# Distribution of Quadrant Errors in Auditory Localization using a Binaural Headphone system

Stephan Werner[1], Martin Rekitt[1], and Florian Klein[1]

[1] *Institute for Media Technology, Ilmenau University of Technology, Ilmenau, Germany*
*Email: stephan.werner@tu-ilmenau.de*

## Introduction

The auditory system of humans enables the perception of spatial audio in real and virtual acoustics using monaural and binaural cues. The perception of direction and distance are two prominent quality features to evaluate the quality of experience of spatial audio systems. Inaccuracies in perception can occur if physical quality elements of the synthesis system are not adequate. In this study a binaural synthesis via headphones is used to re-synthesize single sound sources on several discrete positions on a full circle around the listener. An artificial head (KEMAR) and a two channel spherical microphone setup are used to measure binaural room impulse responses in a real environment. A listening test is performed to measure the number of quadrant errors and perceived externalization of the auditory events in a localization task [1]. The distributions of its frequencies depending on the direction of the re-synthesized sound source are investigated. The results show a continuous relationship between the investigated quality features and the direction of the sound source. Furthermore, an analysis of the relationship between quadrant errors and externalization depending on direction and used binaural room impulse responses are presented.

## Binaural Synthesis

For the test stimuli, binaural recordings of binaural room impulse responses (BRIRs) for the selected room, sound sources and positions have been done. A head and torso simulator (KEMAR) and a spherical microphone array with a diameter of 17.5 cm are used for BRIR recording (see figure 1). The headphones are equalized using headphone transfer functions (HPTFs). HPTFs from the head-and-torso simulator and spherical microphone array are used depending on the used BRIRs. The inverse of a HPTF is calculated by a least-square method with minimum phase inversion [2]. Stax Lambda Pro headphones which fulfill the requirements for open headphones are used for playback [3].

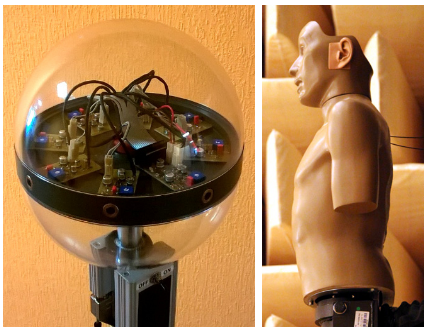

**Figure 1** – Recording devices used in the experiment; left: spherical microphone array (Kugelflächenmikrofon), right: KEMAR head and torso simulator.

## Listening Test

A listening lab (compliant to Rec. ITU-R BS.1116-1, V=179 m³) is used for the listening test and the measurement of the BRIRs at a distance of 2.5 m. Twenty five directions on a circle in the horizontal plane are chosen to measure the BRIRs (see figure 2). The directions are arranged asymmetric for the left and right side and with a higher density at frontal and rear directions. Dry audio signals are convolved with recorded BRIRs from the different azimuth directions. A saxophone part without reverb, a low-pass noise (80 Hz – 600 Hz), and a high-pass noise (6 kHz – 10 kHz) are used as audio signal. Nineteen test persons with an age between 20 and 50 years participate in the listening test. The test persons are trained in the beginning of the test to become familiar with the rating interface and to build up the percepts externalization and localization. The listening room and the loudspeaker positions (dummy loudspeakers) are visible during the listening test.

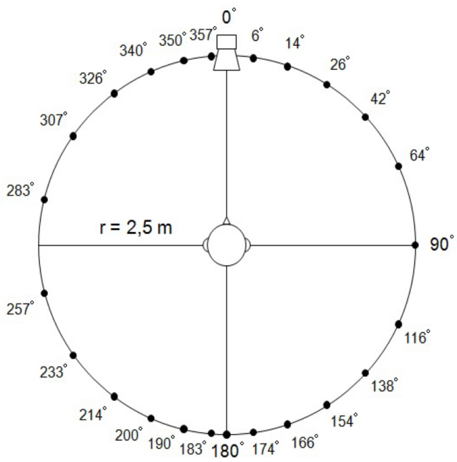


**Figure 2** – Positions of the loudspeakers.

In the experiment, each stimulus is rated on a rating sheet as shown in figure 3. The test design is a single stimulus test for the quality features externalization and localization. Rating for externalization of the auditory event is realized by indicating the inner, middle, or outer circle on the rating sheet. The following definitions are used in the test: a) inner cloud: “The auditory event is entirely in my head and very diffuse.” b) inner circle: “The auditory event is entirely in my head and good locatable.”; c) middle circle: “The auditory event is external but it is next to my ears or head.”; d) outer circle: “The auditory event is external and good locatable.”; e) outer cloud: “The auditory event is external and very diffuse”. (Note that the definitions are given in the test person’s native language, German). The perceived direction of the auditory event can be rated by choosing the frontal or rear direction. If the auditory event is localized close to +/-90° the test person is still instructed to choose a direction.

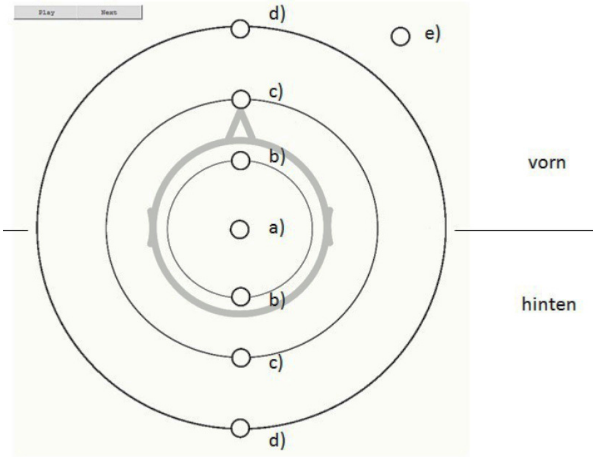


**Figure 3:** Rating sheet for the experiment; a)-c) are related to in-head or near-head localization, d) and e) related to externalization.

## Ratings

An externalization index is calculated as ratio between the number of ratings for "extern" (circle d) and e) on the rating sheet) and all number of ratings. An index of "1" would indicate an external perception of the test condition while an index of "0" indicates total in-head localization. The quadrant error is calculated as ratio between the number of commutated directions and the total number of ratings.

Figure 4 shows the ratings for externalization and localization as example for the saxophone signal and KEMAR BRIRs.

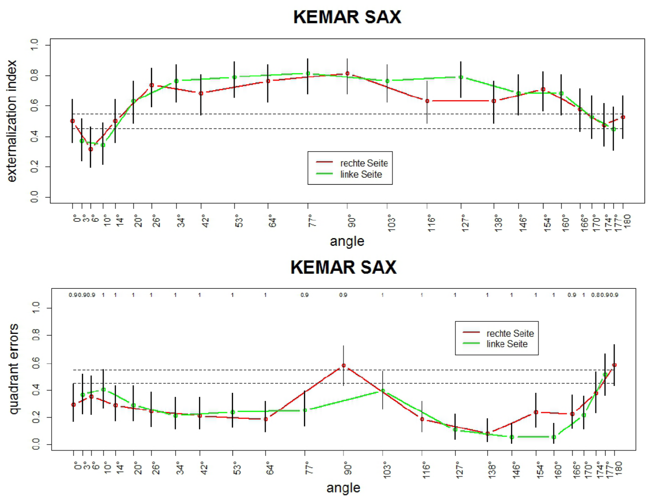


**Figure 4:** Example for the ratings of externalization and localization; top: externalization indices, bottom: quadrant errors; with 95% binominal conf. interval; dotted lines = range of guessing.

The indices for all directions and signals are taken for a combined representation of the externalization indices and quadrant errors shown in figure 5 for the KEMAR and spherical microphone array (KFM). Next to a linear regression line a correlation after Spearman is calculated to show the relationship between externalization and quadrant error.

The ratings for the binaural synthesis using the dummy head show a significant (p<.05, Spearman´s rho) negative correlation between the amount of quadrant errors and the externalization index. This confirms results from own former studies [4]. In contrast, the ratings for the binaural synthesis using the spherical microphone array show no correlation between localization and externalization. The distribution of the occurrence of quadrant errors follows a Gaussian distribution with its mean very close to 0.5. This indicates a guessing of the test persons for the perceived directions. Furthermore, it is noticeable that the rated externalization is comparable to the KEMAR ratings.

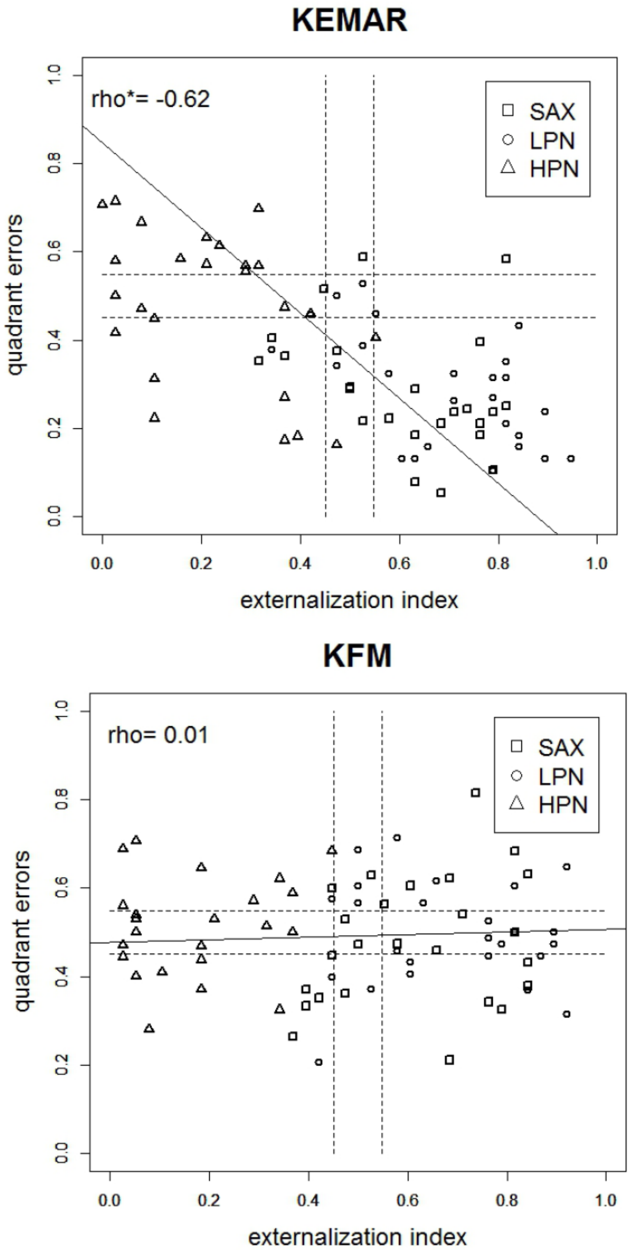


**Figure 5:** Quadrant errors and externalization indices for all directions and signals; spearman´s rho; * with p<.05; solid line=linear regression line; dotted lines = range of guessing; SAX=saxophone, LPN=low-pass noise, HPN=high-pass noise, KFM=Kugelflächenmikrofon.

## Conclusions

The ratings of the test persons show less localization inaccuracies for the binaural synthesis using the dummy head (KEMAR) compared to using BRIRs from the spherical microphone array. The distribution of the occurrence of quadrant errors suggest the assumption that the test persons are guessing for the synthesis using the spherical microphone array. A negative correlation is found between the occurrence of quadrant errors and the externalization indices for the binaural synthesis using the dummy head (KEMAR). But, no correlation is found for using the spherical microphone array.

Georg Plenge formulated in his paper [5] the statement: "[in-head localization occurs if] … wrong knowledge about the sound field is present and/or … the stimuli are mannered in that way that the stimuli cannot be matched with a stimulus pattern." (original paper is in German). The auditory cues given by the dummy head are matching not very well with the individual patterns of each test person [6], [7]. Furthermore, a wrong knowledge about the sound field occurs which results in localization errors like the quadrant error. Following Plenge, an increase of localization errors should yield to an increase of in-head localization.

However, this statement does not describe the ratings for the spherical microphone in a satisfying way. The localization ratings of the test persons show a guessing of the test persons for localization. In contrast, the perceived externalization is comparable with the dummy head synthesis. Welch and Warren formulated in his paper [8] a hypothesis of modality appropriateness. The perceptional system choses this modality which has the most reliable information or cues about the perceptional event. We conjecture that the absence of auditory cues, caused by the missing spectral cues at the spherical microphone array, and the visibility of the listening room and dummy loudspeakers yield to no correlation between the quality features localization and externalization. Furthermore, the room acoustic convergence between the synthesized scene and the listening room also increases the perceived externalization [4].

## References

[1] Rekitt, M.: "Bestimmung der Häufigkeitsverteilung von Quadrantenfehlern bei der Lokalisation von Hörereignissen unter Verwendung einer binauralen Kopfhörerwiedergabe", Bachelor Thesis, TU Ilmenau, 2014.

[2] Schärer, Z. and Lindau, A.: "Evaluation of equalisation methods for binaural signals," 126th AES Conv., preprint 7721, 2009.

[3] Møller, H.: "Fundamentals of Binaural Technology", Applied Acoustics, 36, pp. 171-218, 1992.

[4] Werner, S. and Klein, F. "Influence of Context Dependent Quality Parameters on the Perception of Externalization and Direction of an Auditory Event", 55th AES Conference Spatial Audio, Helsinki, Finland, 2014.

[5] Plenge, G. "Über das Problem der Im-Lopf-Lokalisation", Acustica, 26(5), pp. 241-252, 1972.

[6] Møller, H., Sørensen, M. F., Jensen, C. B., and Hammershøi, D.: "Binaural technique: Do we need individual recordings?", J. Audio Eng. Soc, 44, pp. 451-469, 1996.

[7] Begault, D. R. and Wenzel, E. M.: "Direct comparison of the impact of head tracking, reverberation, and individualized head-related transfer functions on the spatial perception of a virtual speech source", J. Audio Eng. Soc., 49, pp. 904-916, 2001.

[8] Welch, R.B. and Warren, D.H. "Immediate perceptual response to intersenory discrepancy", Psychological Bulletin, Nr. 88, pp. 638-667, 1980.