PROCEEDINGS of the
23rd International Congress on Acoustics

9 to 13 September 2019 in Aachen, Germany

Spatial Cue Distortions Within a Virtualized Sound Field Caused by an Additional Listener

Sergio Luiz Aguirre[1,2]; Lars Bramsløw[2]; Thomas Lunner[2]; William McAllister Whitmer[1]

[1] Hearing Sciences – Scottish Section, University of Nottingham, Glasgow, UK

[2] Eriksholm Research Centre, Snekkersten, DK

ABSTRACT

Realistically, we are rarely alone in a central position with respect to our acoustic environment, yet virtual sound fields are usually evaluated in this manner. Sound presentation with more than one person present using sound source virtualization can be useful to invoke natural behaviors in auditory research. Interaural time and level differences (ITDs and ILDs, respectively) were measured on a symmetric mannequin (HATS). Sound sources were virtualized using vector-based amplitude panning and presented from a horizontal ring of 24 loudspeakers. The influence of a second listener was simulated by positioning a second mannequin (KEMAR) along the midcoronal plane of HATS (i.e., shoulder-to-shoulder). ITDs and ILDs were measured with HATS centered and KEMAR 0.5-1.0 m to the right of center, or with HATS and KEMAR 0.25-0.75 m to the left and right of center. Results were compared to HATS alone. When HATS was centered in the sound field, the spatial cue distortions were small, independent of KEMAR's position. When both listeners were off-center, there were substantial distortions in cues. These results confirm difficulties in virtualizing sound sources for listeners outside of the "sweet spot." However, for a listener in the center, a presentation with an additional listener present is acceptable.

Keywords: Auralization, Additional Listener, Spatial Cues, VBAP.

1. INTRODUCTION

In hearing research, new signal processing techniques, new hardware, and updated parameter settings are theorized, created and evaluated, trying to solve communication problems in everyday situations. This facilitates creating sound fields containing realistic, challenging elements for the listener, such as high background noise, high reverberation, and concomitant sound events from different directions. In order to evaluate new technologies committed to solving this level of complexity, new testing methodologies are proposed. Thus, the need for a more realistic laboratory test environment has increased, and so alternatives using binaural (1) methods to present sounds through headphones using the head-related transfer function HRTF have arisen to the hearing research. However, limitations in the individualized acquisition of HRTF and the reproduction via headphones to users of hearing aids and users of cochlear implants presents a new challenge. It is then that the use of techniques for the creation of three-dimensional virtual sound environments (VSE) applied to psychoacoustic and audiological research is presented as an alternative (2–4). Through a VSE it is possible to simulate more realistic sound scenarios, as well as to obtain an auditory measurement with well-controlled and repeatable parameters (5–7). Additionally, this technology allows the researcher to easily switch between different scenes, signal-to-noise ratio (SNR) among other settings, and to enable a participant to use, for example, the hearing aid itself during the test.

Among the main virtualization methods, there is a subdivision based on three paradigms (8): binaural, panorama and sound field synthesis. Headphone playback resides in the binaural paradigm as well as filter-based crosstalk cancellation in loudspeaker reproduction, both aiming to recreate the sound event at ears as they were recorded. Panorama methods aim to recreate the differences in time and level of relative sound pressure between the ears in a sweet spot creating the impression of spatiality based on the sound perception of the listener. Sound field synthesis methods attempt to recreate the recorded or simulated sound field within the playback area.

[1] sergio.aguirre@nottingham.ac.uk

However, even with the possibility of virtualization of sound fields, often the presence or interaction between humans is neglected in auditory evaluations, for the most part, performed by observing only one individual within the laboratory (9–12). The objective of this work is to study the behavior of the ITD and ILD parameters by including a second person inside the ring of loudspeakers. The vector-based amplitude panning technique was used for the virtualization of sound sources. Although the method paradigm is receptor dependent, the method provides an appropriate sense of sound localization for those with normal hearing. The performance of the method regarding hearing impaired people still needs to be adequately studied.

2. METHODS

This section will present the characterization of the test room (See Figure 1) and the methods used in this experiment.

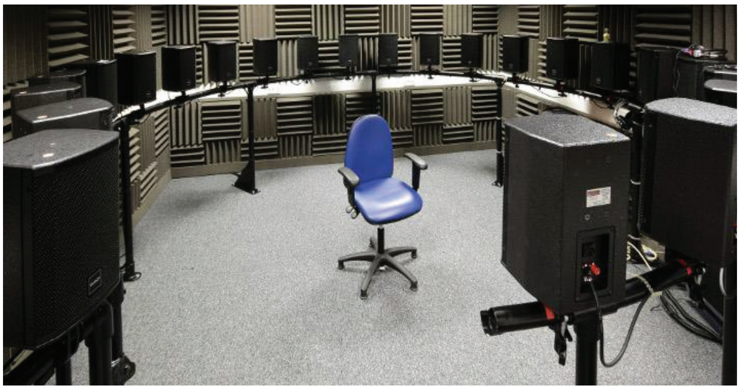


Figure 1: Test Room.

2.1 Room and System Characterization

The experiment was conducted in a large sound-proof audiometric booth (4.3 × 4.7 × 2.9 m; IAC Acoustics). An azimuthal circular array configuration of 24 loudspeakers (3.5-m diameter; 15° of separation; Tannoy VX6) was used. The ceiling and walls were covered with 100-mm deep acoustic foam wedges to reduce reflections; the floor was carpeted with a foam underlay. The AD/DA audio interface that was used was a Ferrofish Model A32. The loudspeakers received signals that were amplified by ART SLA4 amplifiers. The signal acquisition and processing were entirely through ITA-Toolbox (13).

2.1.1 Reverberation Time

The reverberation time (RT60) is one of the most critical objective parameters of a room. The decay of energy to 60 dB below peak (extrapolated from 30 dB below peak) is frequency dependent and provide a subjective perception of the size of the room. For a controlled environment, the values are fractions of seconds. The RT60 of this room in the third octave is presented in Figure 2.

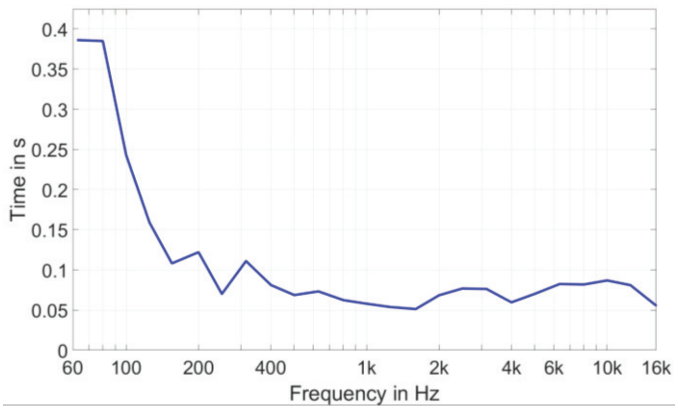


Figure 2: Reverberation Time (RT60) of Test Room as a function of frequency.

2.1.2 Early-Reflections

To ensure that there is no influence of the environment, Recommendation ITU-R 1116-3 determines that the magnitude of the first reflections should be at least 10 dB below the magnitude of the direct sound $\Delta SPL \geq 10$ dB. The differences in the SPL that are determined in the environment of this work meet this requirement.

2.2 Experiment

The experiment studied how the presence of a second person within a loudspeaker ring affects the spatial cues of the reproduced sound field. The data were collected through a B&K TYPE 4128-C Head and Torso Simulator (HATS). The second listener being simultaneously tested was simulated through another mannequin (KEMAR; Knowles Electronics), as shown in Figure 3. The initiative to use the HATS instead of KEMAR to perform data measurement is based on the idea that the objective analysis can use the HATS symmetry as a validation point for the HATS positions.

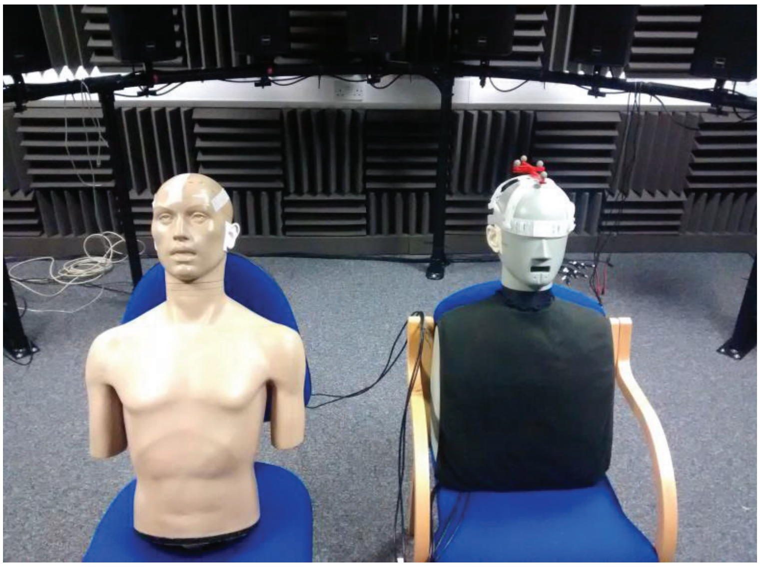
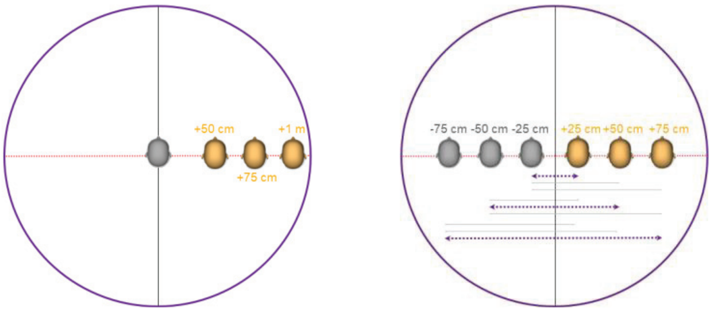


Figure 3: HATS (with motion-tracking crown) and KEMAR inside Test Room.

Using the results for the reverberation time as presented in Section 2.1.1, the appropriate length of a logarithmic sweep signal was calculated as approximately four times larger than the higher value of RT (1.49 seconds). Also, a stop margin of 0.1 seconds was set to ensure the quality of the impulse responses that were obtained. The frequency range covered is from 50 Hz to 20 kHz.

The position of the head has a significant effect on the signals that are measured. To have a reliable assessment of the absolute position of the HATS, its position was measured with a Vicon infra-red tracking system with an accuracy of 0.5 mm. The height position of the microphones in this experiment was 1 meter for all measurements. The first position measured used the HATS in the center, without interference from another obstacle inside the ring, which was taken for future comparison.

A set of positions was proposed to study the influence of a second person inside the ring, keeping the test subject in the center (the sweet spot). Three different positions for the KEMAR (50, 75 and 100 cm of separation) were measured with the HATS fixed at the center of the loudspeaker array. The data collected are from microphones in the HATS ears; the KEMAR was only a physical obstacle to simulate a person inside the ring. Figure 4a illustrates the combinations. Different positions, Figure 4b, maintaining a minimum separation of 50 cm between the center of the heads, were measured. The purpose of these positions with the HATS off-center was to identify the presence of distortions caused by the decentralization of the subject and the effect of the addition of a person within the circle of loudspeakers as a physical obstacle to sound waves. The positioning was standardized so that the movement along the x-axis to the left and right directions of the dummies were annotated as negative and positive, respectively.



(a) (b)

Figure 4: **a**) Measured positions with the HATS centered and the KEMAR present in the room (three combinations) **b**) Measured positions with the HATS in different positions and the KEMAR present in the room (nine combinations).

3. EXPERIMENTAL RESULTS

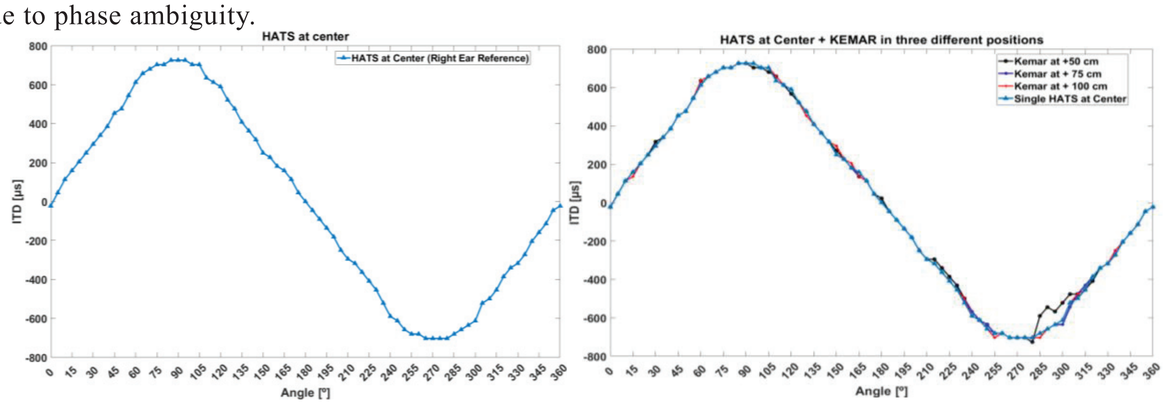
The collected data were analyzed under two objective measures, the interaural time difference (ITD) and interaural level difference (ILD).

3.1 HATS centered

This configuration is intended to collect data from simulating the condition when a person is tested inside the ring accompanied by an actor. Using this type of setting can be useful for analyzing some group influences or disputes between individuals in listening researches.

3.1.1 Centered ITD

The ITD results were obtained after a low-pass filter (LPF) was applied. The cutoff frequency of the filter is 1,500 Hz since the low frequency is predominant in the human hearing concerning the ITD, due to phase ambiguity.



(**a**) (**b**)

Figure 5: ITD as a function of source angle. **a**) HATS alone at center. **b**) Light blue line: HATS alone at center. Black line: HATS centered and KEMAR at 0.5 m to the right. Blue line: HATS centered and KEMAR at 0.75 m to the right. Red line: HATS centered and KEMAR at 1 m to the right

In Figure 5**a**, the data show the results for the ITD from the first setup (HATS alone centered). The system presented a magnitude peak of approximately 650 microseconds corresponding to approximately 0.2 meters for a wave traveling at the velocity of sound propagation in the air. This distance is comparable to the distance between the HATS ear microphones. It is appropriate to note that the symmetry of the HATS is also presented in the results.

The following settings keep the HATS in the center by varying the KEMAR position. The results are presented in Figure 5**b**. The ITD data extracted from this experiment make it possible to note that the second mannequin (KEMAR) has a minor impact as an obstacle on the interaural time difference in the HATS at the center of the loudspeaker ring.

3.1.2 Centered ILD

The interaural level difference was computed to present the second person inclusion effect in higher frequencies. For the following analysis, the effect on ILD can be visualized as a difference from the reference ILDs for HATS alone and centered. The ILD of the setups including the KEMAR were measured and then subtracted from the reference. For a complete match (no difference between ILD measures) all graph should be black. Figure 6 presents the differences between ILD from HATS centered (HC) and the setups with the combination of HATS centered plus KEMAR in each one of the three positions (*e.g.,* HC+K50 to HATS centered and KEMAR at 50 cm to the right).

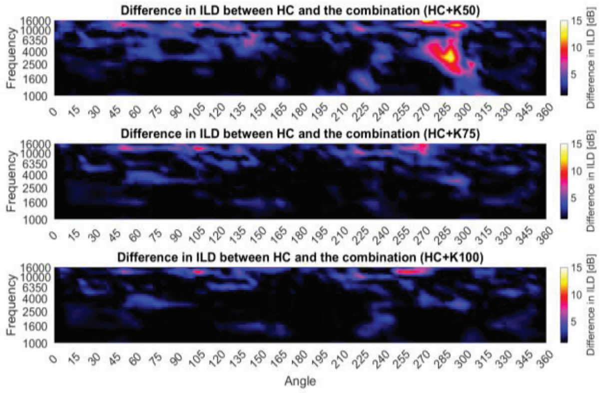


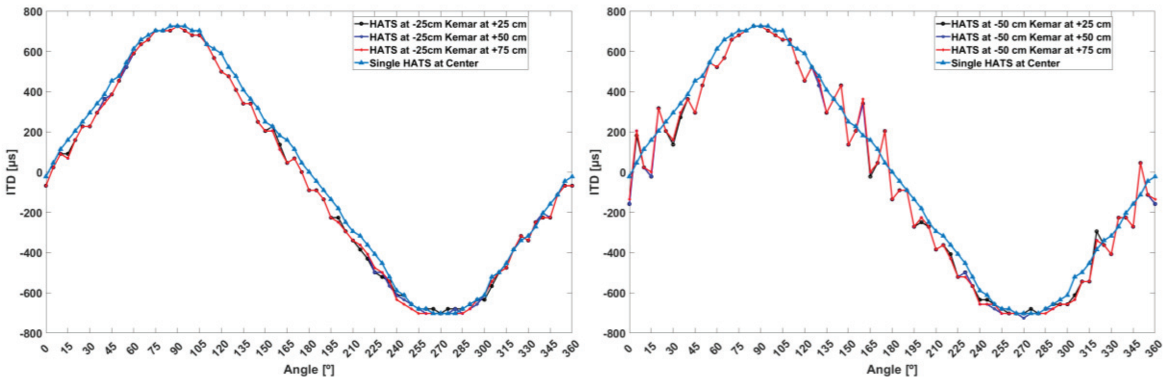
Figure 6: Differences in ILD between HATS at the center and: (Top) HATS at the center plus KEMAR at 50 cm to the right, (Middle) HATS at the center plus KEMAR at 75 cm to the right, (Bottom) HATS at the center plus KEMAR at 100 cm to the right.

3.2 HATS Off-centered

Measurements to study the influence of off-center HATS displacement were performed in nine different configurations: with HATS and KEMAR independently displaced 25, 50, and 75 cm from the center.

3.2.1 Off-center ITD

ITD results are shown in figures 7a 7b and 8a, almost no influence of the second mannequin (KEMAR), even with the HATS off-center can be noted. Nevertheless, a pronounced effect appears by shifting out the HATS off the center, probably due to the vector-based amplitude panning process that was utilized to create the virtual sound sources.

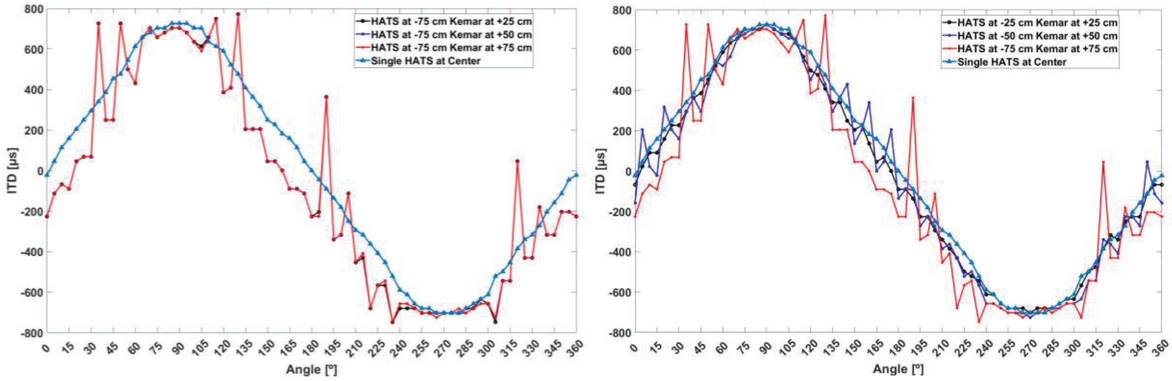


(a) (b)

Figure 7: ITD as a function of source angle **a**) Light blue line: HATS alone at the center. Black line: HATS at -25, KEMAR at +25. Blue line: HATS at -25, KEMAR at +50. Red line: HATS at -50, KEMAR at +75. **b)** Light blue Line: HATS alone at the center. Black Line: HATS at -50, KEMAR at +25. Blue line: HATS at -50, KEMAR at +50. Red line: HATS at -50, KEMAR at +75.

The effect is even more noticeable when the sound is coming from virtual sound sources at angles that are close to the front or rear (0° and 180°) directions. This effect is related to the importance of the sweet spot to the VBAP method since the effect is not present when the HATS is in the center of the ring. Figures 7a, 7b, and 8a show a growth of the difference between the on and off-center in ITDs as the distance increases.

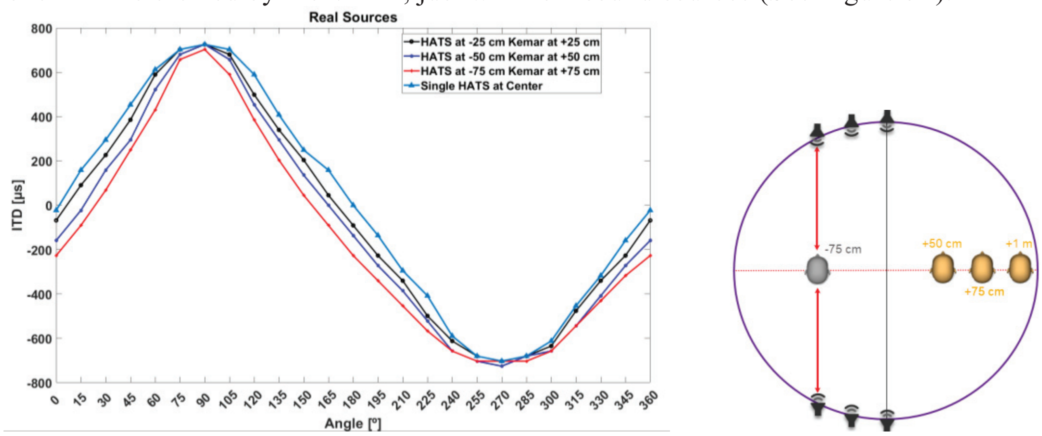
In Figure 10b it is possible to observe more considerable distortions (sharp peaks crossing the reference line) in the ITD for the virtual sound sources. Such distortions increase as HATS is moved away from the central position.



(a) (b)

Figure 8: ITD as a function of source angle. **a)** Light blue line: HATS alone, centered. Black line: HATS at -75, KEMAR at +25. Blue line: HATS at -75, KEMAR at +50. Red line: HATS at -50, KEMAR at +75. **b)** Light blue line: HATS alone, centered. Black line: HATS at -25, KEMAR at +25. Blue line: HATS at -50, KEMAR at +50. Red line: HATS at -75, KEMAR at +75.

The displacement in the off-center position is represented in ITD measurements through a larger lobe to the HATS right ear (270°), and a sharpened lobe at the HATS left ear (90°). This effect occurs because the HATS is not at the center of the ring (See Figure 9b), and the angles and separations between the loudspeakers are modified. The effect is more evident looking at the ITD without the distortions that are created by the VBAP, just with real sound sources (See Figure 9a).

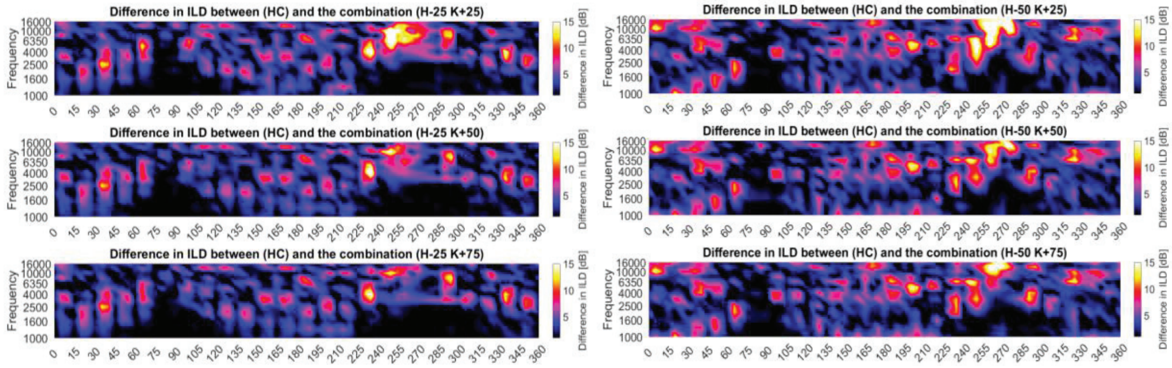


(a) (b)

Figure 9: a) ITD for real (non-virtualized) sound sources as a function of source angle. Light blue line: HATS alone, centered. Black line: HATS at -25, KEMAR at +25. Blue line: HATS at -50, KEMAR at +50. Red line: HATS at -75, KEMAR at +75. b) HATS off-center position -75 cm scheme facing the third loudspeaker

3.2.2 Off-center ILD

Differences in ILD between HATS alone in the center and a configuration with HATS off-center and KEMAR also inside the ring are presented in Figures 15a, 15b. The purpose of this comparison is to study the influence in ILD when the HATS is off-center.



a) b)

Figure 15: Differences in the ILD between centered setup and off-center setups: a) HATS at 25 cm to the left with: KEMAR at 25 cm to the right (top); KEMAR at 50 cm to the right (middle); KEMAR 75 cm to the right (bottom). b) HATS at 50 cm to the left with: KEMAR at 25 cm to the right (top); KEMAR at 50 cm to the right (middle); KEMAR 75 cm to the right (bottom).

Even at best measured off-center setup (HATS at -25 cm and KEMAR at +25 cm), the difference to the ILD results from the centered HATS presents significant differences that can be interpreted as distortions. In the analysis of the acoustic field behavior outside the center of the ring at frequencies above 1 kHz, significant differences are found in all measured configurations.

It is possible to notice a decrease in ILD differences generated by acoustic shadowing caused by KEMAR in the HATS 25 cm and 50 cm off-center position. While these decreases in ILD to positions near 270 degrees (right side of HATS), the differences in other virtualized positions are permanently pronounced.

4. DISCUSSION

When one listener is centered in the loudspeaker array, the ITD was only affected when the other listener is 50 cm away, and only for presentation angles shadowed by the second listener. The ITD for sounds coming from all other angles were barely affected by positioning a second mannequin inside the ring. The same results apply to ILDs for a listener in center position.

In the analysis of off-center listener positions, the displacement effect is apparent in the ITD. The peak of magnitude remains practically the same as the ITD 0 value (sound reaching at the same time in both ears) is shifted. The relative time of sound arrival to positions to the right of the mannequin (between 0 and 180 degrees) increases, while between 180 and 360 (or zero) decreases relative to the centralized position. This effect is expected given that the mannequin is physically in front of another speaker.

For ITDs, no significant difference was observed between the KEMAR positions (25, 50 and 75 cm to the right of the center) in all measurements with HATS to the left of the center. When the off-center HATS position was 25 cm to the left, the ITD had a good approximation of the reference-centered measurement. Subsequent HATS positions of 50- and 75-centimeters present peaks and crossover values that indicate distortion problems at low frequencies.

The decentralized analysis of the interaural level difference parameter was made by subtracting the ILD result from each position combination (both dummies within the decentralized ring) of the ILD of one, centered listener. The shadow effect generated by the presence of a second listener reduces as expected when the first listener is -25 cm or -50 cm off center. However, the high frequencies are mainly affected for only virtual sources, which indicates the difficulty of reproducing the same level of sound pressure in the ears outside the center, independent of the position of the second listener.

It should be noted that the current study did not measure changes in ITD and ILD for off-center listener positions without the presence of a second listener. Based on the effects of having the first listener off-center with a second listener present, coupled with the smaller changes (for actual sound sources) with a second listener when the first listener is centered, it can be deduced from the current results that the off-center position has a greater effect on the ITD and ILD. Considering that many sound-field virtualizations are limited by a "sweet spot" for the listener(s), the off-center position, as opposed to the presence of a second listener, is probably the greatest liability for multi-listener methods in hearing research.

5. CONCLUSION

In the central position, for one test participant, the VBAP technique does not affect the ILD and ITD acoustic cues and the addition of a second person within the ring also does not significantly affect these parameters at the three distances tested, except for the angles usually hidden by the shadow second person. Thus, it is possible to move towards subjective tests with a center participant and an actor on the side.

There is a clear degradation when two test subjects are simultaneously present in all off-center positions, regardless of the distance of a second person, the measurements showed significant differences in ILD. These differences indicate the creation of acoustic artifacts, possibly generated by the method's difficulty in correctly virtualizing high frequencies outside the sweet spot. For the ITD parameter, the displaced position 25 of the center has little difference or evidence of artifacts generated by virtualization errors, while the other distances present significant differences and artifacts.

ACKNOWLEDGMENTS

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under the Marie-Sklodowska-Curie grant agreement Nº 765329. This work was also supported by the Medical Research Council [grant number MR/S003576/1]; and the Chief Scientist Office of the Scottish Government.

REFERENCES

1. Paul S. Binaural Recording Technology: A Historical Review and Possible Future Developments. Acta Acust United Acust. 2009 Sep 1;95(5):767–88.
2. Vorländer M. Virtual Acoustics: Opportunities and limits of spatial sound reproduction for audiology. Deutsche Gesellschaft für Medizinische Physik e.V.39 DGMP-Oldenburg. 2008;
3. Favrot SE, Buchholz J, Dau T. A loudspeaker-based room auralization system for auditory research [phdthesis]. Technical University of Denmark; 2010.
4. Cubick J, Dau T. Validation of a virtual sound environment system for testing hearing aids. Acta Acust United Acust. 2016;
5. Llorach G, Grimm G, Hendrikse MME, Hohmann V. Towards Realistic Immersive Audiovisual Simulations for Hearing Research. In 2018.
6. Seeber BU, Kerber S, Hafter ER. A system to simulate and reproduce audiovisual environments for spatial hearing research. Hear Res. 2010;260(1):1–10.
7. Grimm G, Kollmeier B, Hohmann V. Spatial Acoustic Scenarios in Multichannel Loudspeaker Systems for Hearing Aid Evaluation. J Am Acad Audiol. 2016; 27(7):557-566.
8. Masiero B, Vorlaender M. Spatial Audio Reproduction Methods for Virtual Reality. In Cáceres; 2011.
9. Gandemer L, Parseihian G, Bourdin C, Kronland-Martinet R. Perception of Surrounding Sound Source Trajectories in the Horizontal Plane: A Comparison of VBAP and Basic-Decoded HOA. Acta Acust United Acust. 2018;104:338-350
10. Lundbeck M, Grimm G, Hohmann V, Laugesen S, Neher T. Sensitivity to Angular and Radial Source Movements as a Function of Acoustic Complexity in Normal and Impaired Hearing. Trends Hear. 2017;21:2331-2165
11. Marentakis G, Zotter F, Frank M. Vector-based and ambisonic amplitude panning: A comparison using pop, classical, and contemporary spatial music. Acta Acust United Acust. 2014;100:945-956
12. Guastavino C, Katz BFG. Perceptual evaluation of multi-dimensional spatial audio reproduction. J Acoust Soc Am. 2004;116:1105-1115
13. Berzborn M, Bomhardt R, Klein J, Richter J-G, Vorländer M. The ITA-Toolbox: An Open Source MATLAB Toolbox for Acoustic Measurements and Signal Processing. In 43th Annual German Congress on Acoustics, Kiel (Germany), 6-9 Mar; 2017;43:.