Advanced physical acoustic and psycho acoustic diagnostic methods for innovation in building acoustics (PAPABUILD), grant number 690970.

References

ADOBE (2019). Support for vr workflows. URL: https://helpx.adobe.com/premierepro/using/VRSupport.html accessed on August 25th 2019.

ARTEAGA, D. (2015). Introduction to ambisonics. URL: https://www.researchgate.net/publication/280010078 accessed on August 25th 2019.

AUDIO EASE B.V. (2019). Audioease 360 pan suite. URL: https://www.audioease.com/360/ accessed on August 25th 2019.

BECH, S., & ZACHAROV, N. (2006). Perceptual Audio Evaluation - Theory, Method and Application. John Wiley and Sons.

BHASKAR, R., CHOUDHURI, R., PANDEY, A., BANDOPADHYAY, S., SARANGI, S., & KUMAR GHATAK, S. (2012). Effect of rotating acoustic stimulus on heart rate variability in healthy adults. The open neurology journal, 6, 71–77.

BLUE RIPPLE SOUND LIMITED (2019). A complete audio workflow for virtual reality. URL: http://www.blueripplesound.com/ accessed on August 25th 2019.

BRINKMAN, W.-P., HOEKSTRA, A. R (2015). The effect of 3d audio and other audio techniques on virtual reality experience. Studies in health technology and informatics, 219, 44–48.

CHMELIK, V., BENKLEWSKI, J., RYCHTARIKOVA, M., KISIC, D., JAMBROSIC, K., HORVAT, M., & MULLER, H. (2019). The preliminary study on subjective rating of different floors characterised by ln,w+ci,50-2500. In PROCEEDINGS of the 23rd International Congress on Acoustics. Aachen, Germany.

CIBA, S., WLODARSKI, A., & MAEMPEL, H.-J. (2012). Whisper – a new tool for performing listening tests. 126th Audio Engineering Society Convention 2009, 1.

COCKOS INC. (2019). Reaper digital audio workstation. URL: https://www.reaper.fm/ accessed on August 25th 2019.

DE MAN, B., JILLINGS, N., MOFFAT, D., REISS, J., & STABLES, R. (2016). Subjective comparison of music production practices using the web audio evaluation tool. In Conference: 2nd AES Workshop on Intelligent Music Production.

DE MAN, B., & REISS, J. (2014). Ape: Audio perceptual evaluation toolbox for matlab. In 136th Convention ofthe AES.

DEARVR (2019). dearvr spatial connect: The future of audio production for vr and ar. URL: http://dearvr.com/ accessed on August 25th 2019.

ENGEL, I., HENRY, C., AMENGUAL GAR´I, S., W ROBINSON, P., POIRIER-QUINOT, D., & PICINALI, L. (2019). Perceptual comparison of ambisonics-based reverberation methods in binaural listening. In EAA Spatial Audio Signal Processing Symposium. Paris, France.

FACEBOOK (2019). Facebook 360 spatial audio workstation. URL: https://facebook360.fb.com/spatialworkstation/ accessed on August 25th 2019.

FARINA, A., PINARDI, D., BINELLI, M., EBRI, M., & EBRI, L. (2018). Virtual reality for subjective assessment of sound quality in cars. In Conference: AES 144th International Pro Audio Convention. Milan, Italy.

FRANK, M., & SONTACCHI, A. (2017). Case study on ambisonics for multi-venue and multi-target concerts and broadcasts. Journal of the Audio Engineering Society, 65, 749–756.

GERZON, M. A. (1992). General metatheory of auditory localisation. In Audio Engineering Society Convention 92. URL http://www.aes.org/e-lib/browse.cfm?elib=6827

GINER, A. V. (2013). Scale - a software tool for listening experiments. In AIA/DAGA Conference on Acoustics.

GOOGLE (2019). Resonance audio overview. URL: https://developers.google.com/resonanceaudio/develop/overview
accessed on August 25th 2019.

HELLER, A., LEE, R., & BENJAMIN, E. (2008). Is my decoder ambisonic? In Audio Engineering Society

Convention 125. URL http://www.aes.org/e-lib/browse.cfm?elib=14705

HODGES, P. (2017). History of recording. URL: https://ambisonic.info/history.html accessed on August 25th 2019.

HORVAT, M., JAMBROSIC, K., & DOMITROVIC, H. (2013). Construction of an auralization laboratory based on multichannel loudspeaker system. In Proceedings of the International Conference on Acoustics, (p. p. 2356 2359.). Berlin, Germany.

HYNNINEN, J., & ZACHAROV, N. (2012). Guineapig-a generic subjective test system for multichannel audio. URL: https://www.researchgate.net/publication/229045731 accessed on August 25th 2019.

IAN MCANALLY, K., & LANCE MARTIN, R. (2014). Sound localization with head movement: Implications for 3-d audio displays. Frontiers in neuroscience, 8, 210.

ITU-R (1997). Recommendation ITU-R BS. 1116-1: Methods for the subjective assessment of small impairments in audio systems including

multichannel sound systems. International Telecommunication Union.

ITU-R (2003). Recommendation ITU-R BS.1534-1: Method for the subjective assessment of intermediate quality levels of coding systems.

International Telecommunication Union.

ITU-T (1996). Recommendation ITU-T P. 800: Methods for subjective determination of transmission quality. International Telecommunication Union.

IVANCEVIC, B. (2007). Elektroakustika. Zagreb, Croatia: University of Zagreb, 3.2 ed.

JILLINGS, N., DE MAN, B., MOFFAT, D., & REISS, J. (2016). Web audio evaluation tool: A framework for subjective assessment of audio. In Conference: Proc. 2nd Web Audio Conference At: Atlanta, Georgia, USA.

JILLINGS, N., MOFFAT, D., DE MAN, B., & REISS, J. (2015). Web audio evaluation tool: A browser-based listening test environment. In 12th Sound and Music Computing conference.

KISIC, D., HORVAT, M., & JAMBROSIC, K. (2019). A methodology and a tool for interchangeable reproduction of sound samples in listening tests through different sound reproduction systems. In proceedings of the 23rd International Congress on Acoustics. Aachen, Germany.

KOEHL, V., PAQUIER, M., & DELIKARIS-MANIAS, S. (2011). Comparison of subjective assessments from listening tests through headphones and loudspeaker setups. In 131st Audio Engineering Society Convention 2011, vol. 1.

KRAFT, S., & ZÖLZER, U. (2014). Beaqlejs: Html5 and javascript based framework for the subjective evaluation of audio quality. In Linux Audio Conference.

KRONLACHNER, M. (2019). ambix v0.2.8 – ambisonic plugin suite. URL: http://www.matthiaskronlachner.com/?p=2015 accessed on August 25th 2019.

LAB, A. R. (2019). Step - subjective training and evaluation program. URL: https://www.audioresearchlabs.com/arl-step/ accessed on August 25th 2019.

LEE, H. (2015). Toward the development of a universal listening test interface generator in max. In AES Convention 138.

LOPEZ, M. J., KEARNEY, G. C., & HOFSTADTER, K. (2016). Enhancing audio description: Sound design, spatialization and accessibility in film and television. In Reproduced Sound Conference. Southampton, UK.

MALHAM, D. (2019). The early years of Ambisonics at York. In Audio Engineering Society Conference: 2019 AES International Conference on Immersive and Interactive Audio. URL http://www.aes.org/e-lib/browse.cfm?elib=20404

MCKEAG, A., & MCGRATH, D. S. (1996). Sound field format to binaural decoder with head tracking. In Audio Engineering Society Convention 6r. URL http://www.aes.org/e-lib/browse.cfm?elib=7477

NAWROCKA, A., & HOLEWA, K. (2014). The analysis of the different frequencies sound waves effect on the eeg signal. Solid State Phenomena, 208, 177–182.

PAQUIER, M., & KOEHL, V. (2010). Audibility of headphone positioning variability. 128th Audio Engineering Society Convention 2010, 2.

SCHOEFFLER, M., BARTOSCHEK, S., STÖTER, F.-R., ROESS, M., WESTPHAL, S., EDLER, B., & HERRE, J. (2018). webmushra — a comprehensive framework for web-based listening tests. Journal of Open Research Software, 6.

SHERBOURNE, S. (2017). Ambisonics and vr/360 audio in pro tools — hd. URL: http://www.avidblogs.com/ambisonics-vr360-audiopro-tools-hd/ accessed on August 25th 2019.

SPOOK (2019). Spooksyncvr for gopro vr player. URL: https://www.spook.fm/spooksyncvr/ accessed on August 25th 2019.

STEIN, E., & GOODWIN, M. M. (2019). Ambisonics depth extensions for six degrees of freedom. In Audio Engineering Society Conference: 2019 AES international conference on headphone technology. URL http://www.aes.org/e-lib/browse.cfm?elib=20514

STEINBERG (2019). 3d mixes for ambisonics. URL: https://steinberg.com accessed on August 25th 2019.

VINCENT, E., G. JAFARI, M., & PLUMBLEY, M. (2012). Preliminary guidelines for subjective evaluation of audio source separation algorithms. In ICA Research Network Workshop.

WEFERS, F., PELZER, S., BOMHARDT, R., MÜLLER-TRAPET, M., & VORLAENDER, M. (2015). Audiotechnik des aixcave virtual reality-systems. In Conference: 41. Jahrestagung f¨ur Akustik (DAGA2015). Nuremberg, Germany.

WHATIS.COM (2017). Head tracking. URL: https://whatis.techtarget.com/definition/head-tracking accessed on August 25th 2019.

WIGGINS, B. (2007). The generation of panning laws for irregular speaker arrays using heuristic methods. In Audio Engineering Society Conference: 31st International Conference: New Directions in High Resolution Audio. URL http://www.aes.org/e-lib/browse.cfm?elib=13946

YOUTUBE (2019). Video upload requirements for spatial audio. URL: https://support.google.com/youtube/answer/6395969 accessed on August 25th 2019.

ZACHAROV, N., & WICKELMAIER, F. (2007). Perceptual audio evaluation—theory, method and application. Journal of The Acoustical Society of America, 122.

Summary

An overview of tools, standards, and spatial audio reproduction systems for use in listening tests. Listening tests are the most used method for evaluation of human preferences and perception of sound. This paper presents an overview of main tools and standards for the design of listening tests. Several key standards for ITU-T and ITU-R recommendations are covered, such as ITU-R BS.1116-1 and BS.1534-1 and ITU-T P.800. The list of modern listening test tools for effective design of listening tests is presented as well, showing which standardized listening tests each of the tools covers, along with the associated development software or programming language used for building each tool. Furthermore, the paper gives an overview of the main aspects of Ambisonics and binaural systems as the most commonly used platforms for spatial sound reproduction, and discusses their advantages and disadvantages for specific use in listening tests. Ideas for further study of the influence of different reproduction systems on listening tests are presented as well.

Keywords

Listening tests tools; Listening test standards; Ambisonics; Binaural;