



# Audio Engineering Society

## Convention Paper 7476

Presented at the 124th Convention  
2008 May 17–20 Amsterdam, The Netherlands

*The papers at this Convention have been selected on the basis of a submitted abstract and extended precis that have been peer reviewed by at least two qualified anonymous reviewers. This convention paper has been reproduced from the author's advance manuscript, without editing, corrections, or consideration by the Review Board. The AES takes no responsibility for the contents. Additional papers may be obtained by sending request and remittance to Audio Engineering Society, 60 East 42<sup>nd</sup> Street, New York, New York 10165-2520, USA; also see [www.aes.org](http://www.aes.org). All rights reserved. Reproduction of this paper, or any portion thereof, is not permitted without direct permission from the Journal of the Audio Engineering Society.*

## Frequency dependent Signal Correlation in Surround- and Stereo-Microphone Systems and the Blumlein-Pfanzagl-Triple (BPT)

Edwin Pfanzagl-Cardone<sup>1</sup>, Robert Höldrich<sup>2</sup>

<sup>1</sup> Edwin Pfanzagl-Cardone, Salzburg Festival, Salzburg, 5020, Austria  
[e.pfanzagl@salzburgfestival.at](mailto:e.pfanzagl@salzburgfestival.at)

<sup>2</sup> Robert Höldrich, IEM - Institute of Electronic Music and Acoustics, Graz, 8010, Austria  
[robert.hoeldrich@kug.ac.at](mailto:robert.hoeldrich@kug.ac.at)

### ABSTRACT

With the aim to recreate the original concert-hall sound-field as faithfully as possible in the control- or living-room recordings were made simultaneously with an artificial head and several surround-microphone techniques (among them the new BPT method). The surround recordings were re-recorded using the same dummy-head as in the concert-hall. The results of subjective listening tests (loudspeaker as well as binaural) were assessed using ANOVA and correlation analysis.

Acoustical analysis of the dummy-head recordings was performed by measuring the 'Frequency-dependent Inter Aural Cross-correlation Coefficient' (FIACC): the 'low-correlation' AB-PC microphone system was capable of reproducing the original sound-field better than any of the other systems under test (DECCA, KFM, OCT).

A microphone-systems 'Critical Frequency', below which correlation raises towards 1, is defined.

### 1. INTRODUCTION

*"When I sit in an acoustically perfect hall (full of people), in the best seat for hearing, and listen to an orchestra, I hear such and such sounds. I want to hear precisely this effect from a recording, if that be possible."* – An anonymous music critic ("K. K.") writing in the 1928 issue of *The Gramophone*. [1]

At this moment audio history seems to be at a point where one part of the audio community is trying to push the borders of achievable quality in sound reproduction by using a rather high number of loudspeakers (wave field synthesis – WFS), or with the highest resolution possible (192kHz/24bit or SA-CD), while another – much larger – part of the industry is compromising said quality by intentionally diminishing

signal dynamic range (multiband compression as used on radio and tv), and/or applying data reduction based on psychoacoustic principles (MP3, etc.) for the sake of higher storage efficiency or gain in data transmission capacity. While the first author finds at least the last two practices highly debatable, it also seems clear to him that – at least for the recording and reproduction of acoustic instruments, especially in a multichannel or ‘surround-sound’ context (if a certain ‘minimal-quality’ is guaranteed) the choice of microphone technique (directional characteristics, as well as placement) is the most influential factor on the ‘quality’ which can be achieved upon playback.

While a recent publication [2] points out the deficits of WFS in terms of localization and sound coloration, a double-blind study in the field of high-resolution audio seemed to indicate that there is no evidence that perceived quality has to do with additional resolution or bandwidth [3]. Therefore the authors think it is worthwhile to concentrate on problems related to surround- and stereo-reproduction which employs a much smaller number of loudspeakers than WFS, for example the well established 5.1 Surround loudspeaker-setup as proposed in ITU-R BS. 775-1.

## 1. PROBLEMS RELATED TO LOCALIZATION IN ‘SMALL AB’ MICROPHONE SYSTEMS

A few years ago a study by Wittek and Theile [4] pointed out that there are large discrepancies between results derived from listening-tests concerning localization. As can be seen in Tab.1 these differences are – at times – larger than 100% (compare Wittek and Sengpiel for small AB with 50cm capsule-spacing).

| Setup                                     | Hugonnet    | Sengpiel    | Williams    | Wittek      |
|-------------------------------------------|-------------|-------------|-------------|-------------|
| <b>ab</b><br>omnis<br>50cm                | <b>130°</b> | <b>180°</b> | <b>100°</b> | <b>74°</b>  |
| <b>AB</b><br>omnis<br>100cm               | -           | <b>62°</b>  | -           | <b>36°</b>  |
| <b>ORTF</b><br>cardioids<br>110°,<br>17cm | <b>90°</b>  | <b>96°</b>  | <b>100°</b> | <b>102°</b> |

**Table 1:** Recording-angles as found in literature [5], [6], [7], [8], [9]

In [10] various studies are listed which indicate that the frequency dependence of sound-source localisation between loudspeakers has been known at least since 1934. A more recent study on this topic can be found – among many others – in [11]

Whether the ambiguity of the results found in study [4] are caused by differences in the nature of the test signals or varying physical characteristics of the loudspeakers involved [12], or if it may be based on strong inter-individual differences along with some anomalous localisation effects which occur in the mid-frequency range (see [13] for details), or an a combination of these factors may be an interesting subject to further research in the future.

Looking at the differences in the results of study [4] of more than 100%, it seems clear that ‘small AB’ based recording techniques apparently cannot provide localisation that is ‘solid’ enough so that inter-individual differences (including loudspeaker quality differences, etc.) can be neglected.

## 2. SMALL AB vs. LARGE AB

As is commonly known a time-of-arrival difference between the ears of a human in the range of 1.1 – 1.5ms [4] will result in localizing the sound event entirely in one of the two loudspeakers of a standard 2-channel stereo-setup. Therefore the majority of researchers accepts the ‘small AB’ based microphone configuration of two omni-directional transducers with small capsule spacing as the system of choice, as it is able to provide inter-aural time-differences in the requested range. On the other hand ‘large AB’ based microphone-techniques with capsule-spacings of several meters (as are commonly used with orchestra recordings) are often all too easily discarded by pointing out their inherent ‘hole in the middle’ effect, while in practice this acoustic ‘hole’ can very easily be filled with the help of an appropriate center-fill microphone or microphone-system.

Also – as will be shown further down – the presumed localisation accuracy (which is not that accurate at all, as evidenced by study [4]) of the small AB arrangement comes at the expense of the ‘spaciousness’ contained in the recording.

The preference of many academics for the small AB (and also various coincident XY) microphone systems

may also have to do with the fact that these are relatively easy to grasp mathematically while already a combination of just two (main-) microphone systems (e.g. Large AB with a Blumlein-Pair of crossed figure-of eights) may not be as easy to calculate in all necessary (psychoacoustic) accuracy.

And – in fact – if we take a look at the microphone setups that are being used by practicing sound-engineers, it can be noticed that there is quite a number who utilise not just one main-microphone technique, but rather a combination of two appropriate ones. Many times this includes also a ‚large AB‘-pair of spaced capsules (see the famous ‚DECCA-Tree‘ from the 1950-ies with omni-directional outriggers, as well as the ‚Fukada-Tree‘ for surround-recording [15]).

Other famous examples can be found in the highly acclaimed „Living Stereo“ recordings done by the RCA in the 1950-ies and 60-ies, of which the original 3-channel recordings have been re-issued in a (front signals only) multi-channel SA-CD format during the last few years. Fig. 1 shows an example of the RCA's "A/B - Centerfill" microphone scheme, which was applied for the 1959 recording of the "Pines of Rome" which used a ‚large AB‘ pair of Neumann U-47 microphones in conjunction with a Neumann SM-2 (early stereo microphone) as centerfill system.

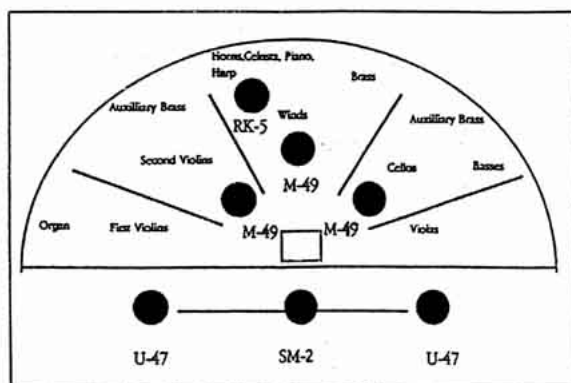


**Fig.1:** RCA's "A/B - Centerfill" microphone scheme (1959) [16]

Simply put, in order to achieve de-correlation of a 2-channel stereophonic sound signal we need to have either sufficient spacing between two omni-directional transducers or employ coincident or near coincident techniques with directional microphones.

Apparently there is a disagreement in respect to the amount of capsule spacing needed for "correct spatial reproduction" of a sound event: On the one hand side authorities of the audio community ([17], [18], [19]) seem convinced that only small microphone spacings ("small AB"), based on psychoacoustic principles, are able to provide correct localization. On the other hand a large percentage of practicing sound engineers favor largely spaced AB techniques ("large AB") [20], or at least use supplementary "outriggers" (largely spaced omni-directional microphones in front of the orchestra), due to the more "open sound" they provide.

### 3. PROBLEMS RELATED TO PERCEPTION OF LOW FREQUENCIES

#### 3.1. Playback Considerations

As is common knowledge of practicing sound engineers, some microphone techniques are better suited for headphone reproduction, while others work better with replay via loudspeakers.

Listening to the signal of a "small AB" microphone recording via headphones, one gets a rather realistic impression of the sound event, as long as the capsule spacing is in the order of the diameter of the human head. (This effect is also true for some other near coincident techniques).

However, when listening to the same stereo signal via loudspeakers, the experienced listener hears that while there is a convincing stereo-panoramic impression of localization and depth of the sound-field at high frequencies, this impression gets lost more and more towards lower frequencies.

Alan D. Blumlein mentioned this effect already in his 1931 patent:

"With two microphones correctly spaced and the two channels entirely separate it is known that this directional effect can also be obtained for example in a studio, but if the channels are not kept separate (for example by replacing the headphones by two loud speakers) the effect is largely lost." [21]

He then proceeds in proposing an electronic compensation circuit ("Blumlein Shuffler"), as well as head-related recording techniques with obstacles between two pressure transducers and his famous "Blumlein Pair" composed of crossed figure-of-eight microphones, among a plethora of other inventions relating to stereophonic sound.

The above mentioned effect -being quite obvious acoustically- needs some empirical as well as theoretical explanation. The change in spatial representation towards lower frequencies can be checked by listening selectively to isolated frequency bands to detect which room impression is provided by a main-microphone signal in various frequency ranges. In this respect the frequency range below approx. 700-800 Hz is of primary interest, since the human head is not yet effective as a baffle in this frequency band and due to signal-diffraction sounds "bend" around it. Above about 800Hz the shadowing capacity of the head becomes more and more important and the human hearing mechanism is therefore mainly based on the interaural differences in terms of signal loudness and frequency response, while - simply put - in the low frequency region it is mainly based on detecting interaural phase- and timing-differences. (see also Fig. 2 in the appendix, from [22])

Research by Yost [23] in 1971 has shown that the low frequency components of transient sounds in a binaural signal are of greatest importance for localization: High-pass filtering of clicks, so as to include only energy above 1500 Hz resulted in a deterioration of localization ability, but the same clicks low-pass filtered to include only energy up to 1500 Hz resulted in little change.

Research by Hirata [24] deals with the phenomenon of localization deterioration for low frequency components of a stereo signal that is replayed via loudspeaker. He proposes a "perceptual inter-aural cross-correlation coefficient" (PICC) as follows:

$$\text{PICC} = \text{DR}_0 + (1-D)\text{R}_E \quad (1)$$

with

D ... definition (directness) of sound

$\text{R}_0$ ... the interaural cross-correlation coefficient of the direct sound (unity for normal incidence)

$\text{R}_E$ ... the interaural cross-correlation coefficient of reverberant (incoherent) sound, expressed by:

$$\text{R}_E = [\text{sinkr}(f)]/\text{kr}(f) \quad (2)$$

where

$k = 2\pi f/c$  ... the wave number

$c$  ... speed of sound, and

$r(f)$  ... the effective acoustic distance between the ears, which is approximately 30cm. [25, 26]

He also defines an index of acoustic spatial impression ASI, expressed by:

$$\text{ASI} = (1-D) \cdot 100 \quad [\%] \quad (4)$$

Full spatial impression is indicated by  $\text{ASI}=100\%$  and no spatial impression by  $\text{ASI}=0\%$ .

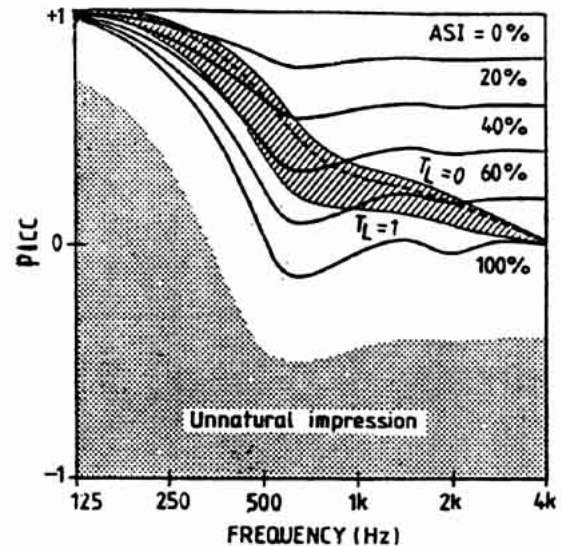


**Fig. 2:** PICC curves for stereo sound reproduction in a listening room of reverberation time  $T_L$  (0 to 1s) shows small ASI in low frequency band compared with an  $\text{ASI}=60\%$  for the middle seat of a concert hall. Broken line shows  $T_L=0.3s$ . (from [24])

Fig. 2 shows that in a standard listening room ( $\text{RT}60=0.3s$ ) the ASI in the stereophonic soundfield is small at frequencies less than 800Hz and large at frequencies greater than 800Hz in comparison with that in the concert hall, where  $\text{ASI}=60\%$ .

Research by Griesinger on envelopment [27] has shown that while for higher frequencies the ideal loudspeaker placement moves towards the median plane, for frequencies below 700Hz the ideal placement is the one which provides the most separation between the radiators, i.e. at the sides of the listener at  $\pm 90$  deg.

The research by Hirata and Griesinger indicates that the standard loudspeaker placement for stereo (as well as for 5.1, see the International Telecommunications ITU-R guidelines [BS.775-1]) with speakers at  $\pm 30$  deg. provides less than perfect conditions in respect to fidelity for low frequency spatial reproduction.

### 3.2 Recording Considerations

Research in relation to concert hall acoustics by Morimoto and Meakawa [28] has shown that spaciousness is strongly related to signal content in the frequency range below 500Hz. It is lower frequency components and the IACC (Inter-Aural Cross-Correlation Coefficient) of a stimulus that affect spaciousness independently; the 100–200Hz frequency range proved to be of particular importance.

Additionally, research by Hidaka, Beranek and Okano [29] on acoustical quality in concert halls found that changes in low frequency levels made greater changes in respect to Apparent Source Width (ASW, traditionally linked to “spaciousness”), than changes in the high frequency levels.

This shows that for reproduction, as well as recording, fidelity (amplitude- and phase-wise) in respect to low frequency signal-components is of paramount importance in order to achieve a sonic impression as natural as possible.

The author is convinced that only microphone techniques, which provide sufficient decorrelation (not only for mid- and high-frequencies, but also for low frequencies) are well suited for this purpose. This essentially rules out quite a number of well-established microphone techniques, among them “small AB”, as will be explained below.

#### 3.2.1. “Small AB” vs. “Large AB” in Orchestral Recording

While it may be possible to achieve a good sounding recording of a small to medium size instrumental ensemble with a small AB system, this technique is not adequate for capturing the full range sound and spatial width of a large symphony orchestra.

The narrow spacing of two omni-directional microphones has one major disadvantage: the stereo microphone signal is not sufficiently decorrelated for low frequencies - in other words: the stereo signal gets more and more monophonic as frequency decreases.

The correlation of the stereo signal is of course determined by the amount of microphone spacing: a spacing of 40 cm for example leads to negative correlation at around 600 Hz and below this frequency the correlation increases steeply. (see Fig. 3)

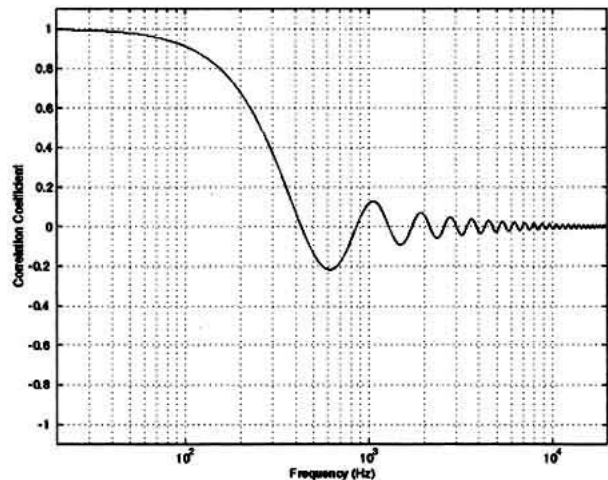


**Fig. 3:** Correlation coefficient vs. frequency for a pair of omnidirectional microphones with a separation of 40cm in a diffuse field (from [30])

Once the microphone spacing is in the order of half the acoustic wavelength at a given frequency, sound (from the side) can still be picked up 180° out of phase. Below that frequency sound will necessarily be picked up with less than 180° phase offset, which means essentially that the correlation coefficient can only get more positive. In other words: as frequency drops below this “critical frequency” ( $f_{crit}$ ) the signal becomes more and more monophonic.

As also proposed by Hecker [31], the “critical frequency” below which the correlation increases towards 1 can be calculated approximately as:

$$f = c / (2r) \quad (5)$$

with:

$f$  ... frequency in [Hz]

$c$  ... speed of sound in [m] (approx. 340m/sec at 20°C)

$r$  ... microphone spacing in [m]

As an example, a “small AB” microphone spacing of 40cm will deliver accurate spatial reproduction (for replay via loudspeakers) only down to approx. 430 Hz.

Therefore, in order to achieve spatial fidelity down to 40Hz, two omni-directional capsules need to be spaced by at least 4.3 meters:

$$r = c / (2f) \quad 343 / (2 \cdot 40) = 4.29\text{m}$$

This spacing is still narrow enough to work well with small to medium size orchestras.

The already very high degree of correlation of a small AB microphone arrangement will be made even higher by the low frequency "spatial deterioration" effect of a standard stereo loudspeaker setup (see Hirata).

With an AB-pair of omni-directional microphones the maximum negative correlation of the 2-channel signal is reached at a frequency  $f$ , where the capsule spacing  $d = 0.66 \lambda$  (see Fig. 4, 5 below)

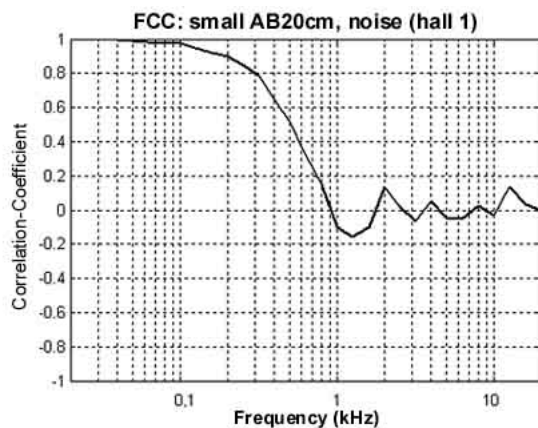


**Fig. 4:** Measured FCC ('Frequency-dependent Correlation Coefficient') for a 'small AB' pair of omni-directional microphones with 20cm capsule spacing. ( $f_{\text{crit}} = \text{approx. } 900\text{Hz}$ )

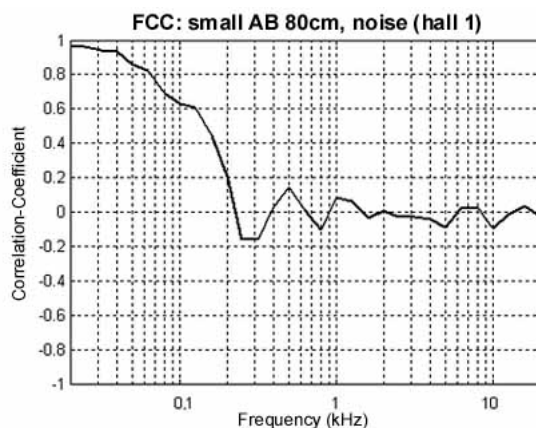


**Fig. 5:** Measured FCC ('Frequency-dependent Correlation Coefficient') for a 'small AB' pair of omni-directional microphones with 80cm capsule spacing. ( $f_{\text{crit}} = \text{approx. } 220\text{Hz}$ )

$$\lambda = c / f \quad \lambda \dots \text{wavelength} \quad (6)$$

Calculation of frequency of the most negative correlation:

Fig.4:

$$\lambda = 34000\text{cm} / 1100\text{Hz} \cdot \text{sec} = 30.90\text{cm}, \text{ therefore } r = 0.66 \lambda \sim 20.4\text{cm}$$

Fig. 5:

$$\lambda = 34000\text{cm} / 290\text{Hz} = 117\text{cm}, \text{ therefore } r = 0.66 \lambda \sim 78\text{cm}$$

The relationship between capsule-spacing ( $r$ ) and the resulting correlation-coefficient  $R$  (in a reverberant sound field) in respect to frequency ( $f$ ) was for the first time calculated and measured in the mid-fifties [32] (This is about around the same time when the recording teams of various record companies experimented with the challenges of stereo-signals and their correlation on a more practical level):

$$R = (\sin k r) / k r \quad (7)$$

with  $k = 2 \pi / \lambda$  (see equation 3)

$$k r = 2 \pi r / \lambda \quad (8)$$

A complete derivation of the results can also be found in [33].

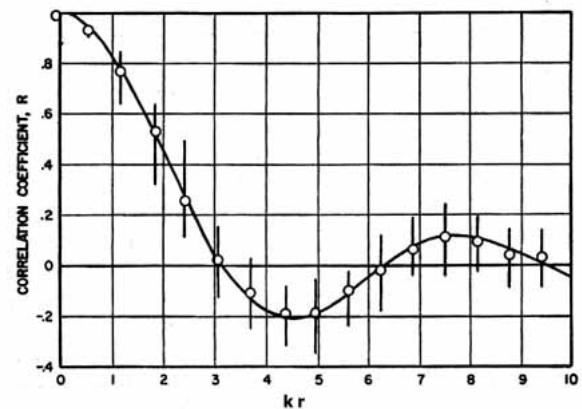


**Fig. 6:** Correlation-Coefficient  $R$  vs. distance  
solid line = calculation; o = measurements (from [32])

#### 4. CORRELATION vs. COHERENCE

A detailed analysis concerning the differences between signal correlation and signal coherence can be found – among others – in [34]., a brief synthax in [35 Martin]

Mathematically the correlation between two signals can be defined as:

$$R = \int_0^T p_1(t)p_2(t)dt / \left[ \int_0^T p_1^2(t)dt \int_0^T p_2^2(t)dt \right]^{1/2} \quad (9)$$

where  $p_1(t)$  and  $p_2(t)$  are the respective instantaneous sound pressures at time  $t$  at the two points of measurement. By definition  $R$  can assume values between  $-1$  and  $1$ : a correlation of  $1$  means that the two signals are identical,  $-1$  that they are identical but of opposite polarity ( $180^\circ$  out-of-phase),  $0$  means that they are not correlated.

Signal-coherence on the other hand is defined as:

$$C_{xy}(f) = \frac{|P_{xy}(f)|^2}{P_{xx}(f)P_{yy}(f)} \quad (10)$$

The coherence is a function of the power spectral density ( $P_{xx}$  and  $P_{yy}$ ) of  $x$  and  $y$  and the cross power spectral density ( $P_{xy}$ ) of  $x$  and  $y$ .

For the measurements carried out in this research the ‘mscohere’-function in MATLAB was used, which finds the magnitude squared coherence estimate  $C_{xy}$  of the input signals  $x$  and  $y$  using Welch’s averaged, modified periodogram method. The magnitude squared coherence estimate is a function of frequency, with values between  $0$  and  $1$  that indicate how well  $x$  corresponds to  $y$  at each frequency.

##### 4.1 Optimal Signal Coherence in Surround-Microphone Systems in Relation to Capsule Spacing

In various of his papers [36], [37] Günther Theile promotes the idea that a minimal coherence of  $k \geq 0.35$  is desirable for the signals of a (quadraphonic) surround microphone system, in order to achieve “...natural imaging of space and envelopment...” (see [36], pg.33).

He draws this conclusion mainly on the grounds of his interpretation of research carried out by Damaske during the late 1960-ies [38]. For most of his experiments Damaske had used bandlimited noise and music signals ( $250\text{Hz} - 2000\text{Hz}$ ) with varying degrees of coherence in order to study the resulting impression of „subjective diffuseness“ dependent on the spatial distribution of the loudspeakers and on the degree of coherence of the signals.

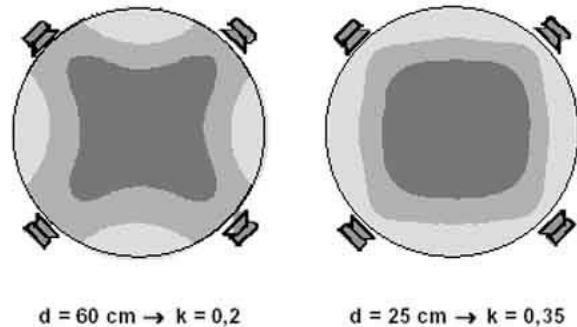


**Fig.7:** Perceived directions in a diffuse field generated by four loudspeakers radiating noise ( $0.25\text{-}2.0\text{kHz}$ ) with inter-channel coherence factors  $k$ . The noise was recorded in a reverberation room by omni microphones, spaced by distance  $d$ . (from [36], after [38])

In his paper Damaske writes (translated): „... to quote a degree of coherence for music is problematic, as it certainly varies with the momentary bandwidth of the music signal.“ He assumes that – according to his experiments - coherence of the music signals should not deviate much from the values which have been measured for the bandlimited signals (but always given a bandlimited nature of the signals - for noise, as well as music)(see [38], pg. 203)

In another part of the paper he writes that - according to the test listeners – with incoherent sound-fields the phantom sound-sources seem to come from ‚various directions‘, not necessarily limited to the baselines connecting the loudspeakers. It was also noted that the resulting „spatial impression“ was rather independent of head movement and judged to be more pleasant than listening to a ‚coherent‘ sound field (see [38], pg. 204)

Given these conclusions of Damaske the first author finds it surprising that Theile simply applies the values found for ‚optimal‘ coherence when using bandlimited noise as test signals also to (normally broadband) music signals, especially since the original author is reluctant to do so. (Rem.: One has to bear in mind that a signal

with a bandwidth limited to 250Hz –2000Hz has a quality that is even inferior to the (old) telephone frequency bandwidth specification...)

And in fact – as will be shown in one of the following sections – according to the experiments carried out by the authors, it was possible to prove that Theile's proposed minimal coherence of  $k \geq 0.35$  is not suited to ensure 'natural spatial impression'; evidence actually points towards decorrelation (or incoherence) for all frequency bands to be preferable for good spatial reproduction.

The first author is afraid that with this questionable interpretation of Damaske's research, Theile – along with other important opinion leaders and authorities in the field of audio [18], [19] – may have pointed a whole generation of sound-engineers as well as researchers in a wrong direction: to look for sufficient correlation in (surround) microphone signals instead of being concerned that they should be sufficiently de-correlated (incoherent) in all frequency bands ...

Many practising sound-engineers know the 'phenomenon' that they may use a (stereo- or) surround-microphone setup for music recording which provides seamless envelopment and a pleasant spatial impression, as long as the recorded signals are music, but – when the audience starts to applaud, suddenly the loudspeakers „come to life“ and the listener is made aware again that he is only listening to a representation of a music performance through speakers. The main reason for this lies – most likely – in the very different nature (spectral distribution and bandwidth) of music and noise (applause) signals. While (acoustic) music signals are usually characterised by reduced power in the higher frequency bands, the coloured noise signal (applause) is more similar to pink noise, which has equal power in any of the (e.g. 1/3-rd octave) frequency bands.

Loudspeakers however usually have a strongly frequency-dependent radiation characteristic with the low frequencies being radiated more or less omnidirectionally, while the mid- and high-frequencies are display a directional radiation pattern. This fact is also responsible for the loudspeaker being more easily discernible ('detectable') if the signal contains a lot of high frequencies. It could be argued that with loudspeakers that were characterised by a „constant directivity“ radiation pattern (as is approximately the case e.g. with B&O speakers using Sausalito Audio's

„Acoustic Lens“ technology for a more omni-directional dispersion of mid/high-frequencies) the results obtained through the tests with Damaske's group of listeners may have had a different outcome.

With traditional n-way (with  $n \geq 2$ ) dynamic loudspeakers (direct radiator principle) sound radiation is split in various frequency bands, which also means that even with signals with equal power distributions in equivalent bands, the electro-acoustic transducers assigned to the various frequency ranges have to radiate different amounts of acoustic energy. More concretely speaking the usually physically relative small high frequency driver has to radiate much more acoustic energy per square-centimeter than mid- or low-frequency drivers. This – in addition to the (by law of physics) higher directional Q for high-frequency dynamic drivers (if used without horns or waveguides) – enhances the effect that the listeners notices the HF-driver as sharp point-source much earlier than the lower-frequency drivers (given the appropriate nature of the signal).

This acoustic effect would probably be less pronounced with loudspeakers of different radiation principles, which use larger membranes sizes (or reflective surfaces) to radiate high frequency signals (e.g. electro-static speakers, or the above-mentioned 'acoustic lens' system). (see also [39])

## 5. MEASURES RELATED TO SPACIOUSNESS

Researchers have proposed various measures for acoustic parameters related to spatial impression and envelopment:

Traditionally spaciousness is regarded to be linked to "apparent source width" (ASW) and "listener envelopment" (LEV). The research of Marshall [40] and Barron [41] has proved the importance of lateral reflections for spatial impression in concert halls. Lateral Energy Fraction (LF) has been used as a measure of apparent source width (ASW) by Barron and Marshall [42], and Kleiner [43] has developed the following measurement method:

$$LF = \left[ \int_{0ms}^{80ms} p_8^2(t) dt \right] / \left[ \int_{0ms}^{80ms} p^2(t) dt \right] \quad (11)$$

where  $P_8(t)$  is the sound pressure measured at a location in the hall by a figure-8 microphone with the null axis pointed to the source, and,  $p(t)$  is the sound pressure measured at the same point with an omnidirectional microphone. Usually the integration for LF is performed over the first 80ms after arrival of the direct sound.

The "Interaural Cross-Correlation Function" IACF ( $\tau$ ) is a binaural measure of the difference in sound at the two ears and, hence, of lateralness:

$$\text{IACF}(\tau) = \frac{[\int_{t_1}^{t_2} p_L(t) p_R(t+\tau) dt]}{[\int_{t_1}^{t_2} p_L^2(t) dt \int_{t_1}^{t_2} p_R^2(t) dt]^{1/2}} \quad (12)$$

where  $P_L$  and  $P_R$  designate the sound pressure level at the entrances to the left and right ear, respectively.

If we select the maximum value of this equation we get the "Interaural Cross-Correlation Coefficient" (IACC):

$$\text{IACC} = |\text{IACF}(\tau)|_{\max} \quad \text{for } -1 < \tau < +1 \quad (13)$$

In 1996 Griesinger [44] proposed the "Lateral Early Decay Time" (LEDT) as a measure for spaciousness:

$$\text{LEDT}_{350} = \frac{60 * 350 \text{ ms}}{(S(0) - SD(350)) * 1000 \text{ ms} / \text{sec}} \quad (14)$$

where  $S(t)$  is the Schroeder integral of the impulse response, and

$SD(t)$  is the Schroeder integral of the Interaural Difference, the IAD.

A lot of factors contribute to spatial impression: more recent research by Griesinger [45] has shown that fluctuations in the Interaural Time Delay (ITD) and the Interaural Intensity Difference (IID) contribute significantly to the spatial properties of a sound event.

While ASW and IACC are useful in concert hall measurement, they do not work equally well in smaller rooms. Therefore Griesinger has proposed two new measures [46]:

The "Diffusefield Transfer Function" (DFT) as a measure of envelopment, which is useful both in small and large rooms, and

the "Average Interaural Time Delay" (AITD) as a measure of "externalization", a sonic property unique to small rooms.

## 6. RECORDING OF PROGRAM MATERIAL:

### 6.1. Surround microphones systems under test:

Taking up the proposal from a previous publication of the first author [47] recordings were undertaken using the following 5 surround microphone techniques:

- OCT-Surround ("Optimal Cardioid Triangle", as proposed by Guenther Theile of IRT, Munich with two parallel spaced cardioids pointing backwards for rear signal pickup (as described in [36])
- a DECCA-triangle style arrangement with three omnidirectional capsules, spaced according to psychoacoustic principles, utilizing interaural-time differences for localization [14]; rear: two cardioid microphones at a distance to the front and separated by approx. 10m (see AB-PC system below)
- the KFM 360 (Kugelflaechenmikrofon – "sphere microphone", invented by Guenther Theile [48] and built by company Schoeps) with two figure-of-eight microphones at the sides to enable dematrixing of the sound to 5.1 surround with an appropriate processor (DSP-4) [49]
- the SOUNDFIELD microphone with the appropriate processor SP451, based on the invention by M. Gerzon [50]
- the AB-PC ("AB-Polycardioid Centerfill") System as proposed by the first author [47]: with a spacing of the 'large AB' omni-directional capsules of 12m, and the use of a so-called "ORTF-Triple" (abbr.: ORTF-T) as centerfill system. For the rear channels the signals of the above described rear microphones for the 'DECCA'-system were used. (see Fig.43)

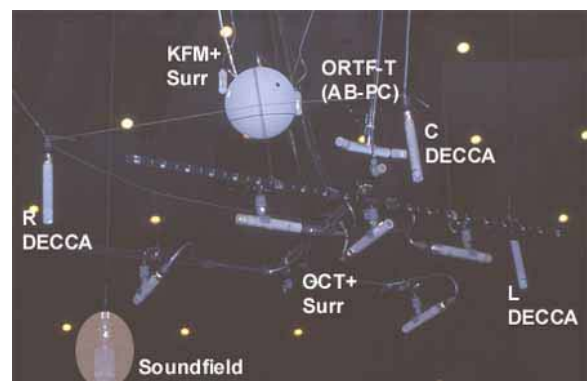


Fig.8: Surround-Microphone systems under test

As the KFM 360 with its 18cm diameter provides a recording angle of  $120^\circ$  for stereo (according to the manufacturer's specifications), the recording angles of the other systems were set accordingly, where possible (i.e. via capsule spacing for the OCT and DECCA Systems, and by means of the SP451 processor for the Soundfield microphone)

All microphones were suspended from the ceiling of the concert hall, between the 2<sup>nd</sup> and 3<sup>rd</sup> row of seats at a height of approx. 3.5m (12 feet) above stage level [except for the Soundfield microphone, which was moved approx. 2m (7feet) further out in the hall, due to its tendency to suppress diffuse-field components and – therefore – sound much drier.]

In order to enable an objective comparison of the microphone systems under test and not get influenced by manufacturer-specific sonic differences only microphones of one manufacturer were used. (Rem.: this was possible, as the recordings of the Soundfield-microphone were excluded from the listening tests; which was due to time constraints and as this system had already scored rather poorly with previous listening tests [14], [51] this decision was taken).

Along with the surround-microphone systems the signals of a Neumann KU81 artificial head were recorded. This dummy head was placed at an appropriate position (in respect to the direct/diffuse-sound ratio) in the hall (6<sup>th</sup> row), at which at least the instruments of the first rows of musicians were still inside the critical distance (or 'reverberation radius', which equals approx. 5.5m or 16.5 feet at mid-frequencies at the "Grosses Festspielhaus", Salzburg).

## 6.2. Program material and venue acoustics

### 6.2.1 ORCHestral recording:

The program material recorded was a symphonic orchestra piece by Lutoslavsky.

The 'Grosse Festspielhaus Salzburg' hall holds approx. 2250 seats, a volume of 14.000 cubic meters and is characterized by an  $RT_{60}=1.5$  sec (at 1kHz), corresponding to a reverberation radius of 5.5m.



**Fig. 8:** Salzburg Festival Hall (,hall 1‘) with suspended surround-microphone systems

The acoustic quality of this concert hall has been classified 'B+' (i.e. good to excellent) in an international study carried out by Hidaka, Beranek and Okano [29].

### 6.2.2. DUO recording:

The recording of the DUO piano-cello (neo-romantic literature) took place in a rehearsal stage with the following dimensions:

20.5m x 11.35m x 7.8m, with an approximate volume of 1815 cubic meters.



**Fig. 9:** Duo-Recording at the rehearsal stage (detail view)

### 6.3. Recording media:

The microphone signals were recorded to digital multitrack tape (24 tracks) in linear PCM resolution at 48kHz/16bit. Later the relevant sections for the listening test were selected and digitally transferred to a PC-based hard-disk recording system for ease of playback.

## 7. SUBJECTIVE LISTENING TESTS

### 7.1. Test design and goals

The double-blind listening test was set up as a randomized sequence of A/B-comparisons of the four surround-microphone systems under test. According to the number of microphones systems to be rated and including a few 'special cases' of A/B comparisons which enabled to evaluate the reliability of the respective test listener (redundant pair, identical stimulus pair, etc.) this resulted in 10 A/B pairings of surround-samples to be judged, with 15 attributes (+ one similarity rating) to be rated each time.

### 7.2. Test-listener selection

The group of test listeners consisted mainly of sound-engineering students of the IEM, as well as practicing sound-engineers and one loudspeaker-designer. The disadvantage with such a non-trained, non-expert panel of test subjects may be that the results have a higher random-error, than would be the case with expert listeners.

On the other hand the results obtained with such a group is likely to be closer to the results which a group of 'average consumers' (or naïve listeners) would produce, and is therefore probably more significant or applicable to the group of potential listeners/consumers (i.e. the buying public) at large. A more detailed analysis about the challenge of subject selection can be found in [52].

### 7.3. Listener 'Calibration' and Scales

Before the actual rating process the subjects were exposed to the same 20sec segment of music recorded by each of the four microphone systems. This initial "subjective calibration section" was repeated at least two times, or more often upon request. The listeners were asked to try to remember the most extreme impressions for all acoustic attributes which they had to judge later on, and use these as minimal and maximal values for the 7-step scale, against which the recordings had to be referenced.

The listeners had to rate their preference for the respective microphone technique and the following 14 acoustic attributes:

naturalness, sound-color, high/low balance, localization, (volume) balance [between the instruments or instrumental groups of the orchestra], source width (stereophonic width), distance, depth, stability, differentiation, spatial impression (spaciousness), diffuse/direct sound balance (amount of reverb or wet/dry balance), rear channel-signals non-intrusive / intrusive, level of the rear-channels (soft/loud)

While for most of the attributes these ratings meant to be of qualitative nature (with smaller value = better), for the following attributes it has to be understood as a merely quantitative judgement in the first place: high/low balance, wet/dry balance, level of the rear-channels.

### 7.4. '5.1 Surround' Listening Test

#### 7.4.1. Orchestral recording

25 listeners rated the four surround-microphones techniques under test in the form of a double-blind A/B comparison. After the initial "Calibration Intro" they were allowed to switch between the A and B sample at any time. The two 80sec surround-samples of the respective microphone techniques were looping simultaneously to enable instantaneous switching between A and B without losing the musical context. The loudspeakers used were 2-way systems, which are found quite frequently in classical music production studios (Genelec 1032A), with a frequency range from approx. 42Hz to 21kHz according to the manufacturer's specification.

#### 7.4.2. Duo recording

A group of 15 listeners had to rate the following four surround-microphone techniques, which were also presented to them via surround-loudspeaker replay: OCT-Surr., DECCA+Surr., KFM and BPT+Surr. (the below described "Blumlein-Pfanzagl-Triple")

In the rating of the DUO recording the attribute 'localization' also became a quantitative (not qualitative) judgement, as the listeners had been asked to mark *where* they localized the cello within the L-R panorama.

#### 7.4.3. Duration of listening test:

Average listening times for both the ORCH and DUO microphone A/B-comparison test were about 60–90min.

### 7.5. Listening Room Acoustics

The acoustically treated listening room normally serves as a control room for surround production at the IEM. This non-rectangular room has the following properties: Approx. dimensions: 7.5m (w) x 6.5m (l) x 3.8m (h) RT60: ca. 0.25sec

### 7.6. Binaural Listening Test

A different group of 10 listeners rated the four microphone systems under test by being presented a binaural recording via headphones, which had been obtained in the following way: the surround-recordings of the four microphone systems had been replayed via the surround-loudspeaker setup at the control-room of the IEM and re-recorded with the same Neumann KU81 dummy head which had been used for the original recording at the concert hall.



**Fig. 10:** Re-recording of the surround-microphone signals at sweet spot of the control room at the IEM (rear channel loudspeakers outside of the picture)

The listeners were then asked to rate the four surround-microphones techniques in the form of a double-blind A/B/REF-comparison. (REF being the original dummy-head recording from the concert hall.) Similarly to the surround-replay situation, they were able to switch between the three microphone techniques, which were looping simultaneously, at any time. At the digital mixing desk, with which the switching was performed, the A, B and REF stimuli were clearly labeled, so the listeners knew which acoustic stimulus was the one (REF) they had to reference the two unknown techniques (A, B) to.

Before the actual rating process the listeners were

exposed to the above-mentioned (in this case: binaural) “calibration intro”. Due to the different replay method (headphone- instead of loudspeaker reproduction) some of the questions related to acoustical attributes did no longer make sense: those judgments related to the rear channels only (signals non-intrusive / intrusive and signals loud/soft) were therefore excluded. Due to this reason the number of pairings was also reduced, since two ‘special cases’ were no longer applicable: the number was reduced to 8.

One additional question however was included at the beginning of the test: the subjects were asked to make a judgment at how different the two presented microphone techniques A, B were in relation to the REF recording. Also, when giving their ratings for the 13 remaining acoustic attributes (+ 2 similarity ratings) the listeners were asked to also give a rating for the REF-sample (where applicable). This meant that the number of ratings to be made was raised by 50 %. Therefore – even though the rating should have been made easier than for the group of listeners with loudspeaker reproduction, due to the presence of the REFERENCE stimulus – in reality the rating took much longer, which was probably not only attributable to the higher number of ratings, but also due to the fact that the binaural representation of the recordings via headphones - of course - differs much more from ‘natural hearing’ than listening to loudspeakers and therefore requires more concentration and causes a greater effort. The average listening time was close to 90min per subject.

#### 7.6.1. Choice of headphones for binaural listening test

The Beyerdynamic DT990 dynamic headphones with diffuse-field compensation were chosen, as they are characterized by almost FEC-properties (FEC – “free-air-equivalent coupling”; see [53] for more details). Due to these properties and according to research by Moller et al. [54] for some high-quality headphones further compensation (equalization) is not necessary, which also applies to the DT990.

#### 7.7. Choice of artificial head

On the way from the original sound-field in the concert hall to the listener’s ears the signal is subject to several changes, also in respect to its frequency response: dummy-head HRTF’s (head-related transfer functions, including outer-ear (pinnae) equalization), the transfer characteristics of the dummy-head microphones, the

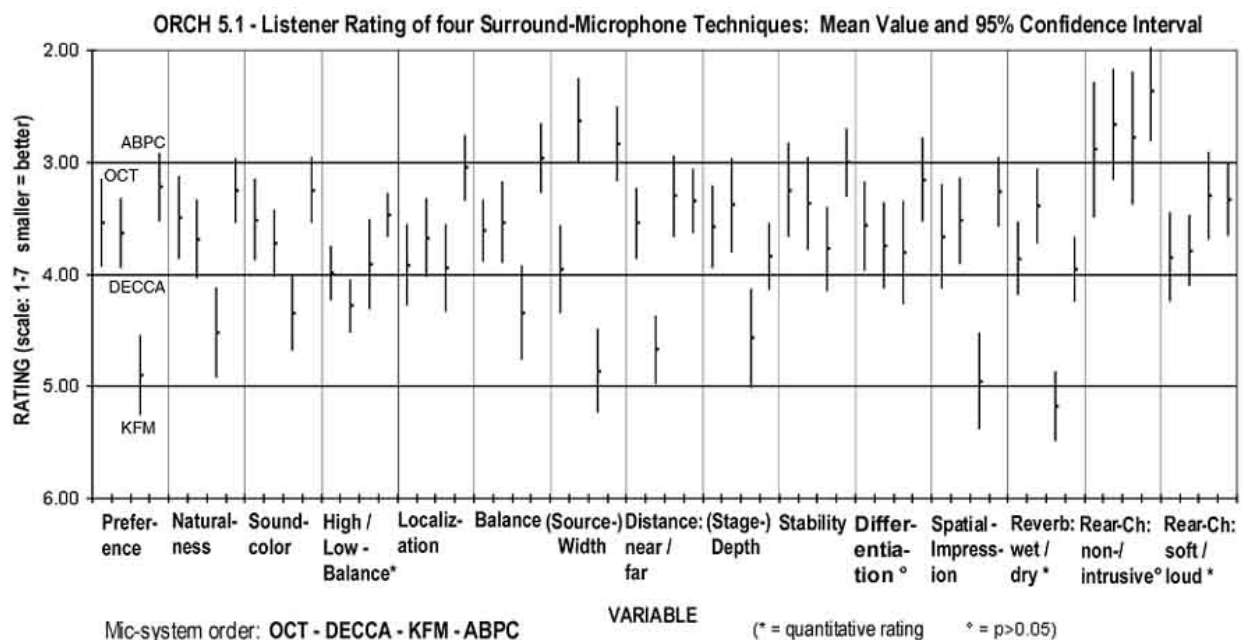


Fig.11: 'ORCH 5.1' test results (25 listeners)

non-linearities of the storage medium or transmission-channel and finally the transfer characteristics of the headphone, before being filtered by the outer-ear/ear-canal of the test listener. Ideally, the signal alterations in the path from the dummy head to the headphones should be compensated for by inverse filtering, which means that the combined transfer function of dummy head and headphones need to be measured.

According to the manufacturer's specification the Neumann KU81i does not need additional frequency compensation, when listened to via diffuse-field equalized headphones, which is the case with the DT990 (see [55]).

## 8. ANALYSIS OF THE LISTENING TESTS

### 8.1. Orchestral Recording

#### 8.1.1. '5.1-Surround' loudspeaker replay

The data obtained through the listening test were subject to ANOVA (ANalysis Of VAriance) using an appropriate function in 'MATLAB'-software. A p-value of  $p < 0.05$  was chosen, the results are displayed in the appendix in Table A1.

Concerning the ratings of the 15 acoustic attributes (or variables) 13 turned out to be statistically significant ( $p < 0.05$ ); (exceptions: 'differentiation' and 'rear-

In Fig.11 the mean values and 95% confidence intervals of the attributes are displayed. For most attributes the differences between the microphone-systems are not statistically significant (overlapping 'error bars' or confidence intervals). For more than half of the attributes however, at least one mic-technique differed significantly from all others. In many cases it was the KFM-microphone, which was 'singled out': this system seemed to be characterized by providing a significantly more "dry" recording, than the others, which also resulted in less spaciousness. Its stage depth, width, balance, sound-color and naturalness have been rated low. Not surprisingly, the overall preference of the KFM-microphone is therefore much lower than for the other surround-microphones, which coincides well with the findings of other studies. (see [51], [14])

The mean values for the OCT and DECCA system are in close proximity to each other for almost all acoustic attributes (except 'source width' and 'distance near/far'). For both, preference appears to be 'above average' (i.e. smaller than the scale middle-value 4), which coincides well with the findings of other studies [51]

The AB-PC system received the best mean-value ratings for two thirds of all attributes. Consequently it also scored best in terms of listener preference (even though

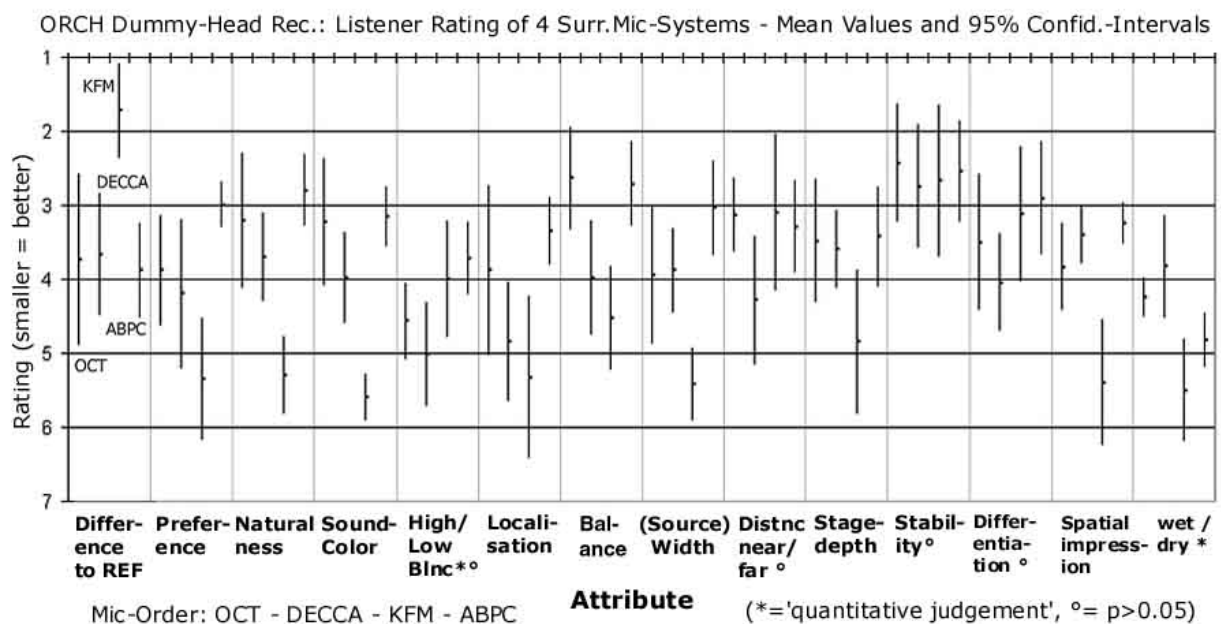


Fig.12: “ORCH – Binaural” test results (10 listeners)

not statistically significant; see overlapping error bars). The ‘sonic picture’ it provides seems to give very good results in terms of localization, stability and balance. It also scored best concerning spatial impression, sound-color and naturalness.

### 8.1.2. Orchestral Recording (‘binaural’ - headphone replay)

The data obtained through the listening test were subject to ANOVA (ANALYSIS OF VARIANCE) using an appropriate function in ‘MATLAB’-software. A p-value of  $p < 0.05$  was chosen, the results are displayed in the appendix in Table A2.

Concerning the ratings of the 14 acoustic attributes (or variables) 10 turned out to be statistically significant ( $p < 0.05$ ); (exceptions: ‘high/low balance’, ‘distance near/far’, ‘stability’ and ‘differentiation’).

In Fig. 12 the mean values and 95% confidence intervals of the attributes are displayed. For many attributes the differences between the microphone-systems are not statistically significant (overlapping ‘error bars’ or confidence intervals). For more than one third of the attributes however, at least one mic-technique differed significantly from all others. In many cases it was again the KFM-microphone, which was ‘singled out’: this system seemed to be

characterized by providing a significantly more “dry” recording, than most of the others, which also resulted in poor spatial impression. It sounds considerably flatter (stage depth) and narrower (source width) than its competitors and is rated inferior in terms of localization. In addition its sound-color and naturalness have been rated low.

Again, in line with all the low attribute-ratings received, also the mean value of preference of the KFM-microphone is therefore much lower than for the other surround-techniques.

Concerning the ‘difference to REF’ ratings it should be noted that the rule “smaller = better” (as applicable to all other ratings of this paper) has to be inverted, which is due to a mistake in the questionnaire (Rem.: the adjectives assigned to the min. and max.-values of the scale on the form should have been swapped). Therefore the fact that the KFM has been rated above value 2 in fact means that it is sonically quite different from the REFERENCE (original dummy-head recording).

The ABPC-technique has received the highest mean-values for 9 out of 12 qualitative attribute ratings. Not surprisingly it has – among those – again received the highest preference rating.

### 8.1.3. Comparison of the test results ‘ORCH 5.1’ and ‘ORCH binaural’

Overall the results of the two double-blind listening tests coincide very well: one surround-microphone technique (KFM) has clearly been singled out due to low ratings. The OCT and DECCA techniques usually have received quite similar ratings (exceptions: balance, localization, distance, source width). Concerning the rating for the attribute ‘source width’ it is interesting to note that in respect to the DECCA-technique there is a significant difference between the ratings received upon loudspeaker playback and headphone playback. As the DECCA technique is based on the psycho-acoustic principle of localization due to time-of-arrival differences (with almost no level-differences between the three front channel signals), it seems that these relatively subtle cues got lost to the listeners of the ‘ORCH binaural’ test (headphone representation), as they did not have the possibility to move (or turn) their head freely in the sound-field, as was the case with the ‘ORCH 5.1’ listeners. Therefore the attributes ‘source-width’ and ‘localization’ got both down-graded in the ‘binaural’ test in comparison to the ‘5.1’ test for the DECCA-system. It is also interesting to note that under headphone reproduction the DECCA signal is now rated to sound significantly less ‘far away’ than with ‘5.1’ loudspeaker reproduction.

The ABPC-technique has in both listening tests received the highest mean-values for the majority of the qualitative attribute ratings. It has consistently scored best in terms of preference, naturalness, sound-color, localization, and spatial impression.

Some readers may criticize that the ‘ORCH 5.1’ listening test is not complete in the respect that only an ideal listening position in the ‘sweet spot’ has been evaluated by the test-listeners. Albeit this is true, the authors would like to offer the proposal that maybe the ‘ORCH binaural’ test, which certainly puts the test listeners in a more difficult acoustic situation, as they have to listen to the surround-recordings by means “someone else’s” ears (i.e. those of the artificial head) and – in addition - through headphones, represents a sonic degradation which may be comparable to the disadvantages of an off-center seat position in a surround-loudspeaker setup.

As has been pointed out above, the DECCA-technique (based on ‘psycho-acoustic’ small time-of-arrival differences) seems to be much more vulnerable in respect to non-ideal listening conditions than the OCT

and ABPC techniques, which rely on strong level-differences (and also much larger time-of-arrival differences in the case of the ABPC) in the signals of the (front-) surround-channels.

This finding coincides well with other studies, which have been made in the field of 2-channel stereo [12].

One could have expected that the re-recording of the surround-microphone signals with the dummy-head and subsequent replay over headphones to the test-listeners may have resulted in a considerable sonic degradation which could have ‘blurred’ the sonic impression and therefore also made differences in the ratings smaller. This does not necessarily seem to be the case, maybe also due to the fact that a REFERENCE stimulus was present. Instead it looks like as if the ‘roughness’ introduced due to the transformation process may have ‘amplified’ the perceived differences between the surround-techniques under test. (Similar to a ‘grayscale’ picture being converted into a ‘black and white’ image.)

The good consistency between the results of the ‘ORCH 5.1’ and ‘ORCH binaural’ test supports the validity of this assumption.

## 8.2. Duo Recording (loudspeaker reproduction):

In addition to the ORCHESTRAL recording also a recording of a musical duo (piano and cello) was carried out with 3 of the above-mentioned surround-microphone systems: OCT, DECCA and KFM. The AB-PC technique was not applied, as it makes sense only with large sound-sources. However, another surround-microphone technique, which is also based on the principle of low signal correlation, but suited for the recording of small sound-sources was used as fourth technique: the newly introduced “Blumlein-Pfanzagl-Triple” (BPT).

### 8.2.1. The Blumlein-Pfanzagl-Triple (BPT) surround-microphone:

As the name already indicates, the BPT makes use of a third figure-of-eight capsule, which is added to the well known ‘Blumlein-Pair’ of two fig.-of-eight capsules, which have an included angle of 90° and are oriented at  $\pm 45^\circ$  towards the sound source. The new (center-channel signal) capsule is pointed directly towards the middle of the sound source at 0°. (see Figure 13 for schematic orientation of the capsules)

Signals for the rear channels can be obtained for example via cardioid microphones which are pointing towards the rear, with a capsule spacing of and at a distance of at least the critical distance (reverberation radius) of the room in order to ensure low correlation between the rear channels and between the front and rear channels. In order to increase signal separation between the front and rear channels (and to shield off unwanted rear-sound from the BPT front-system) it is preferable to place an acoustic barrier (for example in the form of a sound-absorptive acoustic panel) behind the BPT front-microphone. Should the available space be restricted, an alternative (much more compact) microphone set-up consists in a back-to-back arrangement of two BPT microphones, separated by an acoustic panel.

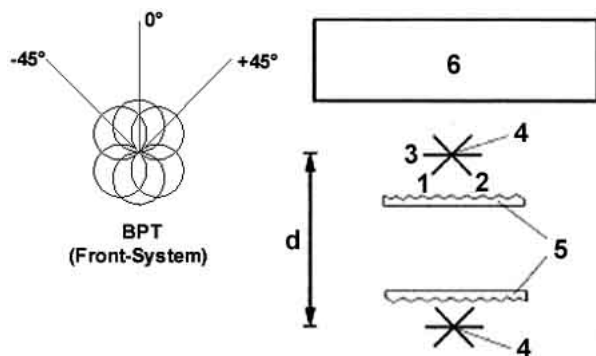


**Fig.13:** BPT – ‘Blumlein-Pfanzagl-Triple’ surround microphone schematic (left side); right side: two BPT’s (4) arranged ‘back-to-back’ with acoustic panels (5) in front of sound source (6). Distance d: 50cm to 10m



**Fig. 14:** Two BPT-Mics mounted ‘back-to-back’ on video-camera , with acoustic panels

The distance between the two BPT microphones can be enlarged (up to approx. 10m or 30feet) if a stronger sense of spaciousness is desired. With larger spacings, the time-of-arrival differences of sound would exceed 30ms and are therefore be likely to create problems

(‘echo effects’) upon playback.

Due to the strong directional characteristics of the fig-of-eight capsules, the middle section of a sound source would be over-represented level-wise when adding the third (center) capsule to the Blumlein-Pair. This ‘zoom’ or ‘center-focus’ characteristic can be of advantage for certain applications, if a microphone system with higher directionality is desired (on-location recording, etc.)

As can be seen from a simulation performed by using the “Image Assistant 2.0” by H. Wittek [56] if the level of the center-capsule is attenuated by approx. 10dB the increase of sensitivity towards the middle axis of the BPT-system is almost entirely compensated.

A similar surround microphone technique was suggested a few years ago by G. Martin at the 113<sup>th</sup> AES Convention in Barcelona, 2005 [57]. The differences between Martin’s technique and the BPT are as follows:

- the center capsule in Martin’s design is normally pointed downwards (by almost 90°) in a way to pick up sound from the front with an attenuation of 9dB in relation to the L and R microphone-capsules.
- he suggests to position the center-capsule at a vertical distance of at least 6cm from the L/R-capsules in order to avoid unwanted sound-coloration (comb-filtering) due to high correlation between the L, C and R-microphone signals
- for obtaining rear-channel signals he suggests a second triple of microphones arranged just behind the front-triple

Concerning the orientation of the center capsule it is interesting to note that Martin’s proposal to tilt the microphone down in order to achieve a level reduction of about 9 dB coincides well with the proposal for the 10dB attenuation for the BPT in order to achieve a more uniform overall L to R level-balance.

The physical separation of the center-capsule in order to reduce comb filtering effects, as proposed by Martin, most likely has its full right, as the correlation-over-frequency measurements (see appropriate section further down) of the BPT section unveiled high correlation between the L/C and C/R channels in the frequency range slightly below 2kHz, which could result in unwanted sound coloration due to comb-filtering. However, if this effect occurs in practice will depend on various factors like distance between the sound-source and the microphones, the radiation characteristics of the sound-source, room reflections,

etc. With test recordings with a hybrid version of the BPT carried out by the first author, this effect turned out to be very unproblematic.

Martins suggestion to use Blumlein-Pair style configuration for the pick-up of rear channel signal information in close proximity to the front unit. However the suggested angle between these two fig-of-eight microphones is smaller than the usual 45°.

The authors question if the resulting channel separation between the front and rear channels is high enough (without the use of acoustic panels or other sound absorptive provisions) to provide a convincing spatial impression, as it can be assumed that (at least for low frequencies) correlation of the channel signals will be high.

The surround-microphone technique proposed by Martin has been made public on the occasion of the 118<sup>th</sup> AES Convention, which took place in Barcelona from May 28th to 31<sup>st</sup>, 2005; the first author has applied for an innovation patent regarding the BPT surround-microphone at the Australian Patent Office (IP Australia) and the patent application was published officially on May 26th, 2005 [Australian Patent Number AU2005100255]. Therefore the first author would like to claim priority for the invention as described in the patent, which is referred to as “Blumlein-Pfanzagl-Triple”. [58]

### 8.2.2. DUO-Recording listener ratings:

Concerning the ratings of the 15 listeners regarding the DUO recording, 13 of the 15 attributes provided statistically significant results (except: differentiation, rear-channels non-/intrusive).

Overall it can be noted that for most attributes the confidence intervals are much larger than with the 'ORCH 5.1' listening test, independent of microphone technique. The two ratings that stick out most are the 'nearness' of the KFM-system (see: distance near/far), and the high amount of diffuse sound (see: reverb wet/dry) picked up by the DECCA system.

Concerning localization, which has been redefined to give a quantitative (instead of qualitative) rating, confidence intervals are much smaller and the OCT system deviates the most from a 'slightly right of center' position, where the cello was physically situated.

With the DECCA recording, informal listening indicates that it is most likely the high amount of diffuse sound (which results also in a highly enveloping room-acoustics) which is of disadvantage for a more precise localization.

What concerns the BPT system, informal listening (as well as the listener ratings; even though not statistically significant) indicate that it provides the highest stage depth, but – as a consequence – this also results in a

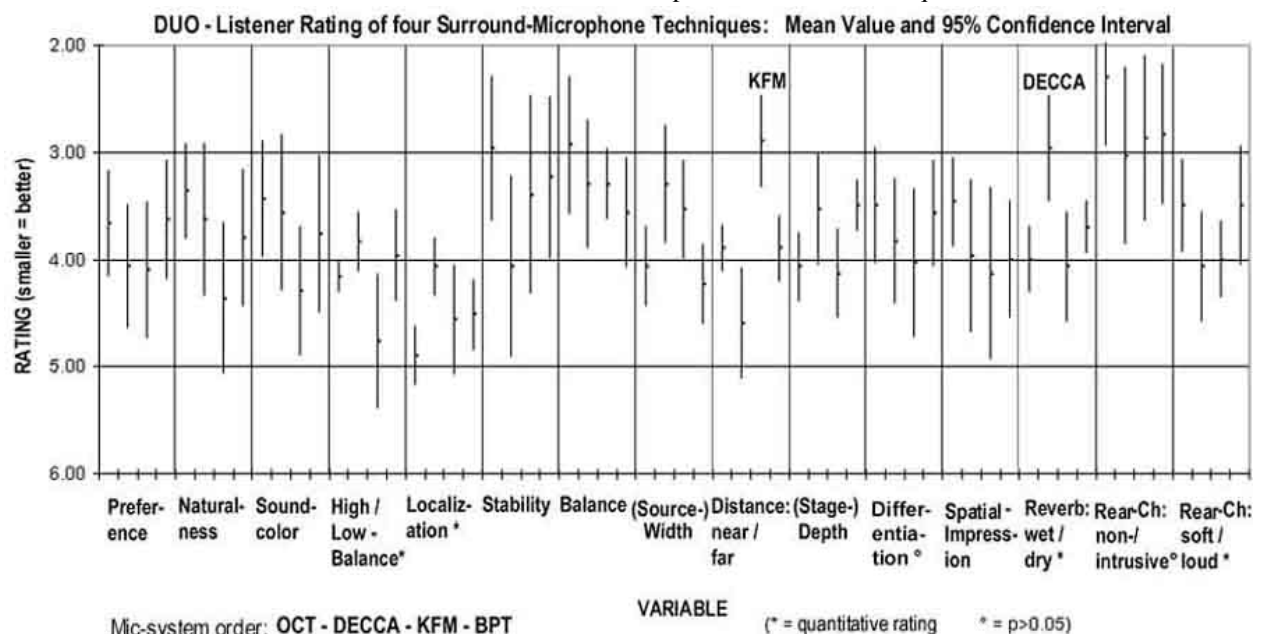


Fig. 15: 'DUO 5.1' test results (15 listeners)

poorer balance than with the other surround-systems.

It is also interesting to note that even though the BPT-system has achieved the highest mean-value only at one of the (qualitative) attributes (i.e. stage depth), it has nevertheless also received the highest mean-value in terms of listener preference.

## 9. CORRELATION ANALYSIS:

The results of the 'ORCH 5.1' listening test were subject to correlation analysis, based on the Pearson product moment correlation coefficient, using MATLAB-software. It seems of primary interest, which attributes have the highest correlation to (listener-) preference: these are naturalness ( $r=0.84$ ), balance ( $r=0.69$ ), spatial impression ( $r=0.68$ ); sound-color ( $r=0.67$ ) and width ( $r=0.63$ ).

The attributes which have the highest correlation to naturalness are sound-color (0.75) and spaciousness (0.64). Stability, balance and localization follow with correlations with a value of around 0.6.

Apart from preference and naturalness, sound-color is most strongly correlated with spaciousness (0.57), followed by localization and width, both with values around 0.51.

Spaciousness is (apart from preference and naturalness) strongly correlated with width (0.65). Not surprisingly it is also quite strongly correlated to reverberation (wet/dry balance). ( $r=0.51$ )

Below a listing of the 6 strongest correlations, which were found:

| Attribute-Combination     | Corr.-Coeff. |
|---------------------------|--------------|
| preference – naturalness  | 0.8383       |
| naturalness – sound-color | 0.7539       |
| preference- balance       | 0.6879       |
| preference – spaciousness | 0.6829       |
| preference – sound-color  | 0.6712       |
| spaciousness – width      | 0.6522       |

For completeness the matrix of all 15 aspects can be found in the appendix (Tab. A3)

The authors are relieved to notice that naturalness has the highest correlation to preference, as this seems to indicate that the listeners still seem to have a strong sense for (and clear idea about) what "natural" sound is like. Together with the results of the acoustic

measurements it will become clear that the listener preference actually goes towards those surround microphone systems which are able to approximate the original sound-field in the concert hall best.

The outcome of the correlation analysis coincides well with the findings of other studies [59], [60].

Among others, the correlation found between the attributes 'preference' and 'envelopment' in [59] coincides well with values for correlation between 'preference' and 'spaciousness' of the current study.

Also, the high correlation between 'spaciousness' and 'naturalness' that was found coincides well with the results of previous research by Toole [61].

## 10. OBJECTIVE MEASUREMENTS

The technically most objective way to measure how good a particular microphone technique is able to translate an acoustic event to the listener at home with the least amount of alteration is probably the following:

1. make a recording of the sound event with the microphone technique of your choice
2. at the same time make a second recording at the "best seat in the house" position of a concert hall with a binaural artificial head (dummy head)
2. while reproducing the first recording through a stereo or 5.1 loudspeaker system make a third recording (using the same dummy head) in the sweet spot of the listening room
3. measure the correlation between the two dummy head recordings.

Of course one has to be aware that with this method signal distortion (concerning amplitude, frequency and phase) introduced by the replay system and the acoustics of the listening room will be included in the evaluation process and might bias the results.

However, the microphone technique, which produces the highest correlation (regarding the entire frequency range) between the original sound event and the re-recorded reproduction would appear to be the one with the highest fidelity.

The measurements carried out in a previous paper by the first author [47] were based on a much simpler approach: signal correlation over frequency was estimated for various 2-channel stereo main-microphone techniques. This was carried out using (frequency-band

specific) signal attenuation, which was caused by phase cancellation, as an indicator for the correlation of 2-channel signals.

The above-mentioned ORCHestral dummy-head recording from the concert hall which had been obtained at a “best seat in the house position” served as a reference for the measurements.

In order to be able to evaluate relationships of the surround signals it was decided to use the dummy-head recording as a kind of “human measure” reference. The surround-microphone technique, which would be able to produce the highest correlation (regarding the entire frequency range) between the original dummy-head recording and the dummy-head re-recording of the surround-microphone replay in the listening room would appear to be the one with the highest fidelity.

### 10.1. Measurement of the Frequency-dependent Correlation-Coefficient (FCC) and Frequency-dependent Inter-Aural Correlation-Coefficient (FIACC):

In order to be able to measure the frequency dependent Correlation Coefficient, it was necessary to program an appropriate tool (function) within MATLAB-software. The two channel-signals of interest were split into 31 independent frequency bands (with center-frequencies according to ISO), which were then subject to correlation analysis.

In addition the correlation of the low-frequency signal content (below 400Hz; important in respect to spaciousness – see section 3.2) and high-frequency signal-content (above 1200Hz) was measured.

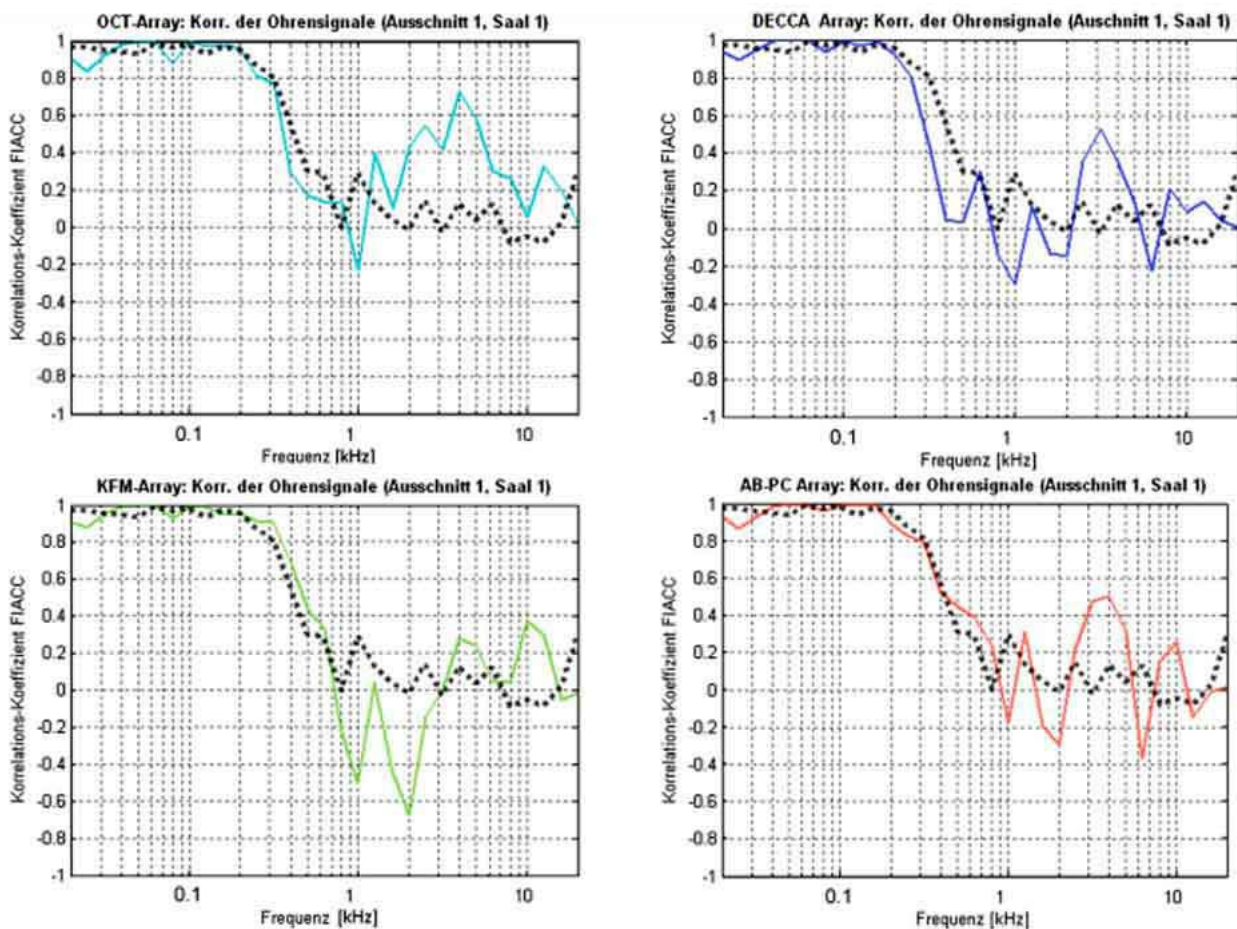


**Fig. 16:** FIACC D.U.T. (solid line), KU81(dotted line) (sample #1, hall 1)

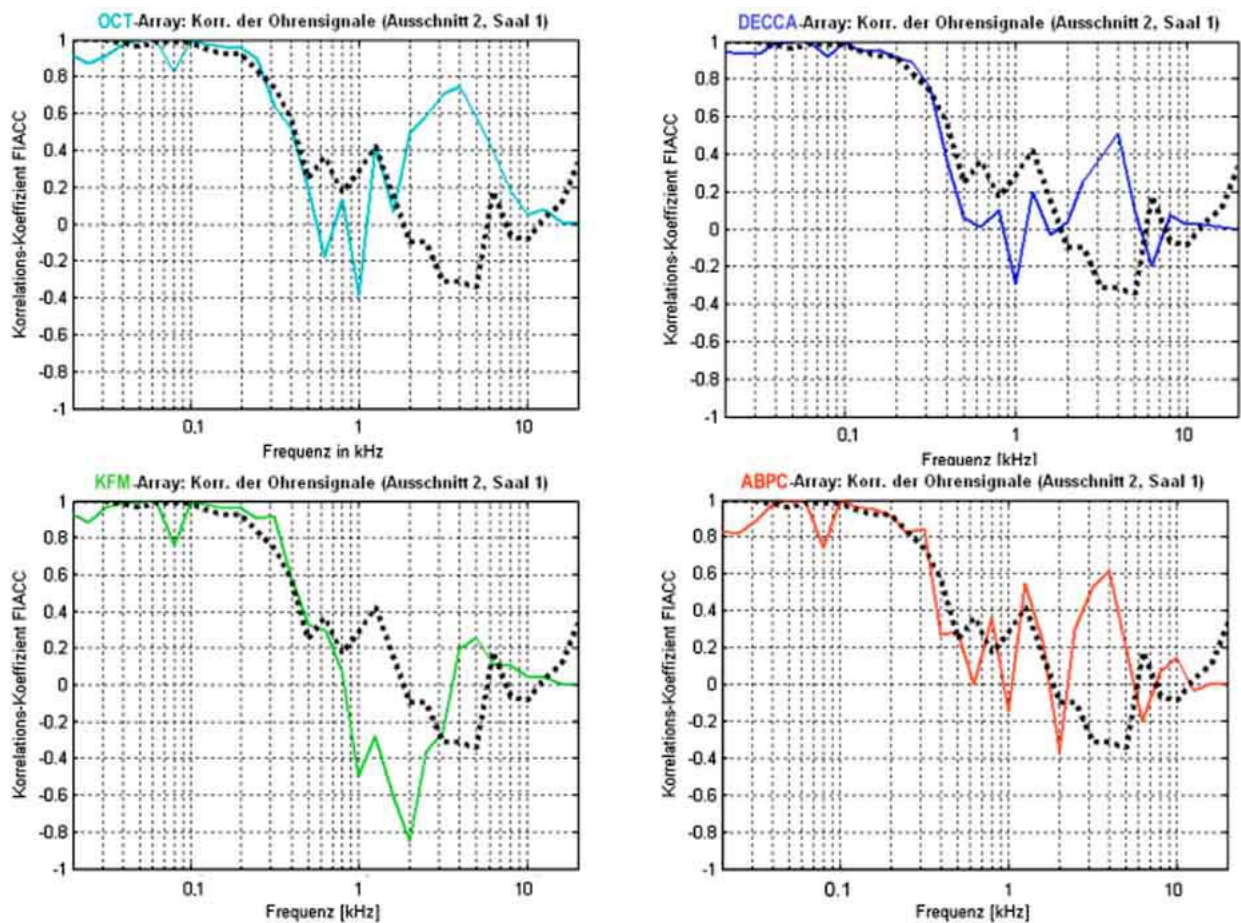


**Fig. 17:** FIACC: D.U.T. (solid line), KU81 (dotted line) (sample #2, hall 1)

For the measurements appropriate sections from the audio material recorded for the listening tests was selected:

ORCHestral recording (hall 1):

- sample #1, 7 sec (forte)
- sample #2, 16sec (piano)
- sample # 4, 60 sec (piano to forte)

DUO recording (hall 2):

- sample #3, 16sec

In figures 16-19 the FIACC's of the original dummy head recordings (KU81) from the concert hall (dotted line) are compared with the dummy-head re-recording of the selected surround-microphone (solid line).

In Fig. 16 it is interesting to note that the microphone-techniques which employ (also) omni-directional microphones (i.e. DECCA, ABPC) are both characterized by moderately pronounced correlation (max. +0.5) in the frequency range around 2-3kHz.

As was already shown by Hiyama, Komiyama, Hamasaki [62] this is a simple consequence of the standard 5.1 loudspeaker-layout, in interaction with the HRTF's of a (dummy) head. (see figure 18 on next page; functions for the 30°(R) and 120°(RS) loudspeakers)

More evidence on this and a calculation of the resulting filter, based on a spherical head-model can be found in a more recent study [63].

In comparison the OCT-system is characterized by a much higher correlation (max. +0.7) in the same frequency range, while the KFM displays prominent *negative* correlation (max. -0.7) in the frequency range from 700Hz up to around 3kHz.

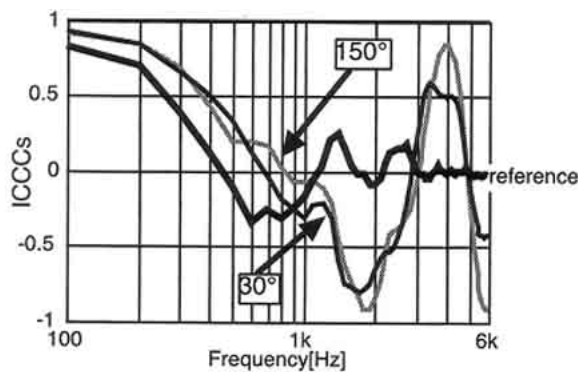


**Fig. 18:** ICCs (=FIACCs) for symmetrical arrangement of a pair of loudspeakers at indicated angles (i.e. at  $\pm 30^\circ$  and  $\pm 150^\circ$ ) (adapted from [62]).

With the 16sec piano orchestra sample #2, it is interesting to see that peak of correlation at the OCT-system has become more pronounced, and so has the negative correlation of the KFM-system (which ranges now from 800Hz to 3.5kHz, with a peak of less than  $-0.8$ ).

In fig. 19 a much longer sample of 60sec duration was analyzed and provided quite astonishing results: while correlation has raised over the entire frequency range for the OCT, DECCA and KFM system, the ABPC system is the only one which still 'wavers' around the non-correlation '0'-axis in relative proximity to the reference (original KU81 dummy head recording from the concert hall, which now displays a much smoother curve, due to the higher sample length).

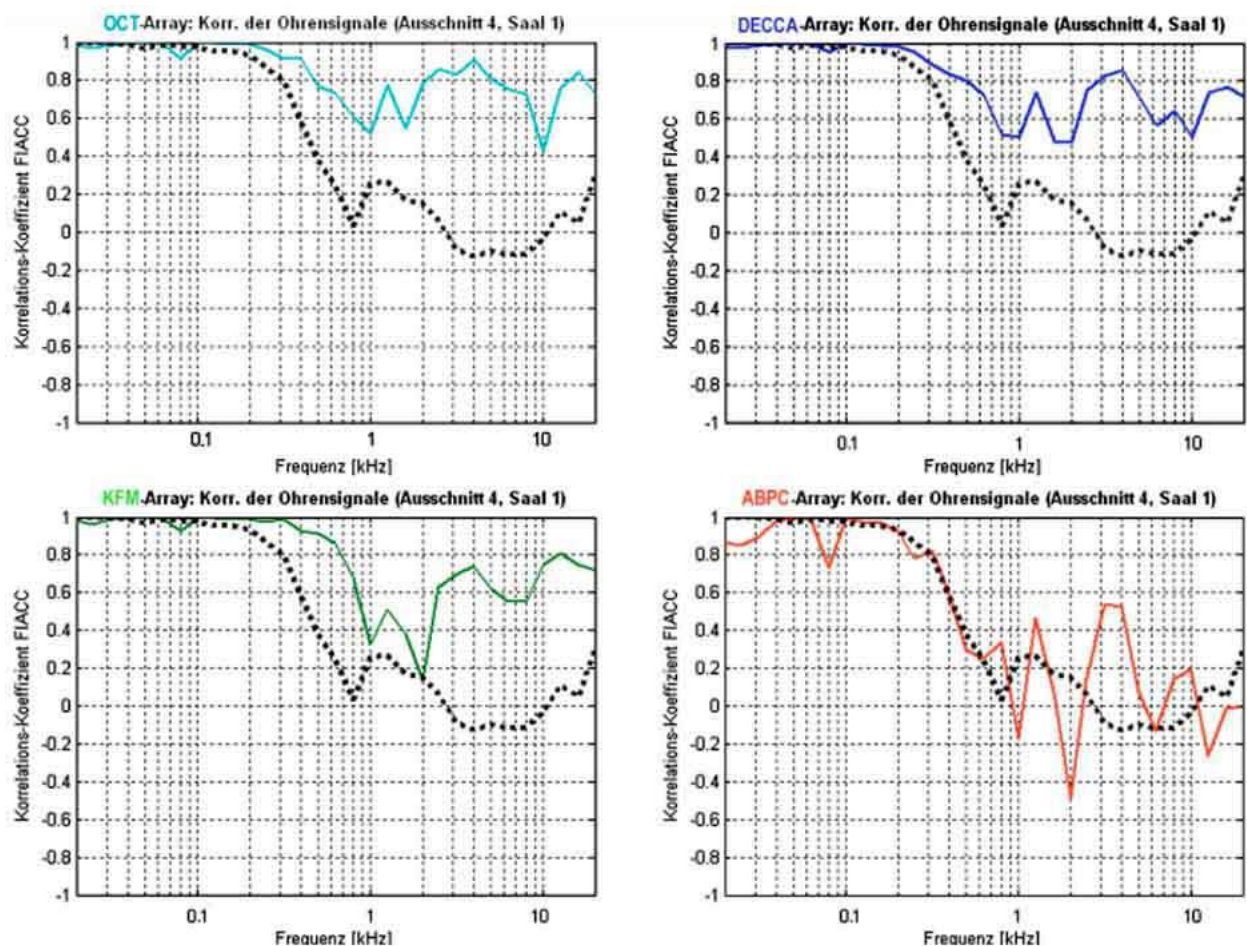


**Fig. 19:** FIACC: D.U.T. (solid line), KU81 (dotted line) (sample # 4, hall 1)

Also, it is extremely interesting to note that the ABPC-system is the only one which follows the KU81-reference function almost exactly below approx. 700Hz (which is the frequency range above which the human or dummy head starts to get acoustically effective as a baffle for higher frequencies).

According to the FIACC-measurements the ABPC-system is therefore the device under test (D.U.T.), which manages to approximate the original sound-field much better than its competitors.

Table 2 on the next page shows the numeric values of the correlation coefficients for the four microphone systems and the three samples: apart from the fact that the ABPC-system is usually closest to the KU81 reference in terms of correlation, it is most interesting to note that it does so -almost always – for the LF-part as well as for the HF- part of its signal (and not only for

the overALL correlation). As can be seen from the values of the ‘ALL’ correlation coefficients, (forte) sample #1 is characterized by much lower overall correlation than (piano) sample#2, which is most likely due to the much higher amount of diffuse sound or reverb (stochastic signal) in the hall, due to the room being excited by the high-level music. With the ‘higher correlation’ (piano) sample #2, the ABPC-system ‘follows foot’ and hence also displays higher overall correlation.

So the ABPC system is not characterized by higher or lower correlation ‘per se’, it simply seems to manage to capture the original soundfield in the hall in a physically more correct way, as can be deduced from the results.

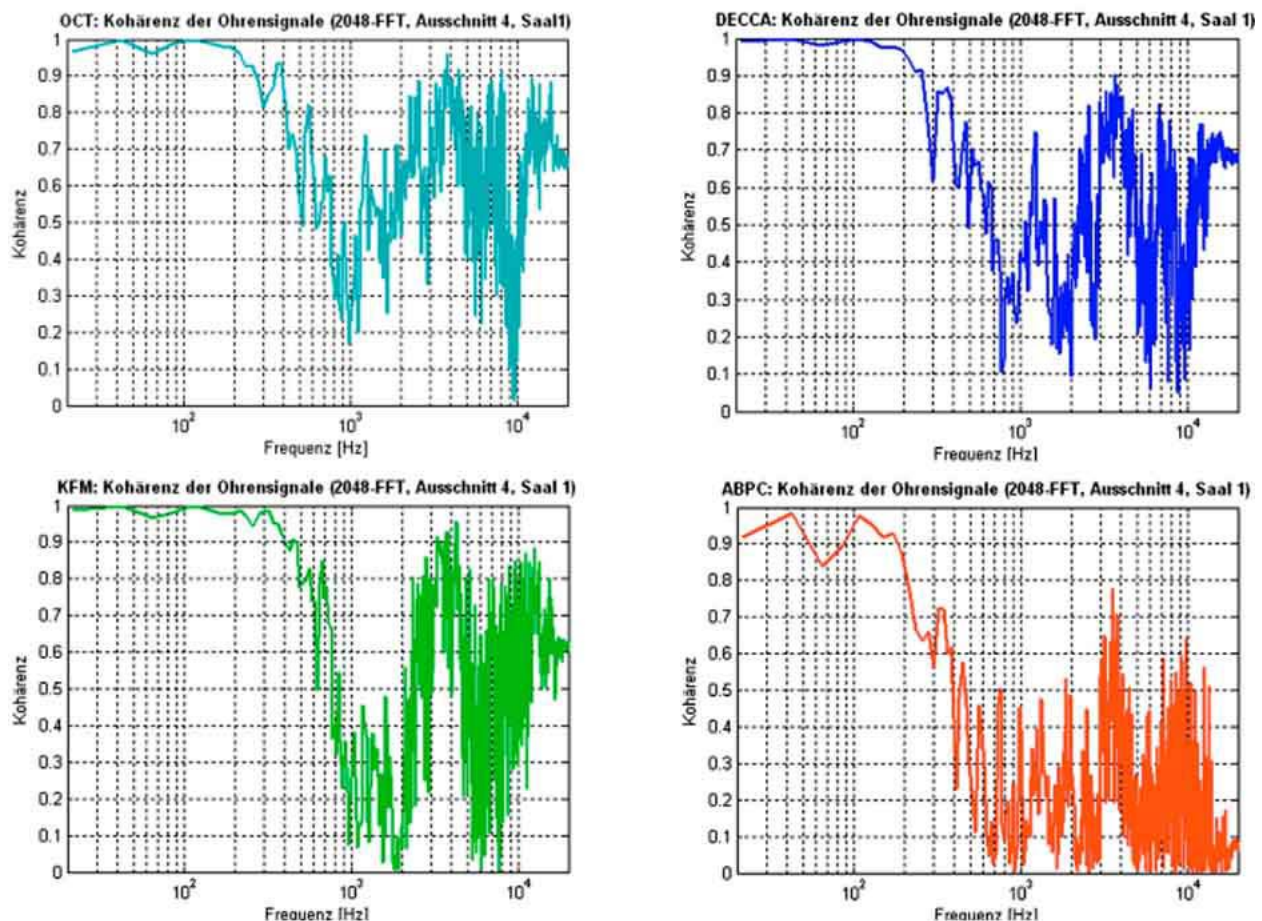


Fig. 20: Estimated coherence for the D.U.T.'s

Corr.-Coeff.: LF(400Hz) HF(1200Hz) ALL

**ORCH-sample 1: (*forte*, 7sec)**

|             |             |             |             |
|-------------|-------------|-------------|-------------|
| OCT         | 0.87        | 0.28        | 0.41        |
| DECCA       | 0.85        | -0.01       | 0.23        |
| KFM         | 0.88        | -0.30       | 0.39        |
| AB-PC       | 0.89        | 0.15        | 0.48        |
| <b>KU81</b> | <b>0.92</b> | <b>0.11</b> | <b>0.55</b> |

**ORCH-sample 2: (*piano*, 16sec)**

|             |             |             |             |
|-------------|-------------|-------------|-------------|
| OCT         | 0.94        | 0.36        | 0.85        |
| DECCA       | 0.96        | 0.11        | 0.89        |
| KFM         | 0.98        | -0.46       | 0.88        |
| AB-PC       | 0.92        | 0.37        | 0.78        |
| <b>KU81</b> | <b>0.85</b> | <b>0.24</b> | <b>0.74</b> |

**ORCH-sample 4: (*piano + forte*, 60sec)**

|             |             |             |             |
|-------------|-------------|-------------|-------------|
| OCT         | 0.98        | 0.71        | 0.88        |
| DECCA       | 0.98        | 0.66        | 0.90        |
| KFM         | 0.99        | 0.44        | 0.87        |
| AB-PC       | 0.88        | 0.27        | 0.54        |
| <b>KU81</b> | <b>0.87</b> | <b>0.20</b> | <b>0.57</b> |

**Table 2:** Correlation-Coefficients for three samples and four surround-microphone systems under test

## 10.2. Measurement of Binaural Signal Coherence:

Some researchers are convinced that it is preferable to measure signal-coherence instead of signal correlation [ 35]. The most basic differences between correlation and coherence have been outlined above in section 4.

For the measurements carried out in this research the ‘mscohere’-function in MATLAB was used, which finds the magnitude squared coherence estimate of the input signals. As this signal processing algorithm is based on FFT, an adequate ‘computational resolution’ has to be chosen in order to achieve sufficient resolution in the frequency domain. With the 48kHz samples an appropriate ‘FFT-point’ resolution was chosen to achieve approx. 20Hz spacing for the frequencies at which the coherence would be calculated, in order to obtain reliable results also for the lowest-frequency bands.

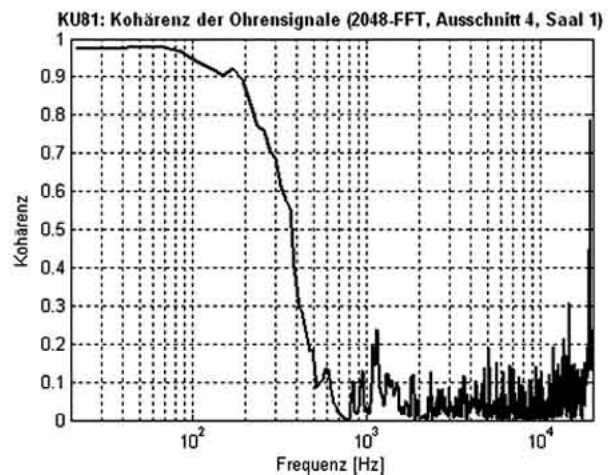


**Fig.21:** Coherence for ‘ear-signals’ of the KU81 original recording (sample #4, hall 1)

Looking at fig. 21 the graph evokes the impression of a much higher accuracy than with the (band-filter bank based) correlation-algorithm. This impression is not necessarily right, as one should bear in mind, that:

- 1) the computation of the signal coherence (- however accurate it may be -) is only an *estimation*,
- 2) the FFT-algorithm calculates coherence at many points along the frequency range of interest, but no matter how high the resolution, there will always be way more frequencies for which the FFT-algorithm did not (or better: could not) calculate coherence, than frequencies for which he did ...

Therefore: while the correlation coefficient gets really *measured*, the coherence gets only *estimated*; and the sometimes quite high spikes along the coherence-‘curve’ within an otherwise relatively ‘calm’ part of the curve shows that there can be vast differences in (calculated/estimated) coherence in a rather small frequency range.

Apart from these ‘system-inherent’ properties of coherence estimation the main point of criticism however becomes evident when taking a closer look at the results the coherence estimation provides for the surround-microphone systems under test: (see Fig. 20 on previous page)

Examining the coherence-functions of the microphone systems it is evident, that they are all characterized by higher coherence in the frequency band around 3-4kHz (this is very prominent for the OCT, DECCA and KFM system, much less prominent for the ABPC). Also the KFM exhibits a relevant 'notch' in terms of coherence mainly in the frequency range of 1-2kHz, which must be related to the relevant phase-changes in that frequency range, which were so evident with the previously performed correlation-analysis.

However, this is – in comparison - only a very 'faint echo' of what was to be seen clearly with correlation analysis before ...

Also with coherence analysis, the appearance of the ABPC system is quite different from the three competitive systems, nevertheless it seems clear to the authors that measuring signal correlation is superior to measuring signal coherence if a detailed analysis of the D.U.T. is desired.

The strong out-of-phase signal components that were found with the KFM-system in a frequency range (1-3kHz) which is of primary importance in relation to human hearing, are most likely responsible for the bad sound-color ratings, which this system has received throughout the listening tests. An in-depth view on the problem of sound-coloration due to comb-filtering upon loudspeaker signal summation can be found in a recent publication dealing with Wave-Field Synthesis [2].

### 10.3. Measurement of Surround Signal Coherence:

Even though correlation and coherence analysis of the (binaural) dummy-head signals has provided quite meaningful results, which make it possible to clearly distinguish the reproduction accuracy of the surround-microphones under test, it seems nevertheless worthwhile to also examine the signal-relationships

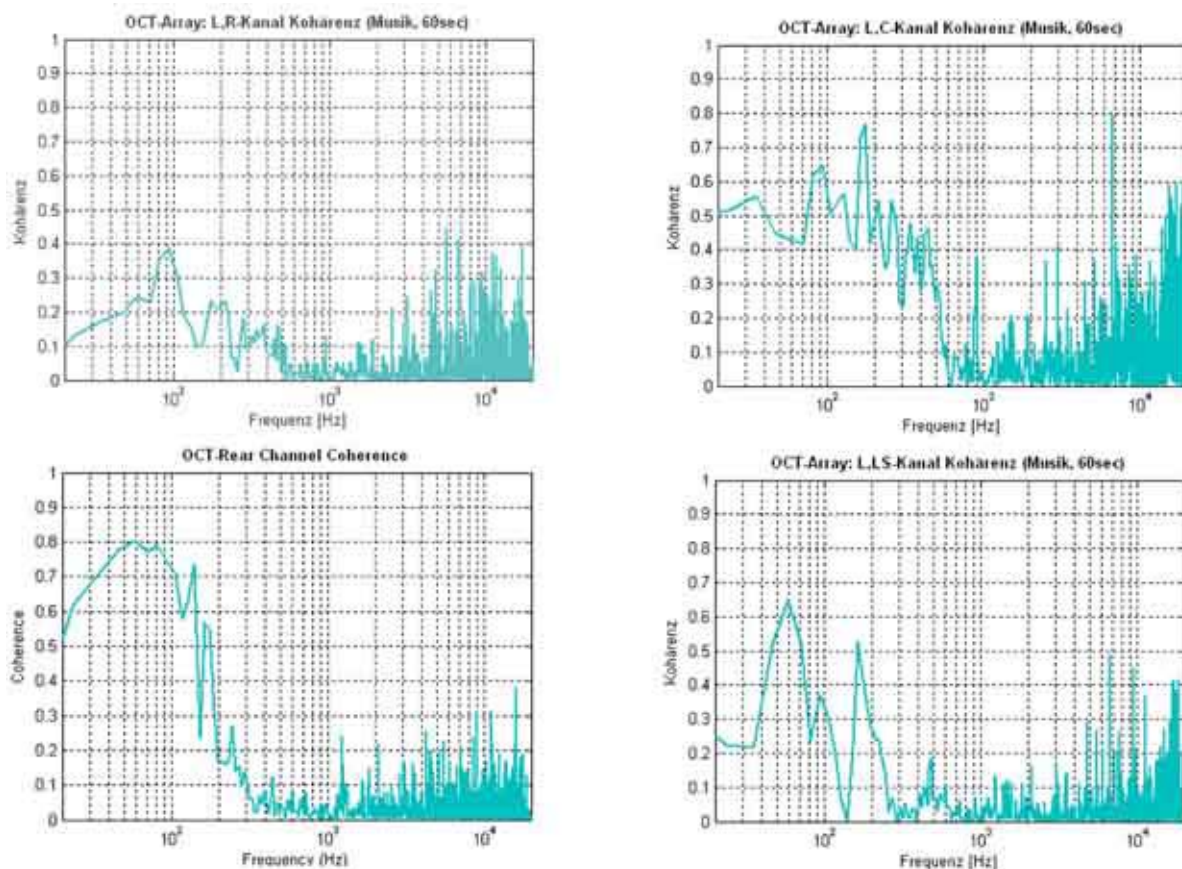


**Fig. 22:** OCT Surr.-Channel Coherence:  
(from L to R, up to down): L/R, L/C, LS/RS, L/LS

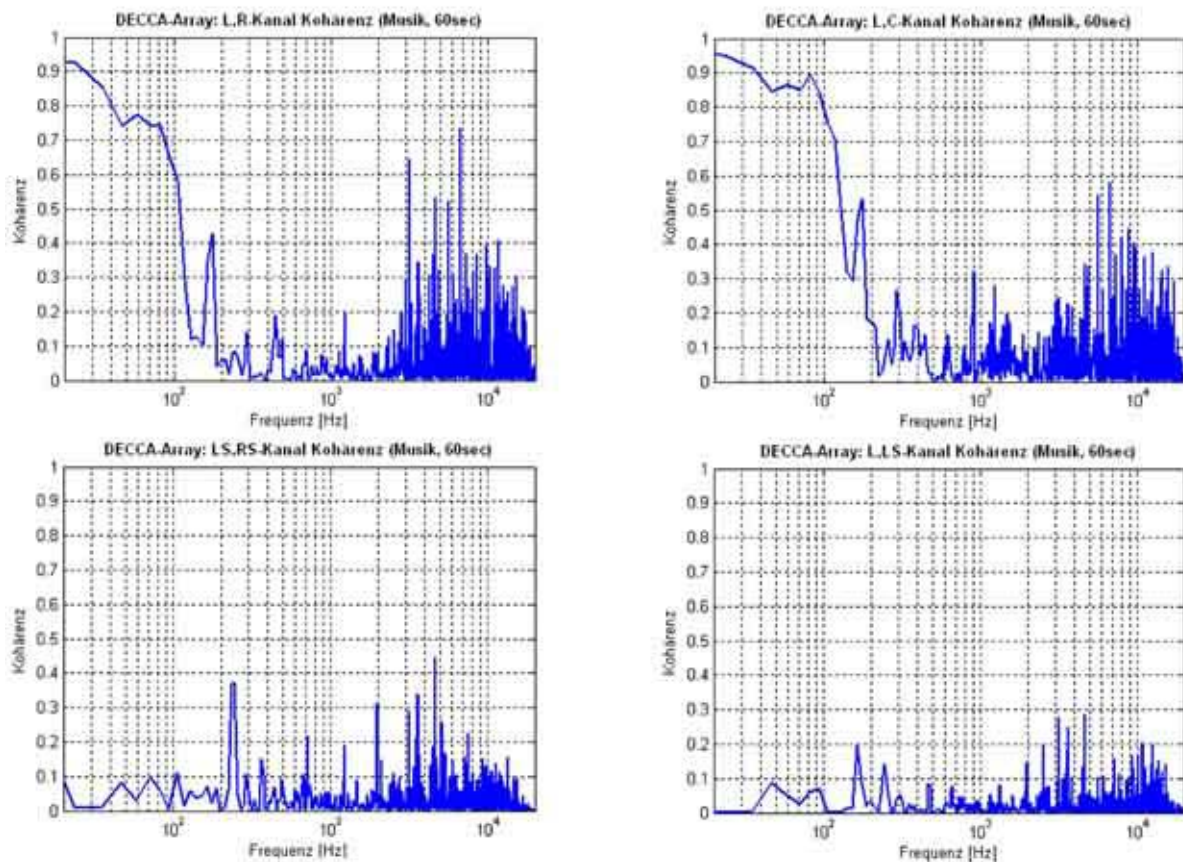


**Fig. 23:** DECCA Surr.-Channel Coherence:  
(from L to R, up to down): L/R, L/C, LS/RS, L/S

between each of the 5 surround signals in order to gain a better understanding.

Examining the OCT surround-system in Fig. 22 it becomes clear that coherence for the L/R-front channels is rather low for the entire frequency range, as can be expected from the two hypercardioids that are pointing side-wise. However coherence is reasonably high between the front-channels and the center-channel below 500Hz with peaks reaching values of 0.75. Interestingly there seem to be some ‘spurious’ very high peaks in the frequency region near 10kHz which reach values of 0.8.

Coherence of the rear-channel LS/RS is very low for most of the frequency range, but below 200Hz there is a steep rise which totals in high coherence of 0.8. Also the coherence between the front L and rear LS channel experiences a relevant rise below 200Hz, which means that in this frequency range signal coherence is overall significantly high for the OCT-system. This high

coherence in the low frequency range is to the disadvantage of good spatial reproduction, as we already know from [28] and [29].

As can be expected, low-frequency coherence is much higher with the DECCA-system, at least between the three front-channel, which employ omnidirectional microphones. Again a steep rise below 200Hz (similar capsule spacing as with the L and R capsule of the OCT-system), but to even higher values of close to 0.9. Due to the fact that the backwards-pointing cardioid microphones for the rear channels are positioned quite far away (approx. 10m), there is almost no coherence between them, as well as to the front channels.

This enables good spatial impression at least between the front and rear channels.

The KFM-system in Fig. 24 on the other hand is characterized by much higher coherence between all channels, mainly at low frequencies, but not only:

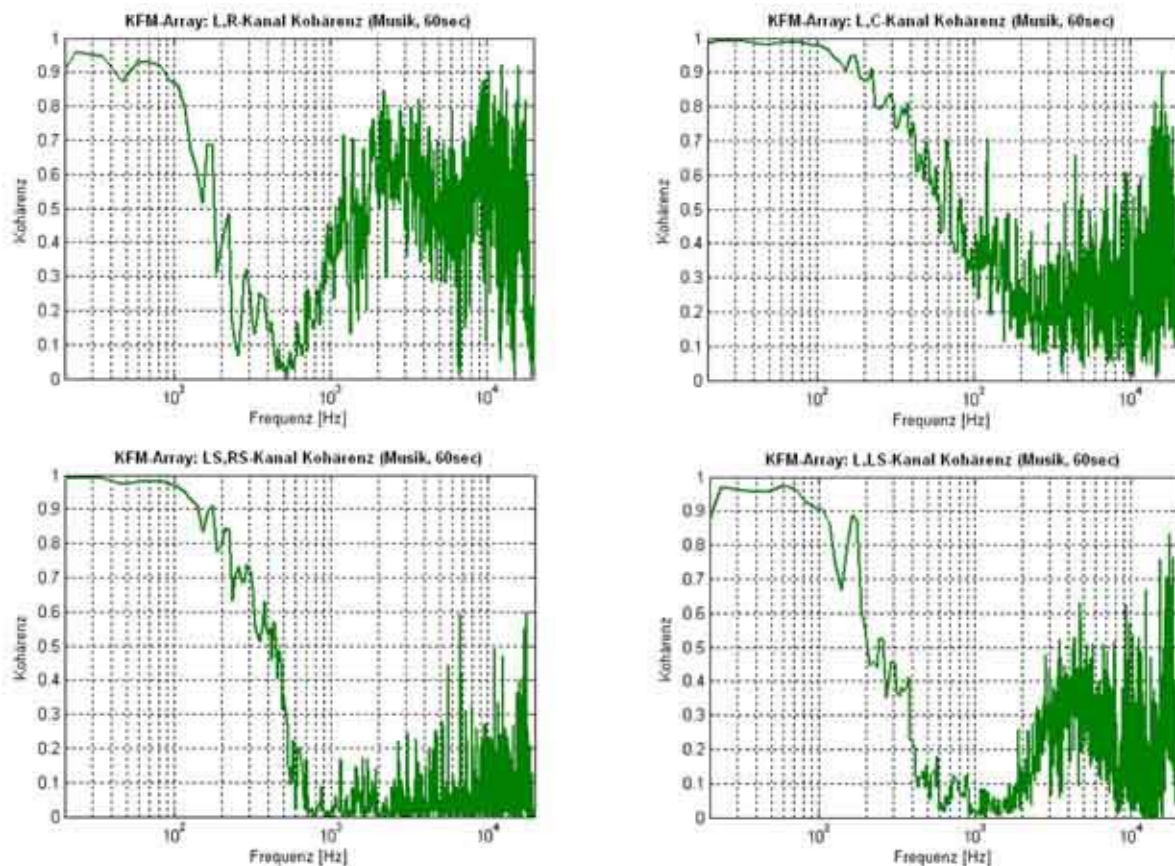


**Fig. 23:** KFM Surr.-Channel Coherence:  
(from L to R, up to down): L/R, L/C, LS/RS, L/LS

It can be seen that L/R coherence is almost down to 0 around 500Hz and then has a steep rise again towards the high frequencies, where max. values of up to 0.9 are being reached. Also coherence between L and C channel is very high not only for low frequencies. For the rear channels LS/RS high coherence is limited to the low-frequency band up to approx. 500Hz. Coherence between L and LS is very high as well for low frequencies, almost 0 from 400Hz – about 2kHz and rises again towards high frequencies.

The ABPC-system in comparisons shows a very different behavior: as can be seen in Fig.24 on the next page, coherence is very low in the low-frequency band. For the L/R front-channels there is a slight rise towards

high frequencies, but coherence stays usually below 0.5 up until almost 10kHz. Also coherence between the L and C channel is rather low: it reaches a high with about 0.5 around 400-500Hz, but then dies away again towards higher frequencies.

The signals of the rear-channels and between the rear and front channels are almost incoherent (apart from a few 'spikes'). The almost complete lack of coherence (i.e. high de-correlation) at lowest frequencies up to about 400Hz is also the reason for the good spatial impression this microphone system provides.

For completeness the numerical values for the coherence between the surround-channels are listed in table 3.

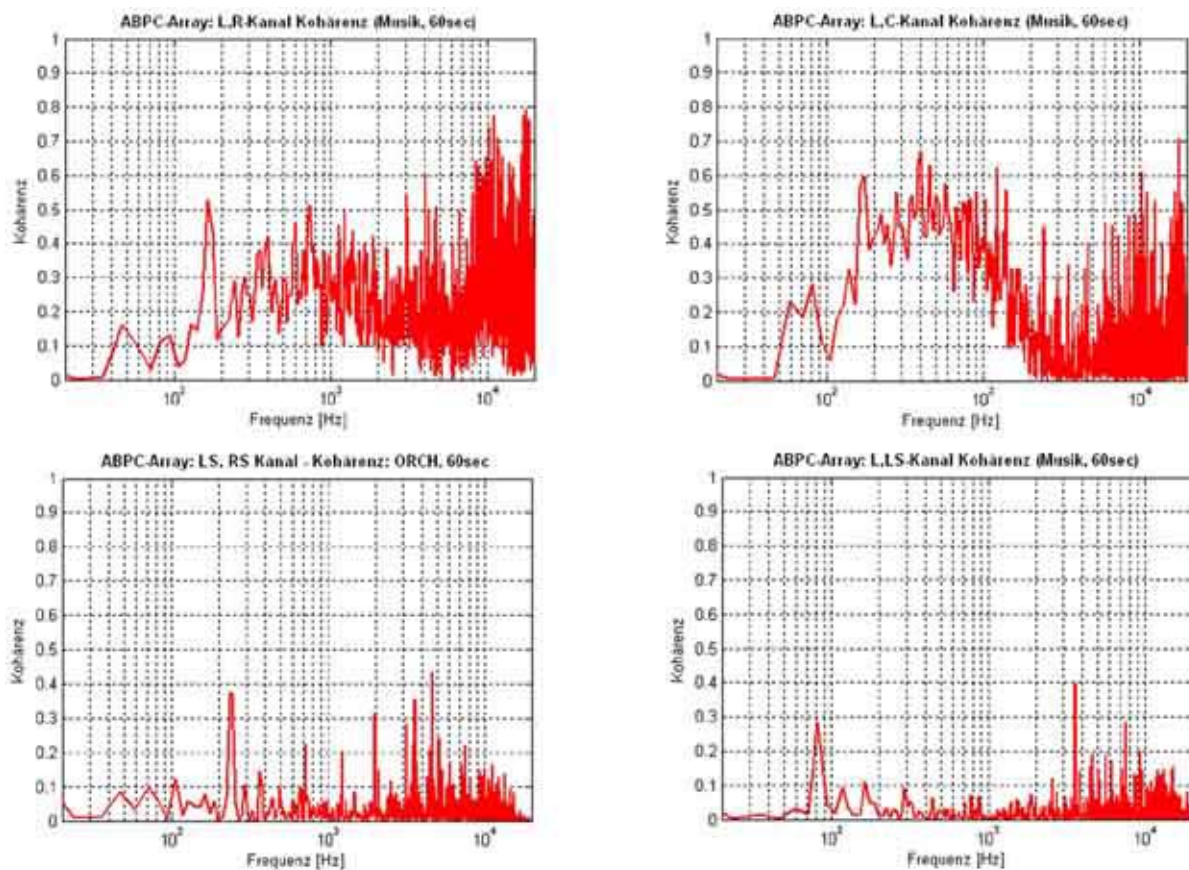


**Fig.24:** ABPC Surr.-Channel Coherence:  
(from L to R, up to down): L/R, L/C, LS/RS, L/LS

| ORCH -<br>Rec. | Binaural<br>(Dummy) | Front<br>L-R | L-C  | C-R  | L-LS | R-RS | Rear<br>LS-RS |
|----------------|---------------------|--------------|------|------|------|------|---------------|
| OCT            | 0.74                | 0.09         | 0.24 | 0.31 | 0.13 | 0.13 | 0.20          |
| DECCA          | 0.68                | 0.19         | 0.24 | 0.24 | 0.03 | 0.02 | 0.05          |
| KFM            | 0.70                | 0.49         | 0.55 | 0.45 | 0.37 | 0.32 | 0.38          |
| AB-PC          | 0.44                | 0.21         | 0.24 | 0.25 | 0.03 | 0.03 | 0.04          |
| KU81           | 0.37                |              |      |      |      |      |               |
|                |                     |              |      |      |      |      |               |
| DUO-<br>Rec.   | Binaural<br>(Dummy) | Front<br>L-R | L-C  | C-R  | L-LS | R-RS | Rear<br>LS-RS |
| OCT            |                     | 0.16         | 0.24 | 0.36 | 0.21 | 0.19 | 0.22          |
| DECCA          |                     | 0.18         | 0.28 | 0.29 | 0.08 | 0.10 | 0.10          |
| KFM            |                     | 0.54         | 0.60 | 0.53 | 0.36 | 0.37 | 0.38          |
| BPT            |                     | 0.24         | 0.42 | 0.44 | 0.08 | 0.09 | 0.10          |

**Table 3:** Coherence between var. channels of the four surround-microphones under test

Also from the numeric values of coherence for the binaural dummy head (re-)recordings it can be seen that the AB-PC system is closest to the reference recording (KU81) from the concert hall. (as was the case with correlation)

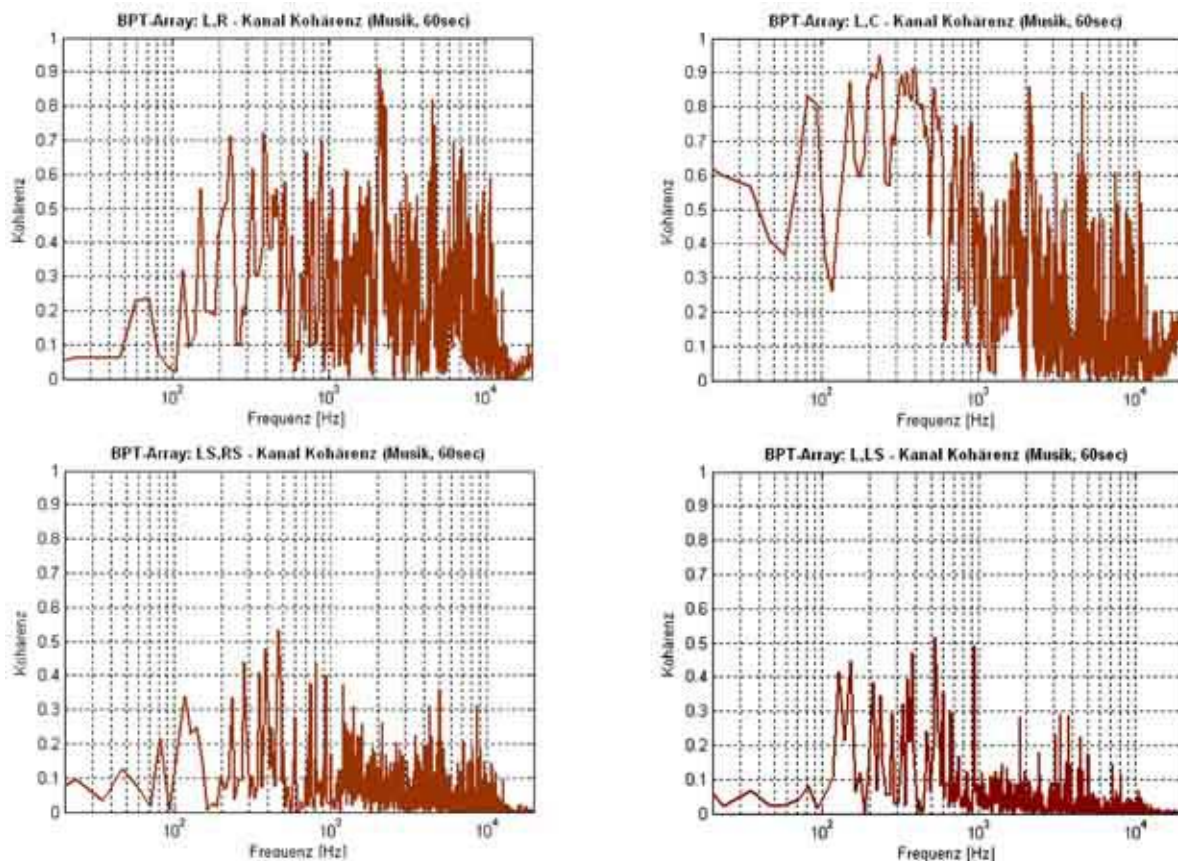


**Fig.25:** BPT Surr.-Channel Coherence:  
(from L to R, up to down): L/R, L/C, LS/RS, L/LS

In Fig. 25 the measurements of the coherence concerning the BPT-surround microphone are documented:

(DUO recording, music signal, 60sec, hall 2):

up to about 100Hz signal coherence is very low, then it rises and from about 200Hz upwards coherence varies between approx. 0.1 (min) to 0.6 (max) with pronounced peaks at 2kHz ( $k=0.9$ ), 4.7kHz and peaks between 6 and 7 kHz.

Due to the strong directional characteristics of the fig.-of-eight capsules coherence between the L and C-channel is quite high, starting already at low frequencies: up until about 200Hz it 'wavers' around a moderate  $k=0.6$ , above 900Hz it drops below 0.5 as an average. However the peaks around 2 kHz, 4.7kHz and around 6kHz and slightly above 10kHz appear also for this combination of capsules and therefore concerns the whole 3-channel front-array.

Due to the widely spaced cardioids which are

pointing towards the rear and are at a distance of more than the reverberation radius from the front system, signal coherence between the rear channels LS/RS and to the front channel (LS-L) is very low. Even if the relatively high coherence at specific frequencies in the mid- and high-frequency bands is not necessarily (psycho-acoustically) desirable, the overall low coherence in the lowest-frequency band below 200Hz) ensures that the BPT microphone is well suited for good spatial reproduction.

## 11. FREQUENCY DEPENDENT SIGNAL CORRELATION (FCC) IN STEREO-MICROPHONE SYSTEMS:

A very concise description of the most important (2- and more- channel) stereo-microphone systems can be found in [64]. Therefore below we present the measurements for some of the most commonly used stereo main-microphone systems, with occasional commentary only.

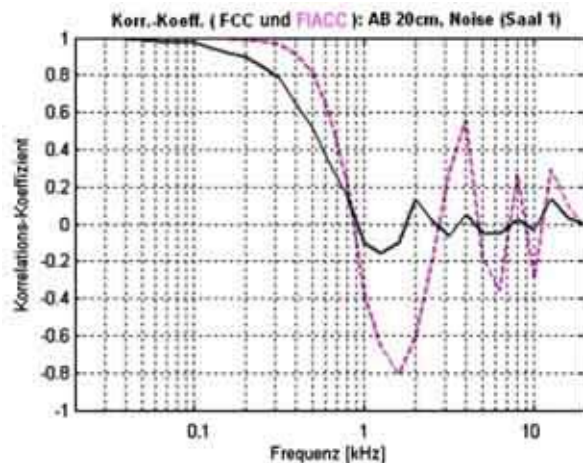


**Fig.26:** Small AB 20cm, FCC (solid line) and FIACC (signal: noise, hall 1)

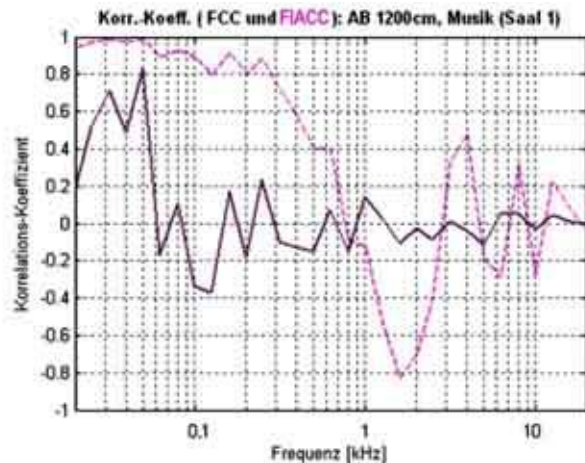


**Fig.27:** Large AB 1200cm, FCC and FIACC (signal: music, hall 1)

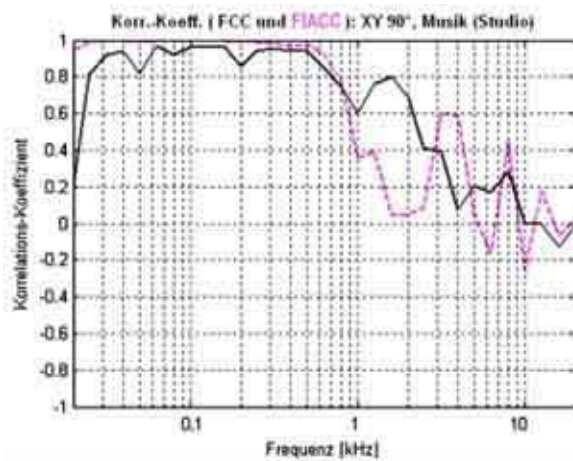


**Fig.28:** XY 90° FCC and FIACC (signal: music, studio)

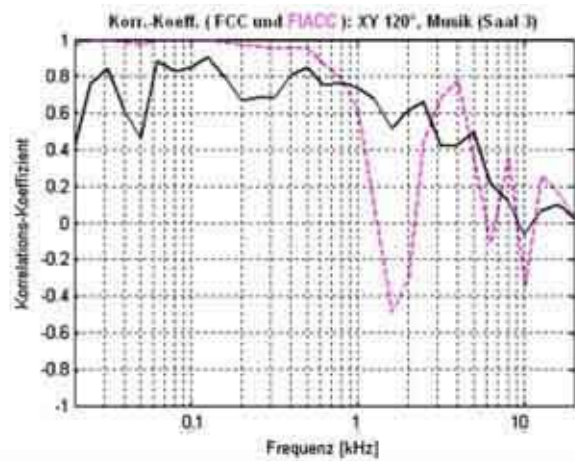


**Fig.29:** XY 120° FCC and FIACC (signal: music, hall3)

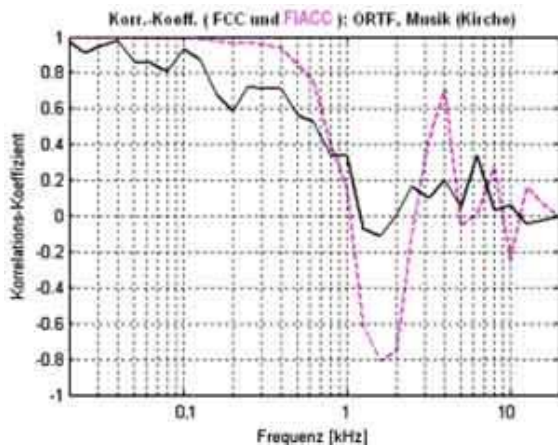


**Fig.28:** ORTF FCC and FIACC  
(signal: music, church)

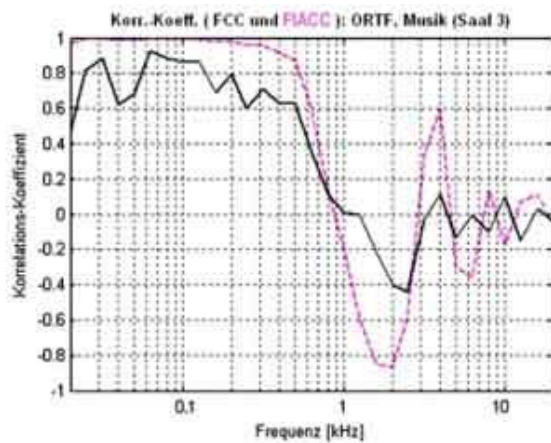


**Fig.29:** ORTF FCC and FIACC  
(signal: music, hall3)

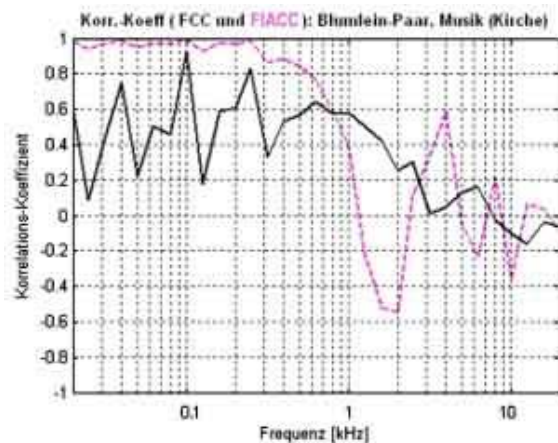


**Fig.30:** Blumlein-Pair FCC and FIACC  
(signal: music, church)

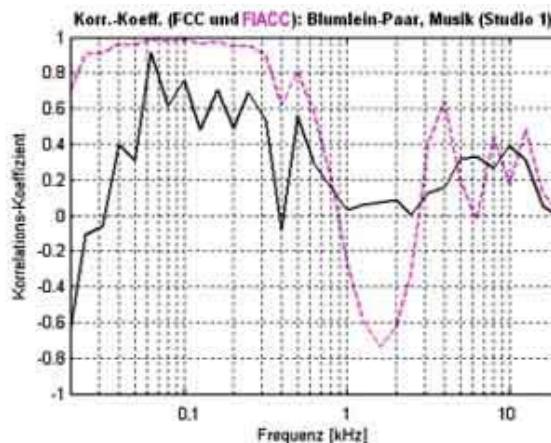


**Fig.31:** Blumlein-Pair FCC and FIACC  
(signal: music, studio 1)

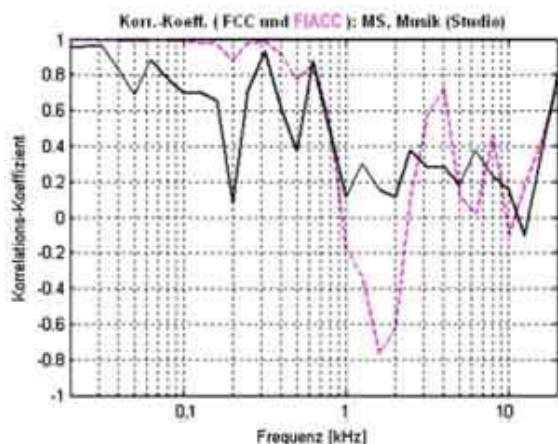


**Fig. 32:** MS FCC and FIACC  
(signal: music, studio)

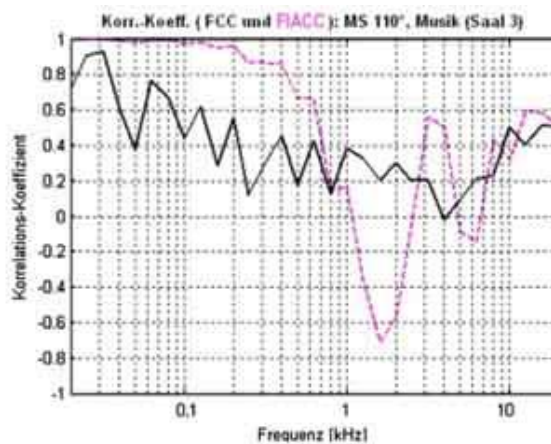


**Fig. 33:** MS 110° FCC and FIACC  
(signal: music, hall 3)

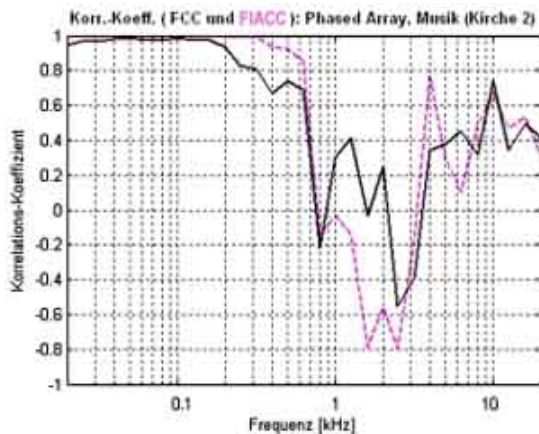


**Fig. 34:** Faulkner “Phased Array” FCC and FIACC  
(signal: music, church)

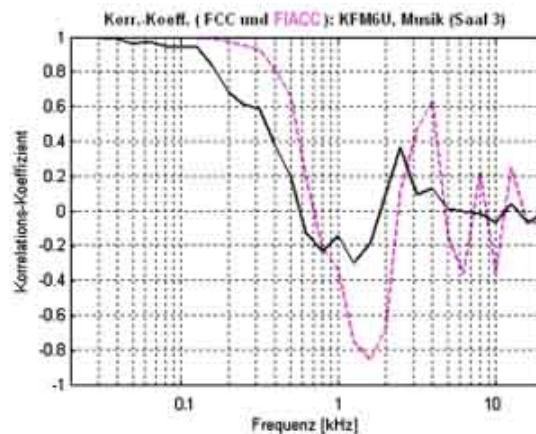


**Fig. 35:** KFM6U FCC and FIACC  
(signal: music, hall 3)

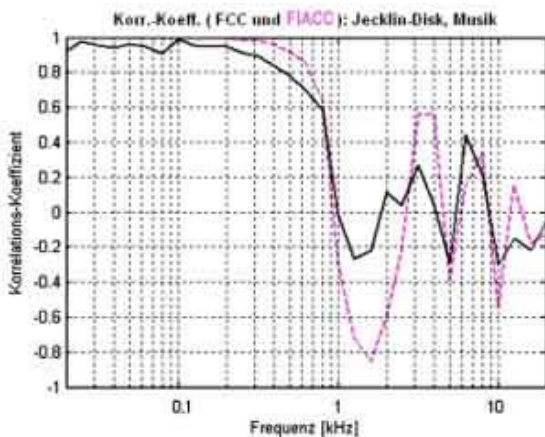


**Fig. 36:** Jecklin-Disk (OSS), FCC and FIACC  
(signal: music, hall)

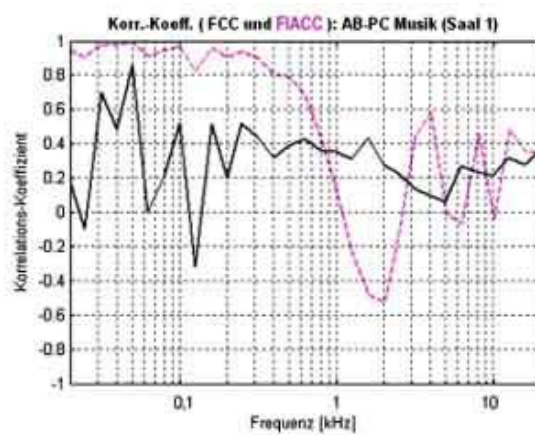


**Fig. 37:** AB-PC (with ORTF-Triple) FCC + FIACC  
(signal: music, hall 1)

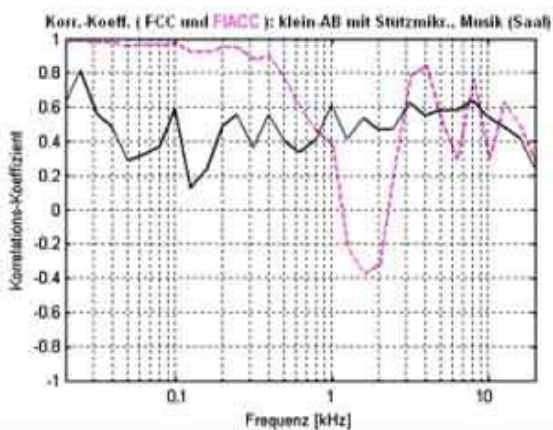


**Fig. 38:** small AB + spot-mics, FCC and FIACC  
(signal: music, hall) TELDEC

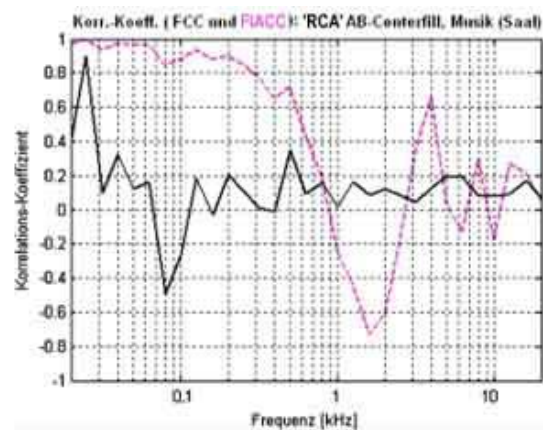


**Fig. 39:** AB-Centerfill, FCC and FIACC  
(signal: music, hall) RCA “Living Stereo”

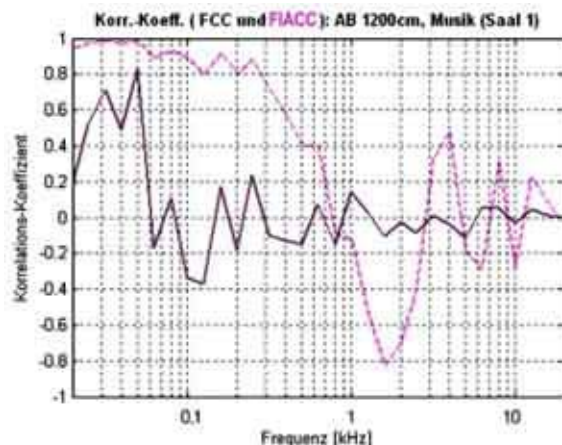


**Fig. 40:** Large AB 1200cm, FCC and FIACC  
(signal: music, hall 1)

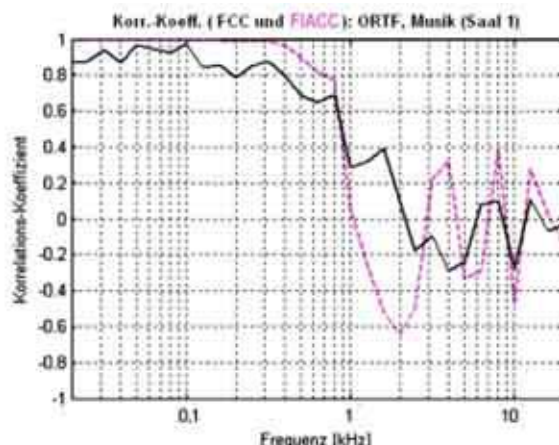


**Fig. 41:** ORTF, FCC and FIACC  
(signal: music, hall 1)

In Fig. 26 – 41 the solid line represents the frequency dependent correlation coefficient FCC which is obtained when measuring the electrical signal, which is derived from the two capsules of a stereo microphone (or from the recording of such). As it may also be of interest to see which frequency dependent inter-aural correlation coefficient FIACC (dotted line) would result at a human or dummy head when replaying the signal of a stereo-microphone through the standard L/R stereo loudspeaker setup, the microphones signals were processed using a software-plugin which can transform a stereo-signal into an appropriate binaural signal. This plugin is based on the HRTFs of the MIT “Kemar” dummy-head. The resulting binaural signal was analyzed using to the same MATLAB –function, as for the measurement of the FCC.

If available, two different recordings with the same microphone technique (but in different locations) were analyzed in order to get a more balanced impression about their ‘real world’ behavior. Many of the microphone techniques which were tested are characterized by rather high signal correlation at low frequencies (below approx. 500Hz), which impedes good spatial impression.

Among these are:  
Small AB, XY, ORTF, Faulkner Phased Array, KFM, and Jecklin Disk

What should a ‘healthy’ FCC signal look like ?

In order to enable good spatial impression it should be de-correlated at low frequencies. In addition, it should preferably also display low correlation at frequencies above 1 kHz, as higher correlation in the upper frequency band in the electrical (replay) signal will also lead to high correlation at the FIACC, which is not desirable, as we have already seen when examining the surround recordings.

## 12. DISCUSSION AND CONCLUSION:

As we know from results presented in [4] and [12] small AB ‘time-of-arrival-difference’ based recordings are not able to provide reliable localization. In addition the small capsule spacing leads to high correlation at low frequencies, which is to the disadvantage for achieving good spatial impression. Research carried out in this study has provided similar results: a surround microphone technique based on ‘psycho-acoustic’ principles of small time of arrival differences (“DECCA”) has suffered more severe degradation when being listened to under non-ideal conditions (binaural headphone replay), than the competing systems.

Four surround microphone techniques (OCT, DECCA, KFM, AB-PC) have been analyzed in terms of frequency dependent correlation, as well as frequency dependent coherence. It was shown that coherence is not able to unveil potential weaknesses of a microphone system (e.g. in respect to phase anomalies; see KFM), as clearly as correlation.

Therefore the authors would like to encourage the use of FCC and FIACC measurements, when appropriate (instead of measuring coherence).

An assumption (or postulate) made by an important opinion-leader of the audio-community, (i.e. that optimal signal coherence should have a value of at least  $k \geq 0.35$  in order to ensure "...natural imaging of space and envelopment..."[36]) was proved wrong: not only that one of the more recent surround microphone techniques of his (OCT-Surround) does not reach this proposed value, but in fact the only surround microphone system, which does achieve such a value (KFM) was clearly downgraded by two different groups of test listeners in double-blind listening tests (see Table 3 and Fig. 11 and 12).

All this seems to indicate that microphone systems with high signal correlation should probably better be abandoned for the sake of systems that are largely de-correlated. As was found in research [65] and can be seen in Fig. 42, low correlation of the playback signals (L/R vs. C vs. LS/RS) ensures a better compatibility between different listening environments and also leads to an enlarged 'sweet spot' [66].

Therefore microphone techniques should be used which are able to provide low signal-correlation over the entire frequency range, if possible.

Among the four surround microphone systems (OCT, DECCA, KFM, ABPC) which were evaluated via subjective listening tests, as well as objective measurements (FCC and FIACC) the low-correlation AB-PC system turned out most successful: in terms of listener preference and as it has achieved several 'best mean value' ratings for various acoustic attributes.

In addition it was the microphone system which managed to physically approximate the original sound-field of the concert hall better than any of the other systems under test.

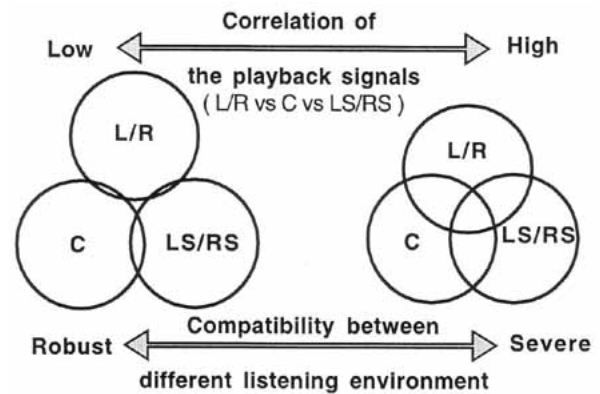


**Fig. 42:** Amount of signal correlation in surround playback systems and compatibility between different listening environments (from [65])

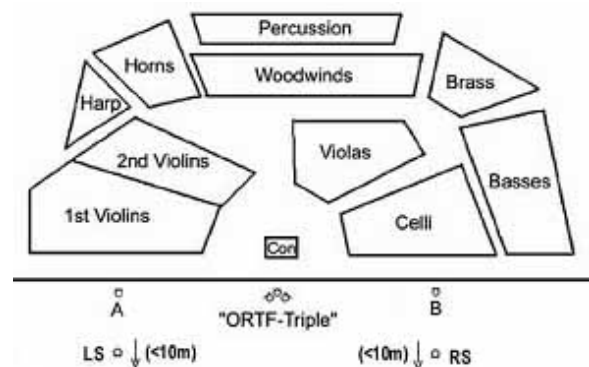


**Fig. 43:** 'AB-PC' (AB-Polycardioid-Centerfill) surround microphone technique (cardioid rear microphones should be separated by and at a distance from the front system of at least the reverberation radius) (see [47] for further details)

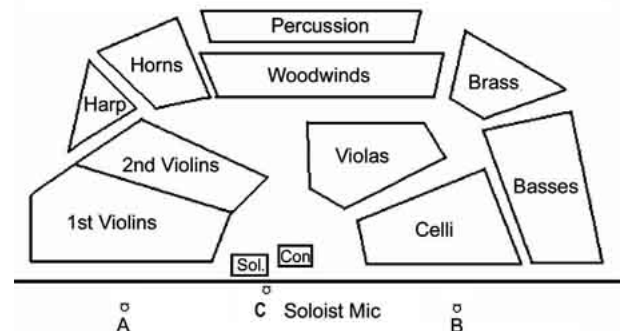


**Fig. 44:** (large) AB-Centerfill system in the style of the early Mercury "Living Presence" recordings

### 13. ACKNOWLEDGEMENTS

This work was supported by: 'IEM-Institute of Electronic Music and Acoustics' (at the KUG - University of Music and Performing Arts, Graz, Austria) and the 'Salzburg Festival'.

Thanks goes to: microphone manufacturer Schoeps (Germany), microphone manufacturer Soundfield (Great Britain), acoustic consultants Müller-BBM Schalltechnik, Munich (Germany) for loan of the Neumann KU81 dummy head; Markus Noisternig, Günther Harner and Ilya Bajerl for operational support (recording), Alois Sontacchi for room acoustic measurements of the IEM control-room, Markus Waldner (literature support), to the musicians involved and - last but not least – all the test listeners who have devoted their time.

### REFERENCES

- [1] E. Torick, "Highlights in the history of Multichannel Sound", *JAES*, vol. 46, No.1/ 2, 1988, pg. 27
- [2] H. Wittek, F. Rumsey, and G. Theile "Perceptual Enhancement of Wavefield Synthesis by Stereophonic Means", *JAES*, vol.55, 2007, pg. 723-751
- [3] E. B.Meyer and D. R. Moran "Audibility of a CD-Standard A/D/A Loop Inserted into High-Resolution Audio Playback", *JAES*, vol. 55, 2007, pg. 775-779
- [4] H. Wittek, G. Theile "The Recording Angle – Based on Localization Curves", paper 5568, 112<sup>th</sup> AES-Convention, Munich, 2002
- [5] C. Hugonnet, P. Walder "Stereophonic Sound Recording", John Wiley & Sons, 1998 ISBN 0-471-9787-0
- [6] E. Sengpiel, Informations on website: [www.sengpielaudio.de](http://www.sengpielaudio.de), 2002
- [7] M. Williams "Unified theory of microphone systems for stereophonic sound recording", AES-Preprint No. 2466
- [8] H. Wittek, G. Theile "Investigations into directional imaging using L-C-R stereo microphones" 21. Tonmeistertagung 2000 (German), Proceedings ISBN 3-598-20362-4, p.432-454
- [9] H. Wittek, "Image Assistant", JAVA-Applet and documentation on website: [www.hauptmikrofon.de](http://www.hauptmikrofon.de), 2002
- [10] J. Knothe, G. Plenge "Panoramaregler mit Berücksichtigung der frequenzabhängigen Pegeldifferenzbewertung durch das Gehör", paper presented at the 11. Tonmeistertagung (VDT), Berlin 1978, pg.136
- [11] D. Griesinger „Stereo- and Surround-Panning in Practice“, Preprint # 5564, AES 112<sup>th</sup> Convention, Munich 2002
- [12] A. Gernemann, T. Rösner, "The Dependence of the Stereophonic Localization on the Sound Quality of Monitor Speakers", Proceedings (German), 20. Tonmeistertagung, Karlsruhe 1998 (ISBN 3-598-20361-6), pp.828
- [13] E. Benjamin, R. Brown „The effect of head diffraction on stereo localisation in the mid-frequency range“, Preprint 7018, 122<sup>nd</sup> AES-Convention, Vienna, 2007
- [14] U. Hermann, V. Henkels „Comparison of 5 surround microphone methods“ (German) Proceedings, 20. Tonmeistertagung VDT 1998, pg. 517 (ISBN 3-598-20361-6)
- [15] M. Sawaguchi, A. Fukada „Multichannel sound mixing practice for broadcasting“, proceedings, IBC Conference, 1999
- [16] J. Valin, "The RCA Bible – A Compendium of Opinion on RCA Living Stereo Records", Second Edition, The Music Lovers Press (1994), pp. 123,124
- [17] Rumsey F., Segar P., "Optimisation and Subjective Assessment of Surround Sound Microphone Arrays", paper #5368, 110<sup>th</sup> AES Convention, Amsterdam, May 2001

- 
- [18] E. Sengpiel, lecture notes at the "Hochschule der Künste", Berlin
- [19] Wuttke J. "The Microphone between Physics and Emotion", paper presented at 20. Tonmeistertagung des VDT, Karlsruhe 1998, Proceedings (ISBN 3-598-20361-6), pp.460
- [20] J. La Grou, "Orchestral Recording ", *Mix*, Feb. 1994, pp. 32
- [21] Blumlein A., "Improvements in and relating to Sound-transmission, Sound-recording and Sound-reproducing Systems", British patent, 394,325, Dec.14 1931, (reprinted in the "Anthology of Stereophonic Techniques", AES, 1986, pp.32-40)
- [22] Steinberg J. C., Snow W. B., "Auditory Perspective – Physical Factors", *Electrical Engineering*, vol.53, no.1., Jan. (1934), pp12-15
- [23] Yost W.A., Wightman F.L., Green D.M., "Lateralisation of filtered clicks", *Journal of the Acoust. Soc. of America*, 50, pp. 1526-1531, (1971)
- [24] Hirata Y., "Improving stereo at L.F.", *Wireless World*, pp.60, Oct. 1983
- [25] Yanagawa H., Higashi H., Mori S., "Interaural correlation coefficients of the dummy head and the feeling of widness", *A.S. J. Tech. Rep. H-35-1*, (1976)
- [26] Suzuki A., Tohyama M., "Interaural cross-correlation coefficient of Kemer head and torso simulator", *IECE Japan, Tech. Rep. EA80-78*, (1981)
- [27] Griesinger D., "Objective Measures of Spaciousness and Envelopment", paper at the AES 16<sup>th</sup> Int. Conf. on Spatial Sound Reproduction
- [28] Morimoto M., Meakawa Z. "Effects of Low Frequency Components on Auditory Spaciousness", *Acustica Vol. 66* (1988)
- [29] Hidaka T., Beranek L. and Okano T., "Interaural cross-correlation, lateral fraction, and low- and high-frequency sound levels as measures of acoustical quality in concert halls", *J. Acoust. Soc. Am.* 98 (2), August (1995)
- [30] Griesinger D., "The Theory and Practice of Perceptual Modeling – How to use Electronic Reverberation to Add Depth and Envelopment Without Reducing Clarity", paper presented at the 21. Tonmeistertagung of the VDT, Hannover 2000, pp.766-795 (see also [www.world.std.com/~griesngr](http://www.world.std.com/~griesngr))
- [31] Hecker P., "The Decision of the Microphone Spacing and its Creative Benefit", (German) paper presented at the 21. Tonmeistertagung of the VDT, Hannover 2000, pp.796-804, Proceedings (ISBN 3-598-20362-4)
- [32] R. K. Cook, R.V. Waterhouse, R.D. Berendt, S. Edelman, and M.C. Thompson Jr., "Measurement of Correlation Coefficients in Reverberant Sound Fields", *Journal Acoust. Soc. Amer.*, vol.27, no.6, Nov. 1955, pp.1072
- [33] G. W. Elko, "Spatial Coherence Functions for Differential Microphones in Isotropic Noise Fields", in: 'Microphone Arrays' (M.Brandstein, D. Ward, Editors), Springer, 2001, (ISBN 3-540-41953-5), pp.61
- [34] J. Blauert "Zur Verwendung der Worte 'Korrelationsgrad' und 'Kohärenzgrad'" (German), *ACUSTICA*, vol.35, 1976, pp.215
- [35] G. Martin, "A New Microphone Technique For Five-Channel Recording", paper # 6427, 118<sup>th</sup> AES Convention, 2005
- [36] G. Theile, "Multichannel Natural Music Recording Based on Psychoacoustic Principles", Paper AES 19th Int. Conference, May 2001
- [37] G. Theile "Microphone and mixing concepts for 5.1 music recordings", paper presented at the 21. Tonmeistertagung of the VDT, Hannover 2000, ISBN 3-598-20362-4, pg. 384-431
- [38] P. Damaske "Subjektive Investigation on Sound Fields" (German) *ACUSTICA*, vol. 19 (1967/68), pg. 199-213
-

- 
- [39] E. Pfanzagl "Vom Punkt zum Flächenstrahler", (German), *PROSPECT – Magazin für Bühnen und Veranstaltungstechnik*, Dez. 2007, pg. 32 (ISSN: 1686840)
- [40] Marshall A.H., "Acoustical Determinants for the Architectural Design of Concert Halls", *Arch. Sci. Rev.*, Australia 11, 1968, pp.81-87
- [41] Barron M., "The Subjective Effects of First Reflections in Concert Halls – The need for Lateral Reflections", *J. Sound Vib.*, 15, 1971, pp. 475-494
- [42] Barron M., Marshall A. H., "Spatial impression due to early lateral reflections in concert halls", *J. Sound Vib.* 77, 1981, pp. 211-232
- [43] Kleiner M., "A new way of measuring the lateral energy fraction", *Appl. Acoust.* 37, 1989, pp.321-327
- [44] Griesinger D., "Spaciousness and Envelopment in Musical Acoustics", paper presented at the "19<sup>th</sup> International Convention on Sound Design" of the VDT in Karlsruhe, 1996, *Proceedings* (ISBN 3-598-20360-8), pp. 375-391
- [45] Griesinger D., "Spatial Impression and Envelopment in Small Rooms", paper #4638 (H-2), 103<sup>rd</sup> AES Convention, 1997
- [46] Griesinger D., "General overview of spatial impression, envelopment, localization and externalization", *Proc. 15<sup>th</sup> Int. Conf. AES on Small Rooms Acoustics*, pp. 136-149, Denmark Oct/Nov 1998, (see also: [www.world.std.com/~griesngr](http://www.world.std.com/~griesngr))
- [47] E. Pfanzagl-Cardone, "In the Light of 5.1 Surround: Why "AB-Polycardioid Centerfill" (AB-PC) is superior for Symphony Orchestra Recording", Preprint 5565, 112<sup>th</sup>-AES Convention, Munich 2002
- [48] G. Theile "Das Kugelflaechenmikrofon", paper, 4. Tonmeistertagung des VDT, Munich 1986, pg. 277
- [49] C. Langen, "Signal Processing for the KFM 360", paper (German) 21. Tonmeistertagung des VDT, Hannover 2000, pg.329-346 (ISBN 3-598-20362-4)
- [50] M. Gerzon, "Periphony (with Height Sound Reproduction)", paper M07, AES-Convention 2ce, March 1972
- [51] F. Camerer "Classical music in Radio and TV – a multichannel challenge", paper, 2001 – see also: [www.irt.de/wittek/hauptmikrofon/ORF\\_und\\_FH\\_D.htm](http://www.irt.de/wittek/hauptmikrofon/ORF_und_FH_D.htm) (ORF-Listening Test also performed at FH Düsseldorf)
- [52] S. Bech, N. Zacharov "Perceptual Audio Evaluation – Theory, Method and Application", Wiley & Sons, 2005, pp. 118, (ISBN-13: 978-0-470-86923-9)
- [53] D. Hammershoi, H. Moller "Binaural Technique – Basic Methods for Recording, Synthesis, and Reproduction", from *Communication Acoustics* (J. Blauert, Editor), pg.232, Springer 2005, ISBN-13: 978-3-540-22162-3
- [54] H. Moller, D. Hammershoi, C.B. Jensen and M. F. Sorensen "Transfer Characteristics of Headphones Measured on Human Ears", *JAES*, vol. 43, pp. 203-217, April 1995
- [55] H. Moller, D. Hammershoi, C.B. Jensen and M. F. Sorensen "Evaluation of Artificial Heads in Listening Test", *JAES*, vol. 47, No.3, 1999
- [56] Helmut Wittek, 2002: "Image Assistant 2.0" see <http://www.hauptmikrofon.de/ima2-folder/ImageAssistant2.html>
- [57] G. Martin "A New Microphone Technique For Five-Channel Recording", paper # 6427, 118<sup>th</sup> AES Convention, 2005
- [58] E. Pfanzagl-Cardone "3.0 Microphone for Surround-Recording", Australian patent No. AU2005100255, IP Australia ([www.ipaustralia.gov.au](http://www.ipaustralia.gov.au))
- [59] J. Berg and F. Rumsey "Verification and correlation of attributes used for describing the spatial quality of reproduced sound", Preprint, AES 19<sup>th</sup> Int. Conference, (2001)
-

- 
- [60] A. Gabrielsson and B. Lindström “Perceived Sound Quality of High-Fidelity Loudspeakers”, *JAES*, vol. 33, No.1/2, pg.33 (1985)
- [61] F. E. Toole “Subjective Measurements of Loudspeaker Quality and Listener Performance”, *JAES* vol. 33, No1/2, pg. 2-32 (1985)
- [62] Hiyama, Komiyama, Hamasaki “The minimum Number of Loudspeakers and Its Arrangement for Reproducing the Spatial Impression of Diffuse Sound Field”, paper 5674, AES 113th Convention, Los Angeles, Oct. 2002
- [63] B. Shirley, P. Kendrick, C. Churchill “The Effect of Stereo Crosstalk on Intelligibility: Comparison of a Phantom Stereo Image and Central Loudspeaker Source”, *JAES*, vol. 55 No. 10, Oct. 2007, pg. 825ff
- [64] R. Streicher, W. Dooley, “Basic Stereo Microphone Perspective – A Review”, *JAES*, vol.33, no.7/8, pp.548 – 556, Jul/Aug. 1985
- [65] M. Nakahara “Multichannel Monitoring Tutorial Booklet” (M2TB) rev. 3.5.2, Yamaha Corp. 2005, SONA Corporation 2005, pg. 41
- [66] E. Prokofieva “Relation between correlation characteristics of sound field and width of listening zone”, paper 7089, 122<sup>nd</sup> AES-Convention, Vienna 2007
- [67] J. C. Steinberg, W. B. Snow „Auditory Perspective – Physical Factors“, *Electrical Engineering*, vol. 53, no. 1Jan (1934), pp.12-15

## APPENDIX

| 'Dependent Variable' | Sum of squares | df | Mean square | F       | p             | stat. significant differences between mean-values of the surr.-mic ratings |
|----------------------|----------------|----|-------------|---------|---------------|----------------------------------------------------------------------------|
| Preference           | 38.5897        | 3  | 12.8632     | 16.7695 | 0             | 1 & 3, 2 & 3, 4 & 3                                                        |
| Naturalness          | 20.7831        | 3  | 6.9277      | 8.4185  | 0             | 1 & 3, 2 & 3, 4 & 3                                                        |
| Sound-Colour         | 14.5719        | 3  | 4.8573      | 7.5008  | 0.0001        | 1 & 3, 4 & 3                                                               |
| High/low Balance     | 9.7675         | 3  | 3.2558      | 6.3025  | 0.0006        | 1 & 4, 2 & 4                                                               |
| Localisation         | 14.9475        | 3  | 4.9825      | 6.3767  | 0.0005        | 1 & 4, 2 & 4, 3 & 4                                                        |
| Balance              | 21.6322        | 3  | 7.2107      | 9.4913  | 0             | 1 & 3, 2 & 3, 4 & 3, 1 & 4                                                 |
| Width                | 79.2342        | 3  | 26.4114     | 30.7490 | 0             | 1 & 2, 1 & 3, 1 & 4, 2 & 3, 4 & 3                                          |
| Distance             | 35.5233        | 3  | 11.8411     | 16.8673 | 0             | 1 & 2, 2 & 3, 2 & 4                                                        |
| Stage depth          | 17.6275        | 3  | 5.8758      | 5.9987  | 0.0009        | 1 & 3, 2 & 3                                                               |
| Stability            | 8.4431         | 3  | 2.8144      | 2.8690  | 0.0405        | 3 & 4                                                                      |
| Differentiation      | 7.6853         | 3  | 2.5618      | 2.3699  | <b>0.0754</b> | none                                                                       |
| Spatial Impression   | 43.7697        | 3  | 14.5899     | 13.4332 | 0             | 1 & 3, 2 & 3, 4 & 3                                                        |
| Wet/dry Balance      | 43.0233        | 3  | 14.3411     | 22.2303 | 0             | 1 & 3, 2 & 3, 4 & 3                                                        |
| Rear non-/intrusv.   | 4.0756         | 3  | 1.3585      | 0.7353  | <b>0.5335</b> | none                                                                       |
| Rear soft/loud       | 7.2100         | 3  | 2.4033      | 2.7409  | 0.0475        | none                                                                       |

Rem.: 1=OCT, 2=DECCA, 3=KFM, 4=ABPC

**Table A1: ORCH-5.1 Surround Listening Test; 1-way ANOVA concerning the influence of the variable „microphone-array“ on the rating of 15 attributes (25 listeners)**

| 'Dependent Variable' | Sum of squares | df | Mean square | F       | p             | stat. significant differences between mean-values of the surr.-mic ratings |
|----------------------|----------------|----|-------------|---------|---------------|----------------------------------------------------------------------------|
| Difference to REF    | 31.3394        | 2  | 10.4465     | 5.6872  | 0.0027        | 1&3, 2&3, 4&3                                                              |
| Preference           | 31.4602        | 3  | 10.4867     | 5.5144  | 0.0032        | 3&4                                                                        |
| Naturalness          | 35.4           | 3  | 11.8        | 10.5526 | 4.02E-05      | 1&3, 2&3, 4&3                                                              |
| Sound-Colour         | 37.8214        | 3  | 12.6071     | 14.0776 | 3.14E-06      | 1&3, 2&3, 4&3                                                              |
| High/low Balance     | 9.4583         | 3  | 3.1528      | 2.8363  | <b>0.0517</b> | none                                                                       |
| Localisation         | 23.7339        | 3  | 7.9113      | 3.5448  | 0.0239        | 2&4, 3&4                                                                   |
| Balance              | 26.4394        | 3  | 8.8131      | 7.1988  | 6.61E-04      | 3&4                                                                        |
| Width                | 29.2201        | 3  | 9.74        | 8.2485  | 2.63E-04      | 1&3, 2&3, 4&3                                                              |
| Distance             | 10.1352        | 3  | 3.3784      | 2.0932  | <b>0.1183</b> | none                                                                       |
| Stage Depth          | 14.5589        | 3  | 4.853       | 3.251   | 0.0329        | none                                                                       |
| Stability            | 0.4451         | 3  | 0.1484      | 0.0775  | <b>0.9718</b> | none                                                                       |
| Differentiation      | 7.6208         | 3  | 2.5403      | 1.4245  | <b>0.2516</b> | none                                                                       |
| Spatial Impression   | 29.0688        | 3  | 9.6896      | 11.5494 | 1.89E-05      | 1&3, 2&3, 4&3                                                              |
| Wet/dry Balance      | 15.7089        | 3  | 5.2363      | 6.9765  | 8.08E-04      | 1&3, 2&3                                                                   |

Rem.: 1=OCT, 2=DECCA, 3=KFM, 4=ABPC

**Table A2: ORCH-Binaural Listening Test; 1-way ANOVA concerning the influence of the variable „microphone-array“ on the rating of 14 attributes (10 listeners)**

|              | pref.   | natrl.  | colour | hi/low  | local. | bal.    | width   | distnc. |
|--------------|---------|---------|--------|---------|--------|---------|---------|---------|
| preference   | 1.0000  | 0.8383  | 0.6712 | 0.0238  | 0.5875 | 0.6879  | 0.6289  | -0.0964 |
| naturalness  | 0.8383  | 1.0000  | 0.7539 | -0.0377 | 0.5967 | 0.5964  | 0.5010  | -0.0530 |
| sound-colour | 0.6712  | 0.7539  | 1.0000 | 0.0935  | 0.5113 | 0.5178  | 0.4978  | 0.0392  |
| high/low bal | 0.0238  | -0.0377 | 0.0935 | 1.0000  | 0.2035 | -0.0645 | -0.0054 | 0.2603  |
| localisation | 0.5875  | 0.5967  | 0.5113 | 0.2035  | 1.0000 | 0.5292  | 0.4410  | 0.1248  |
| balance      | 0.6879  | 0.5964  | 0.5178 | -0.0645 | 0.5292 | 1.0000  | 0.4748  | -0.0219 |
| width        | 0.6289  | 0.5010  | 0.4978 | -0.0054 | 0.4410 | 0.4748  | 1.0000  | -0.2470 |
| distance     | -0.0964 | -0.0530 | 0.0392 | 0.2603  | 0.1248 | -0.0219 | -0.2470 | 1.0000  |
| depth        | 0.4858  | 0.3996  | 0.4308 | 0.1067  | 0.3497 | 0.3813  | 0.4637  | -0.2565 |
| stability    | 0.5407  | 0.6025  | 0.4845 | -0.0402 | 0.4676 | 0.5312  | 0.3021  | 0.0675  |
| differ.tn    | 0.5011  | 0.5017  | 0.4772 | 0.1603  | 0.5234 | 0.5742  | 0.2438  | 0.1955  |
| spaciousness | 0.6829  | 0.6398  | 0.5724 | 0.0083  | 0.4336 | 0.4867  | 0.6522  | -0.0901 |
| wet/dry bal  | 0.4288  | 0.2912  | 0.2808 | 0.0112  | 0.1330 | 0.2790  | 0.6483  | -0.4005 |
| r. non/intr  | 0.2988  | 0.3512  | 0.3836 | 0.1406  | 0.2875 | 0.2021  | 0.2987  | 0.0471  |
| r. soft/loud | -0.0069 | 0.0827  | 0.1738 | 0.0777  | 0.0718 | 0.1135  | 0.0220  | 0.0777  |

|              | depth   | stabil. | difftn | spatial | wet/dry | r.n/intr | r.sft/ld |
|--------------|---------|---------|--------|---------|---------|----------|----------|
| preference   | 0.4858  | 0.5407  | 0.5011 | 0.6829  | 0.4288  | 0.2988   | -0.0069  |
| naturalness  | 0.3996  | 0.6025  | 0.5017 | 0.6398  | 0.2912  | 0.3512   | 0.0827   |
| sound-colour | 0.4308  | 0.4845  | 0.4772 | 0.5724  | 0.2808  | 0.3836   | 0.1738   |
| high/low bal | 0.1067  | -0.0402 | 0.1603 | 0.0083  | 0.0112  | 0.1406   | 0.0777   |
| localisation | 0.3497  | 0.4676  | 0.5234 | 0.4336  | 0.1330  | 0.2875   | 0.0718   |
| balance      | 0.3813  | 0.5312  | 0.5742 | 0.4867  | 0.2790  | 0.2021   | 0.1135   |
| width        | 0.4637  | 0.3021  | 0.2438 | 0.6522  | 0.6483  | 0.2987   | 0.0220   |
| distance     | -0.2565 | 0.0675  | 0.1955 | -0.0901 | -0.4005 | 0.0471   | 0.0777   |
| depth        | 1.0000  | 0.2660  | 0.3283 | 0.4892  | 0.5055  | 0.2367   | -0.0427  |
| stability    | 0.2660  | 1.0000  | 0.6547 | 0.2910  | 0.0415  | 0.2213   | 0.0217   |
| differ.tn    | 0.3283  | 0.6547  | 1.0000 | 0.2661  | 0.0052  | 0.2588   | 0.1391   |
| spaciousness | 0.4892  | 0.2910  | 0.2661 | 1.0000  | 0.5099  | 0.3826   | 0.0632   |
| wet/dry bal  | 0.5055  | 0.0415  | 0.0052 | 0.5099  | 1.0000  | 0.1328   | 0.0108   |
| r. non/intr  | 0.2367  | 0.2213  | 0.2588 | 0.3826  | 0.1328  | 1.0000   | 0.5092   |
| r. soft/loud | -0.0427 | 0.0217  | 0.1391 | 0.0632  | 0.0108  | 0.5092   | 1.0000   |

**Table A3:** 'ORCH 5.1' listener ratings - Correlation-matrix for 15 attributes (Pearson product moment corr. coeff.)

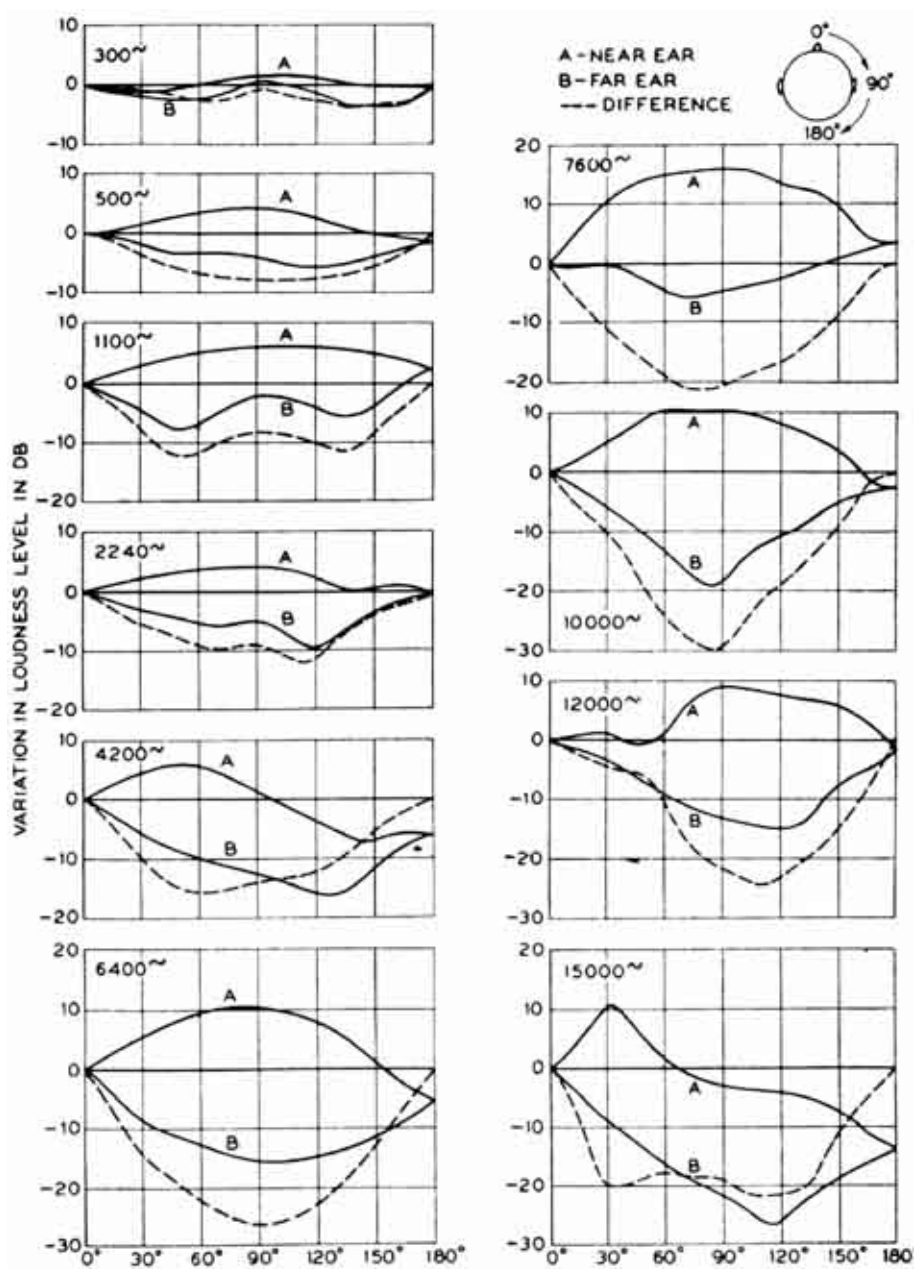


Fig. 45: Variation in loudness level as a sound source is rotated in a horizontal plane around the head (from [67])