Assessment of Sound-field parameter differences in studio listening conditions

G. Spikofski
IRT

Measurements and psychoacoustic investigations have been carried out in ten studio listening environments in European broadcasting organizations. The measurements were based on EBU Doc. Tech 3276 while the psychoacoustic tests related to previous investigations carried out by the author. The material used for the psychoacoustic assessment was dummy head recordings made under specific studio listening conditions. Correlations between the relevant physical and psychoacoustic parameters are given.

1. Introduction

It can be generally assumed that, in professional audio environments, the "neutrality" of the listening conditions is an indispensable requirement. However, the definition of "neutrality" is a complex issue as the sound field in listening rooms is influenced by a variety of parameters. In a listening room – in contrast to headphone listening, which is determined solely by the properties of the transducer used – the loudspeakers, the room and local relationships between the loudspeakers and listeners interact in a complex fashion. This is true of monophonic, stereophonic and multichannel listening. In other words, if one considers a conventional music transmission from a single sound source in a concert hall, the listening experience from the reproduced sound will be characterized by an overlapping of the "primary room information" (i.e. from the concert hall) and the "secondary room information" (from the listening room). Seen in this way, "neutrality" can be defined approximately as follows: the listening room enables the sound producer to work in a familiar "room" environment, rather than in an unfamiliar "anechoic" environment. The influence of the room may be just perceptible enough for the "primary room information" not to be drowned out.

"Neutral listening conditions" are a necessary instrument for aesthetic quality assessment in music production. In broadcasting, this is of particular importance in the field of international programme exchanges. With the introduction of new transmission and storage systems based on aurally-adapted data compression, neutral listening conditions are a necessity from another angle: only in neutral listening conditions is an objective assessment of the quality of the new generation of audio systems possible. Technically, the definition of "neutral listening conditions" involves setting minimum requirements for the relevant electroacoustic and acoustic parameters. Standards bodies have been working on the optimization of these parameters for decades.

The three relevant international recommendations for the assessment of studio listening conditions are ITU-R Recommendation BS.1116: [1], EBU doc. Tech 3276 [2] and EBU doc. Tech 3276 (Supplement 1) [3]. The ITU and EBU standards differ in structure and content. However, with regard to the relevant electroacoustic and acoustic parameters, the differences are insignificant and so, in this respect, it is possible to assume identical requirements.

The present study refers only to EBU doc. Tech 3276 [2], which was drawn up by the EBU P/LIST Project Group. This document has acquired current importance through a supplementary internal EBU report, BPN 014, which is available on the EBU's website (to Members only). BPN 014 contains the measurement results from ten reference listening environments in European broadcasting organizations, all of which meet the minimum requirements in respect of the relevant electroacoustic and acoustic parameters described in [2]. The measurement results described in this article allow a differentiated technical assessment of the listening conditions investigated. The inclusion of dummy-head recordings from the investigated listening conditions means that the results can also be correlated with the subjective listening impression.

Within the framework of IRT work on audio systems engineering, differential assessment of the listening conditions is of particular importance to the current project [4][5][6][7][8]. Binaural Room Scanning (BRS) is based on scanning and virtual presentation of real listening rooms with the aid of a dummy head or corresponding binaural HRTF (head-related transfer function). In the BRS processor, the signals to be synthesized (e.g. the left loudspeaker) are generated by convolving the source signal with the corresponding HRTF. The processor output signals which have been calculated for headphones are the same as the dummy head signals that, for example, are received at the reference listening position.

This means that, by reference to a suitably-defined interface dummy head / head-phone transfer function, the dummy head – like the virtual dummy head, the BRS processor – delivers the same sound perception as in the original sound field. However, with this static scanning of the listening room, the typical dummy head front-rear inversions occur. To eliminate this disturbant perception feature, we have incorporated head movement. The head movement is carried out by means of a head-tracking system which transfers the listener's head movements to the BRS processor. For dynamic classification of the HRTF which corresponds to the head movements, the data-store of the BRS processor contains – in addition to the corresponding standard directions – the scanned-in HRTF of the front horizontal plane over the range of ± 42° with a definition of 6°. Through dynamic allocation of the HRTF as a function of the listener's head movements, front/rear inversions can be effectively avoided [8]. With the aid of the BRS processor, it is thus possible to virtually "clone" the real listening conditions. This means that quality assessment of the chosen real listening conditions assumes particular importance with respect to the target definition of the BRS process.

An unusual complement to this study is that interested readers have the opportunity to listen on the Internet to the dummy head recordings that formed the basis for the psycho-acoustic assessment described here, thus allowing comparative assessment of the results presented. Dummy head recordings need to be listened to on headphones (not loud-speakers). The virtual impression thus arises that one is sitting in the optimum listening position where the dummy head was situated during the recording. These dummy head signals can be called up via http://www.irt.de/IRT/indexpubli.htm. The corresponding Internet page (click on "*Zur Bewertung von Schallfeldparameter-Unterschieden in Studio-Abhörsituationen*" to find the English version) provides further information on listening to the dummy head recordings and the relevant links to the audio players used. The listening conditions are not given in the same order as in the EBU report.

2. The problem and an approach to a solution

Measurement data were available from the listening rooms listed in the table below:

Listening room	Organization	Location
Acoustic test room	BBC	Kingswood Warren
Listening room 1	BBC	Kingswood Warren
Studio	IRT	Munich
Listening room	Secondary School of Musical Art	Budapest
Listening room	Hungarian Radio	Budapest
Reference listening room	YLE	Helsinki
Reference listening room	NRK	Oslo
Listening room 1	Deutsche Telekom/TZ-D	Berlin Adlershof
Non-reflecting room	IRT	Munich

The listening rooms/conditions considered in BPN 014 represent the state of the art with respect to the reference listening conditions. Although some of the listening rooms investigated went into use some years ago – bearing in mind the standards required at that time – the majority are new constructions. The demand for the creation of reference listening conditions meant that most of the listening rooms contain no disturbant elements (e.g. mixing consoles).

BPN 014 contains the following measurement results, corresponding to EBU doc. Tech 3276 [2], for each listening room:

- room dimensions;
- measurement situation;
- early reflections;
- reverberation time;
- loudspeaker/room response curves;
- background noise.

The question the present study sets out to answer is as follows: *do the differences in measurement values arising from the tolerance limits have a bearing on subjective assessment?* On this subject the following approach to a solution is adopted. With the aid of the available dummy head recordings, the investigated listening conditions are also assessed subjectively. The subjective assessment is based on our own investigation at the IRT [9]. The result of this investigation shows that for the assessment of studio listening conditions, the following aspects are of relevance:

- image quality;
- tonal balance / bass and treble reproduction;
- reverberation component;
- disturbant noise.

The disturbant noise aspect, which in the IRT investigation [9] was regarded as irrelevant, has been reintroduced in the present study. In the light of these assessment aspects, a psychoacoustic quality profile has been drawn up for each listening room, which provides the basis for computing the correlations between the subjective and objective parameters.

3. Objective measurements – definitions according to EBU doc. Tech 3276

3.1. Listening conditions

The definitions according to EBU doc. Tech 3276 are given in the table below. The measurement results are presented and discussed in *Section 4* together with the results of the subjective assessment.

	Room volume V (m^3)	Reverb. time T_m (s)	Stereo basis B (m)	Loudspeaker
AHS_1	190 (RAR = reference)	0.00	2.7	B&W 801
AHS_2	267	0.39	3.75	Genelec 1038
AHS_3	175.5	0.23	3.5	Genelec 1037
AHS_4	175	0.31	4.5	Genelec 1025C
AHS_5	72.5	0.33	2.05	BBC LS5/8
AHS_6	106.65	0.35	2.55	BBC LS5/9
AHS_7	177	0.26	2.9	K&H O92
AHS_8	177	0.26	2.9	K&H O108 + Genelec 1094A (Sub woofer)
AHS_9	177	0.26	2.9	Spendor 120-A
AHS_10	255.61	0.27	3.5	B&W 801

3.2. Direct sound

Direct sound is defined as the sound field which is measured using the same loudspeakers, but under anechoic conditions, i.e. without the early reflections and the reverberation caused by the listening room. The tolerance range of the frequency-response curve, between 40 Hz and 16 kHz, must not exceed 4 dB (measured in 1/3-octave sub-bands). It has been assumed that all the loudspeakers used in the tests meet this requirement.

3.3. Early reflections

Early reflections are defined as reflections from boundary surfaces in the room which reach the listening area within the first 15 ms after the arrival of the direct sound. The levels of these reflections should be at least 10 dB below the level of the direct sound for all frequencies in the range 1 kHz to 8 kHz. Since information on amplitude, time and frequency is important for the analysis of early reflections, the measurement data are given in the form of ETF or "waterfall" diagrams. Apart from strong broadband early reflections, which did not occur in the listening conditions investigated, the assessment of individual reflection models is usually difficult. For this reason, these aspects are not discussed in the psychoacoustic interpretation of measurement data.

3.4. Reverberation time

Fig. 1 shows the range of *reverberation times* encountered in the measured listening conditions, and individual results with the relevant tolerance limits. Measurement was carried out in 1/3-octave bands. The nominal value T_m of the measured reverberation times is the average of the 1/3-octave sub-bands between 200 Hz and 4 kHz.

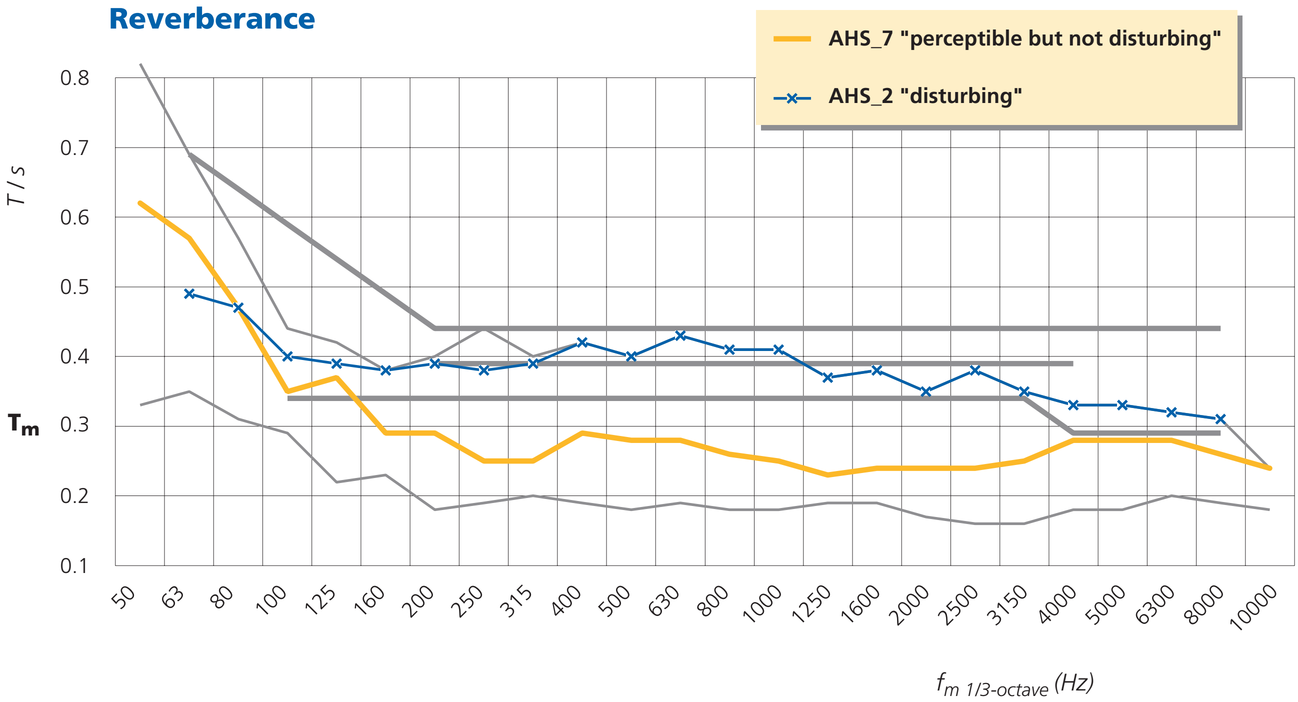


Figure 1
Reverberation – reverberance.

3.5. Loudspeaker / room response curves

Fig. 2 shows the range and individual results for these response curves, along with the relevant tolerance limits. Measurement was carried out using noise in 1/3-octave sub-band, or equivalent measuring methods. The tolerance range relates to every listening position in the listening area. The measurement data refer limitatively to the reference listening position. The reference level L_m is calculated from the mean levels of the 1/3-octave sub-bands, with centre frequencies ranging from 200 Hz to 4 kHz.

3.6. Background noise

The background noise in the listening room was calculated in 1/3-octave sub-bands between 50 Hz and 10 kHz. The tolerance limits shown are noise rating curves NR 10 and NR 15; high-quality studio listening conditions may not exceed NR 10 and other studio listening conditions, NR 15. *Fig. 3* likewise shows the range and individual results for these curves, with the corresponding tolerance limits.

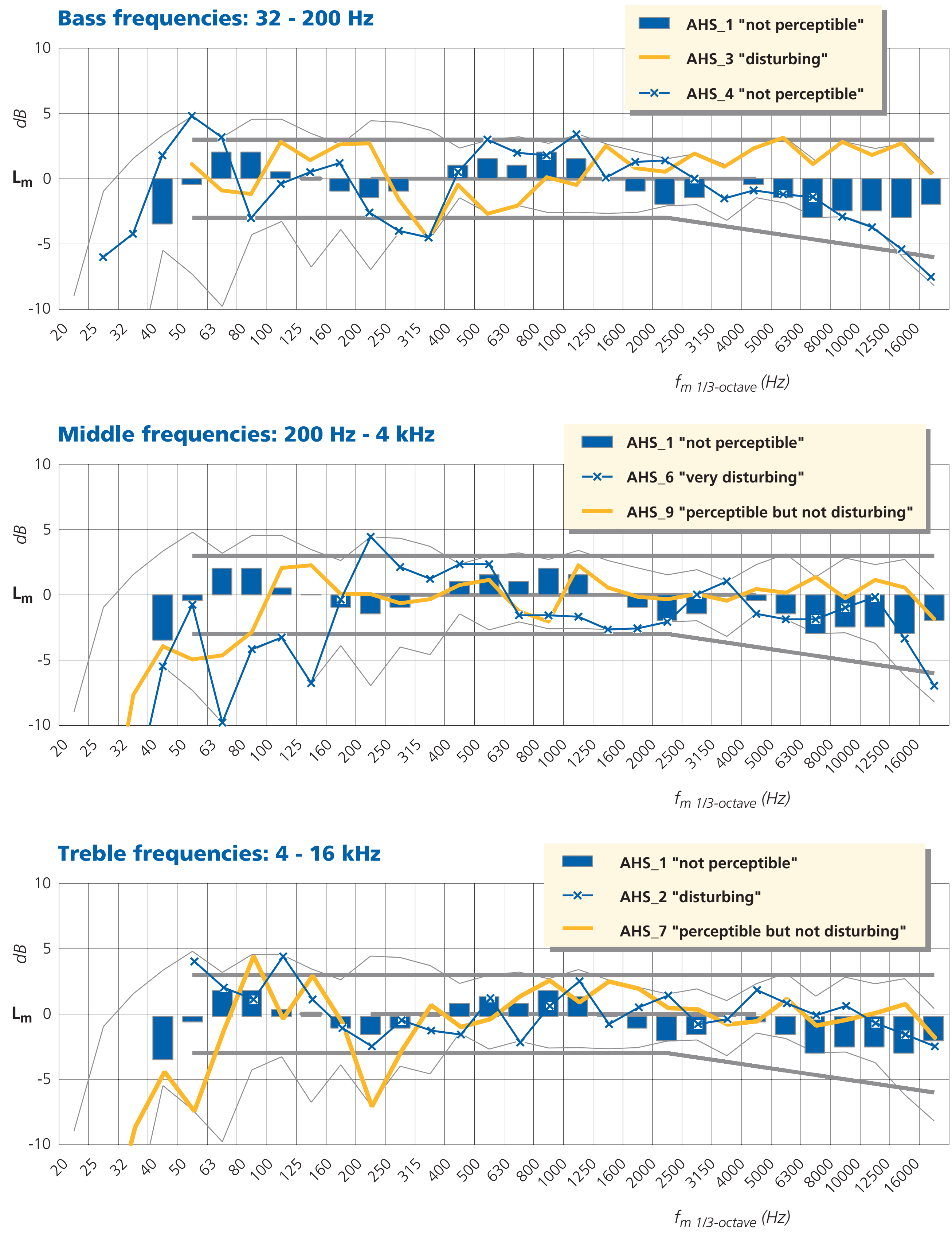


Figure 2
Loudspeaker / room response curves – coloration at bass, middle and treble frequencies.

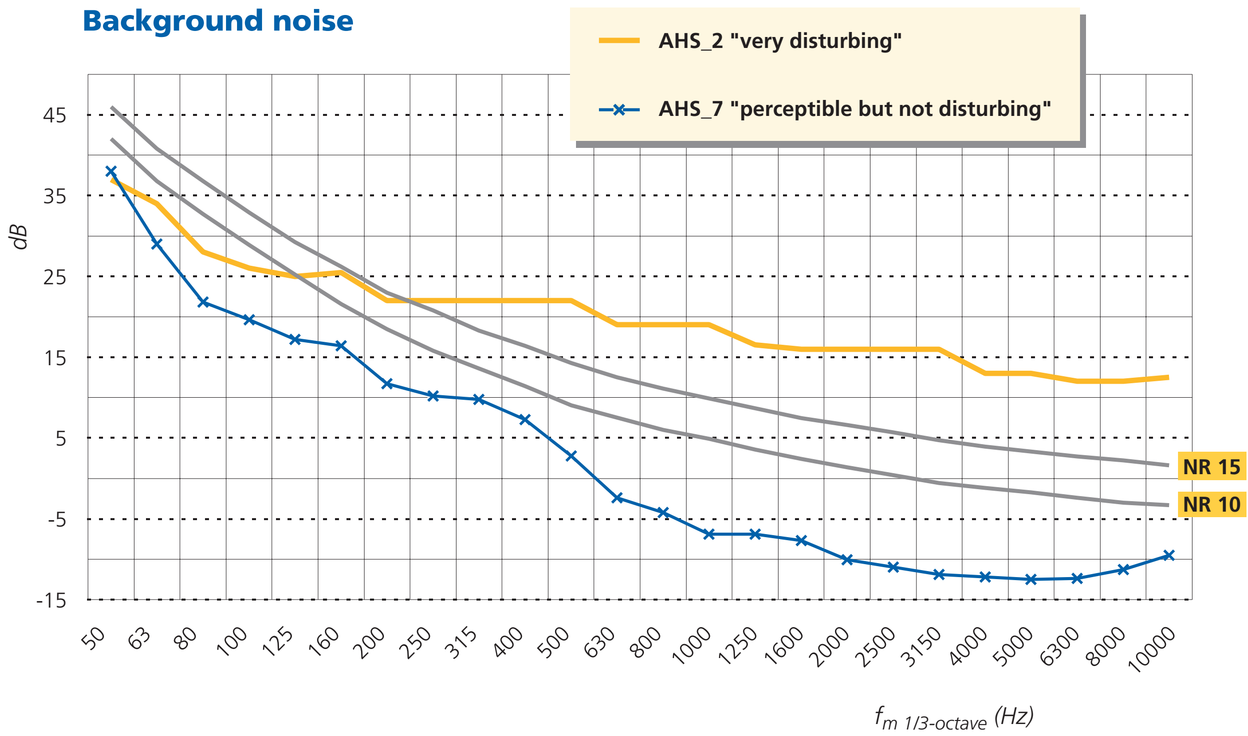


Figure 3
Background noise – disturbant noise.

4. Subjective measurements – definitions

In the psychoacoustic assessment, we refer to our own study at the IRT on the definition of psychoacoustic parameters in various studio listening conditions [9]. The investigation related to seven listening rooms using three studio loudspeaker types, and four listening rooms including an anechoic chamber. The seven listening rooms were assessed with the aid of 24-dimensional quality profiles. The results of the factor analysis which was performed can be summarized in the form of the following four psychoacoustic factors: *coloration*, *tonal balance*, *imaging quality* and *reverberance*.

4.1. Coloration

The typical attributes of this factor relate to sound colorations ("coloured", "uncoloured") due to linear (frequency-dependent) distortions. This aspect is of particular importance for the quality assessment of listening conditions, since the general quality attributes "natural" and "well-balanced" correlate closely with coloration.

4.2. Tonal balance

The typical attributes of tonal balance relate to the reproduction of bass and treble frequencies. Apart from the reproduction of bass and treble sounds ("emphasized" – "de-emphasized"), another factor to consider is the relationship between bass and treble sounds. This becomes clear if one considers the correlations between the attributes "treble emphasis – lacking bass" and "bass emphasis – muffled" (in the sense of "lacking treble").

4.3. Imaging quality

The attributes relevant to interpretation of this factor relate to the room and the (phantom) sound sources which are simulated with the aid of stereophony in the loudspeaker base area. The attributes "stage width" and "stage depth" describe the imaging of the simulated sound source, e.g. an orchestra, relative to the dimensions "spatial breadth" and "depth". Good spatial depth of the sound source is described with attributes such as "exactly localizable", "transparent", "detailed" and "spacious", which are contrary to the attribute "blurred".

4.4. Reverberance

It can be assumed that reverberance, as a relevant psychoacoustic factor, correlates with the sound-field parameter "reverberation" in the listening room.

4.5. Assessment parameters – test procedure

On the basis of these results, the following assessment parameters have been defined for the present study:

- ⇨ coloration of middle frequencies (coloured – uncoloured);
- ⇨ coloration of bass frequencies (coloured – uncoloured);
- ⇨ coloration of treble frequencies (coloured – uncoloured);
- ⇨ imaging quality (low – high imaging quality);
- ⇨ reverberance (reverberant – dry);
- ⇨ background noise (soft – loud).

The background noise aspect included here relates to all noise which is not part of the recording, e.g. the drone of an air-conditioning system, hum caused by the electric mains supply, etc., which are perceptible in the pauses between signals or are not part of the music signals.

Assessment was carried out with the aid of the dummy head recordings from the listening conditions investigated. Reproduction of a dummy head recording via headphones delivers the same sound experience as in the original sound field – provided the requirements of the reference interface dummy head / headphones transfer function are met [10][11][12]. In the present case (interface: non-frequency-dependent diffuse-field transfer function), this was verified physically for the dummy head/headphone combination used [dummy head: KU100 (Neumann); headphone: SR Lambda/SRM Monitor with diffuse-field equalizer (Stax)].

The programme material selected for the psychoacoustic measurements consisted of the following five programme examples of 10 - 20 seconds' duration:

Example	Recording system
Radio drama	Neumann SM 69 / Philips MCR unit
Jazz	Vienna Art Orchestra (CD)
Baritone & piano	Two omnis 35 cm / Brüel & Kjaer 4003
Pop	Tri Cycle (CD)
Mozart symphony	Two omnis 60 cm / Schoeps MK 2 S

The programme examples were part of a collection already used successfully in comparable investigations [9][13][14]. The most important selection criterion was the clear differentiation between the listening conditions investigated.

Comparative assessment was carried out in the form A-B-A-B, with A being the reference and B the listening condition to be assessed. This meant that, for each example and each assessment aspect, room B was to be assessed comparatively with room A. In making the assessment, the following 4-rank scale was adopted.

- ⇨ **4** = not perceptible
- ⇨ **3** = perceptible but not disturbing
- ⇨ **2** = disturbing
- ⇨ **1** = very disturbing

The reference used was listening condition AHS_1 (RAR/B&W 801) since, in the investigation mentioned previously [9], it was assessed as neutral with respect to the above-defined psychoacoustic parameters. Assessment was carried out by a panel of five experts. The results given represent the median value of the assessor group and the programme examples. The purpose of psychoacoustic assessment is to draw up an appropriate psychoacoustic quality profile for each listening condition. This psychoacoustic subjective quality profile is placed in relation to the objective quality profile represented by the measurement results given in EBU report BPN 014.

5. Results

The results of the subjective assessment, the psychoacoustic quality profiles, are presented in the form of net diagrams. The chosen form of presentation provides clarity in assessing the quality of the investigated listening conditions in relation to the psychoacoustic parameters examined. In the chosen form of presentation, the size of the enclosed surface is equivalent to the degree of neutrality.

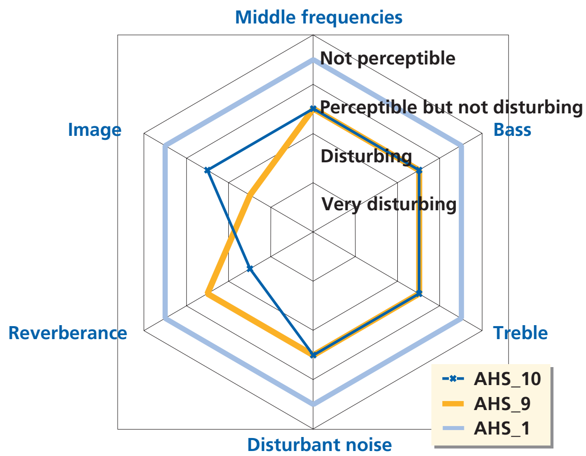
On the basis of this criterion, a simple classification is possible. For interpretation of the results and comparison of subjective and objective measurements, two listening rooms from different quality classes were selected and discussed in each case. In the light of the psychoacoustic quality profiles, three quality classes were defined:

- **2** = “rather good”;
- **3** = “rather satisfactory”;
- **4** = “rather adequate”.

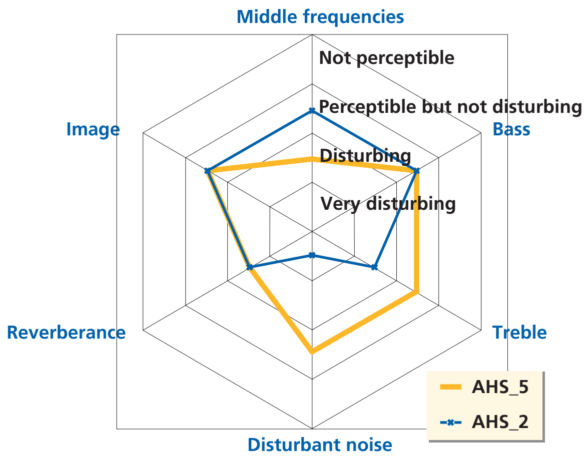
With class “2”, the listening conditions AHS_10 and AHS_9 were assessed (*Fig. 4a*). It is striking that the quality of the reference condition was not attained. In *Fig. 4b* the quality profiles of listening conditions AHS_2 and AHS_5 are represented by assessment “3” and in *Fig. 4c*, listening conditions AHS_6 and AHS_8 by assessment “4”.

To establish a relationship with the objective measurements given in *Figs 1* to *3*, the acoustic parameters are considered individually and in such a way that, in each case, two listening conditions with opposing assessments are compared. In *Figs. 5a - 5c*, the psychoacoustic parameters “coloration of bass, middle and treble frequencies” are considered; in *Figs 5d* and *5e* the parameters “reverberance” and “background noise”, and in *Fig. 5f* the parameter “imaging quality”. For all of the objective parameters, the tolerance limits according to [2] and the range of variation are given, so that assessment of the individual curve in relation to the overall results is possible.

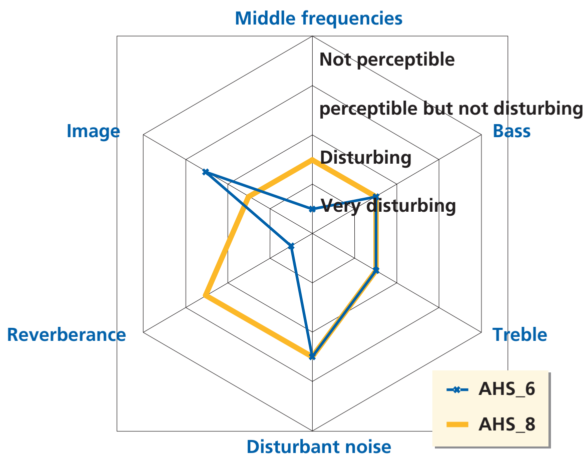
While a simple correlation between the subjective parameter “coloration” and the objective parameter “loudspeaker/



a) Assessment "2 = rather good"



b) Assesment"3 = rather satisfactory"



c) Assessment "4 = rather adequate"

Figure 4
Quality profiles – assessments 2, 3 and 4.

room response curve" is not observable, correlations between "reverberance" and "disturbant noise" are, on the other hand, clearly recognizable. The correlations are verified mathematically on the basis of Spearman rank correlation. The corresponding objective data records are obtained by simple calculations. In the case of "tonal colouring", the sums of the square deviations (SD) are calculated in the following frequency ranges:

- bass frequencies: 32 - 200 Hz;
- middle frequencies: 200 Hz - 4 kHz;
- treble frequencies: 4 - 16 kHz.

In the case of background noise, the SD is calculated in the range 200 Hz - 4 kHz while, in the case of reverberation, the mean reverberation time T_m is taken as the basis.

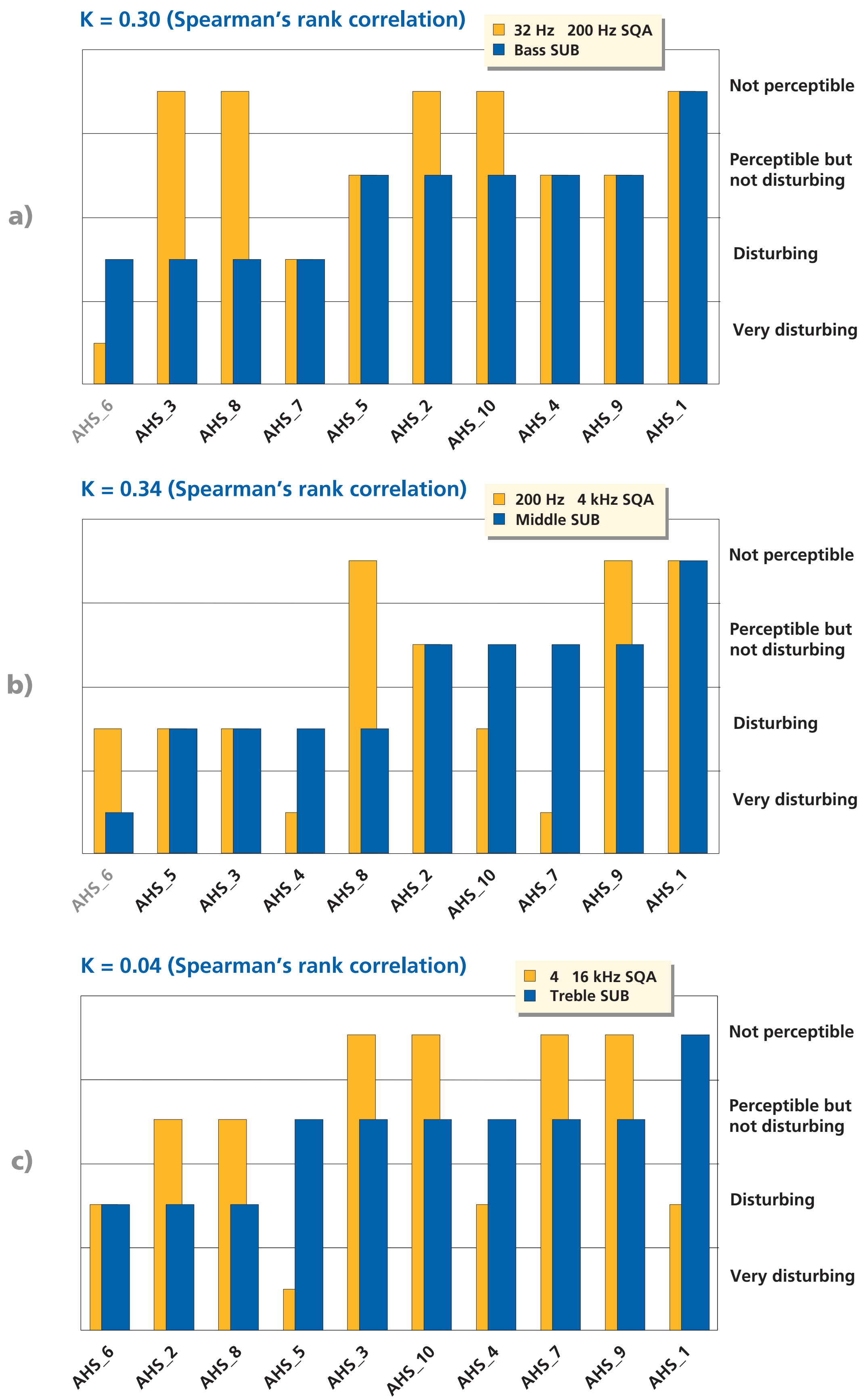
A comparison of the objective and subjective measurement data computed in this way is shown in *Fig. 5*. There is indeed a high correlation between "background noise" and "disturbant noise" on the one hand, and between "mean reverberation time" and "reverberance" on the other. The formally calculated negative correlation between "coloration of trebles" and "SD 4 to 16 kHz" is not held to be significant here, but simply shows some degree of dependence between the reproduction of treble frequencies and imaging quality.

To sum up, it may be observed that the psychoacoustic aspects "coloration" and "imaging quality" appear to be too complex to be simply computed from the objective measurement data. In regard to "background noise", the correlation with "disturbant noise" is not noteworthy, since here the EBU recommendation [2] is clearly overstepped and, on the other hand, observance of the tolerance limits NR 10 or NR 15 ensures disturbant noise that is "not perceptible" or "perceptible, but not disturbing". In regard to reverberation, these results do however suggest that the ITU and EBU recommendations [1][2][3] need to be reviewed. This is regarded as an important outcome of the present study. Although the "room volume" aspect cannot be neglected here, the mean reverberation time in studio listening conditions should not markedly exceed 0.3 seconds.

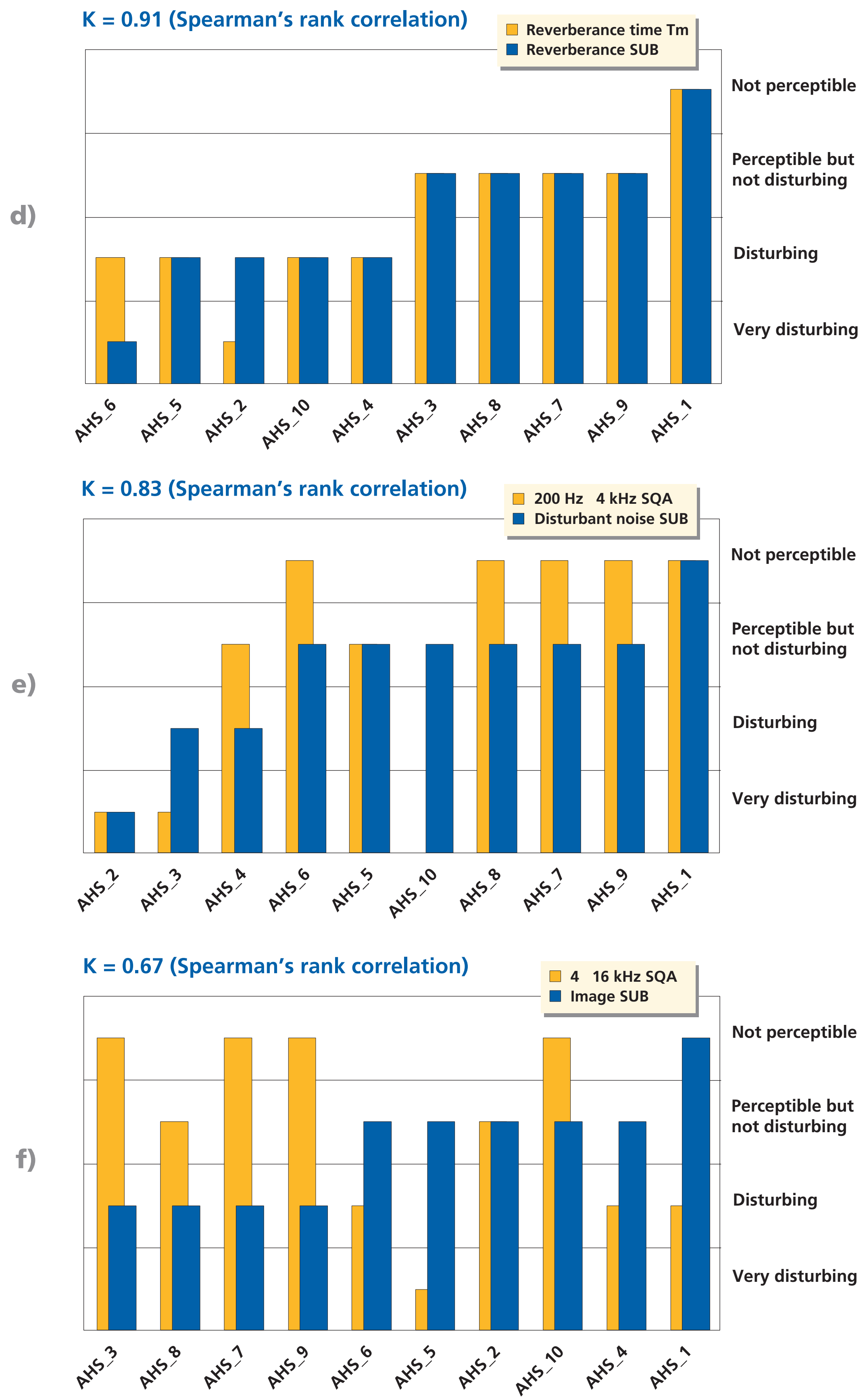
6. Conclusions

The results show that EBU Doc. Tech 3276 is to be regarded as a minimum requirement for studio listening and does not necessarily ensure neutral listening conditions. It is apparent that psychoacoustic assessment, on the basis of dummy head recordings, allows a more differentiated quality rating of studio listening conditions than objective assessment on the basis of acoustic measurements.

It should be borne in mind, however, that such differentiated results can only be obtained by A-B-A-B comparison. The A-B comparison, i.e. direct switching between two listening rooms, presupposes virtual imaging of the real situation, e.g. with dummy head technology. With the dummy head / headphones combination used here, it can be assumed



Figures 5a, b and c
Subjective / objective assessment – Coloration of bass, middle and treble frequencies.



Figures 5d, e and f
Subjective / objective assessment – Reverberance, disturbant noise and image quality.

Gerhard Spikofski studied electrical engineering at Berlin Technical University, one of his main areas of study being technical acoustics. Since 1980, he has been on the scientific staff of the Institut für Rundfunktechnik, Munich (IRT). His field of interest covers development and optimization of audio systems in broadcasting, with special reference to the psychoacoustic aspects.

Dipl.-Ing. Spikofski has published many articles in national and international specialist journals and is a regular speaker at national and international technical conferences. He is also a member of various national and international standardization bodies.

that the virtual imaging is authentic. However, when using a fixed dummy head, it is not always possible to exclude front/rear inversions. This drawback can be avoided by virtual imaging of the listening conditions with the aid of the BRS processor and inclusion of head movements to stabilize localization. Corresponding experiments with the BRS processor show a convincing relation with reality, especially in regard to comparison of different listening conditions. An important finding of the present study is that requirements with regard to reverberation time need to be reviewed.

The study further shows that the quality of listening conditions is open to improvement. The quality level of reference condition AHS_1 (RAR/B&W801) was not attained by any of the listening conditions. With regard to the psychoacoustic aspect, "coloration" and "imaging quality" in particular cannot be correlated with physical parameters. This means that this important psychoacoustic aspect cannot, objectively, be described with precision. Hence the subjective assessment of experts – as a complement to the measurement of sound-field parameters – will remain indispensable. We believe that, in this context, the BRS processor is a useful tool in optimizing the sound quality.

7. Bibliography

[1] ITU-R Recommendation BS.1116-1: **Methods for the subjective assessment of small impairments in audio systems including multichannel sound systems** http://www.itu.int/itudoc/itu-r/rec/bs/1116-1.html.

[2] EBU Tech. 3276 (2nd edition): **Listening conditions for the assessment of sound programme material: Monophonic and two-channel stereophonic** http://www.ebu.ch/tech_t3276.pdf.

[3] EBU Tech. 3276 (Supplement 1): **Listening conditions for the assessment of sound programme material: Multichannel sound** http://www.ebu.ch/tech_t3276_s1.pdf.

[4] G. Theile: **Virtuelle Lautsprecher in einem virtuellen Raum** (Virtual loudspeakers in a virtual room)
Fernseh- und Kino-Technik, Volume 52, No. 11/1998, (page 670 and following).

[5] P. Mackensen, U. Felderhoff, G. Theile et al.: **Binaural room scanning - A new tool for acoustic and psychoacoustic research**
Proceedings of the 137th Meeting of the Acoustical Society of America, Seattle, June 1998 (CD-ROM).

[6] U. Horbach, R. Pellegrini, U. Felderhoff and G. Theile: **Ein virtueller Surround Sound Abhörraum im Ü-Wagen** (A virtual surround sound listening room in an outside broadcast van)
Proceedings of 20.Tonmeistertagung [20th Conference on sound mixing], November 1998, Karlsruhe, Germany, pp. 238 - 245.

[7] U. Horbach et al.: **Design and applications of a data-based auralisation system for surround sound**
106th AES Convention, Munich, May 1999, (Preprint 4976).

[8] P. Mackensen, K. Reichenauer and G. Theile: **Einfluß der spontanen Kopfdrehungen auf die Lokalisation beim binauralen Hören** (Binaural recording: effect on sound localization of spontaneously turning the head)
Proceedings of 20.Tonmeistertagung [20th Conference on sound mixing], November 1998, Karlsruhe, Germany, pp. 218 - 228.

[9] G. Spikofski: **Psychoakustische Faktoren zur Beurteilung der Abhörsituation Regielautsprecher/Regieraum** (Psychoacoustic factors for the assessment of loudspeaker / control room listening conditions
17.Tonmeistertagung [17th Conference on sound mixing], November 1992, Karlsruhe: Report ed.: Papers of the Association of German Sound Mixers, Berlin 1992, pp. 646 - 661.

[10] G: Theile: **On the standardization of the frequency response of high-quality studio headphones**
Journal of the Audio Engineering Society, Vol. 34, 1986, pp. 956 - 969.

[11] G. Spikofski: **The diffuse-field probe transfer function of studio-quality headphones**
EBU Technical Review, No. 229, June 1988, pp. 1 - 16.

Abbreviations

AHS *Abhörsituation* (listening condition)

BRS Binaural room scanning

ETF Energy time frequency

HRTF Head-related transfer function

IRT *Institut für Rundfunktechnik* (Germany)

ITU-R International Telecommunication Union, Radiocommunication Sector

NR Noise rating

NRK *Norsk rikskringkasting* (Norway)

SD Sum of square deviation

YLE *Oy Yleisradio Ab* (Finland)

[12] ITU-R Recommendation BS.708: **Determination of the electro-acoustical properties of studio monitor headphones**
http://www.itu.int/itudoc/itu-r/rec/bs/708.html

[13] G. Spikofski and G. Karamalis: **Neue Aspekte zur Beurteilung von Regielautsprechern** (New parameters for the assessment of monitoring loudspeakers)
Rundfunktech (RTM), Issue 35 (1991), pp. 102 - 115.

[14] M. Wöhr and B. Nellessen: **Untersuchung zur Wahl des Hauptmikrofonverfahrens** (Study on the selection of main microphone procedures)
14.Tonmeistertagung [14th Conference on sound mixing], November 1986, Munich: Report ed.: Papers of the Association of German Sound Mixers, Berlin 1984, pp. 106 - 120.