# How do we perceive Early Reflexions ?
## Some Notes on the Directivity of Music Instruments

*(Rezeption früher Reflexionen – ein paar Notizen zur Richtcharakteristik von Musikinstrumenten)*

*Urban Schlemmer*

University of Music and performing Arts, Vienna
schlemmer@nusurf.at

## Abstract

Psychoacoustic research of early reflexions (ER) and their spectral attributes leads to the question how directivity of sound sources contributes to spatial impression. This paper investigates the relationship of the limit of the precedence effect to echo-perception (LPE) to source-characteristics (as given by the musical score and the directivity of playing instruments) in the presence of six or more reflexions reproduced in the listening room by five or more loudspeakers.

## 1. Introduction

The precedence effect describes the observation that if two or more delayed signals are presented to the hearing system an assessment process takes place, inhibiting redundant information already contained in the preceeding signal [1] - [3]. The inhibited information includes angle of occurrence, delay times and spectral contents of the delayed sound event. A hearing event evolves with the sensation of space surrounding it.

This sensation is referred to as 'auditive Räumlichkeit' in German by Blauert and many other authors[1]. In English and Japanese literature there are several labels for this sensation: the term auditory spacial impression (ASI) is the umbrella term referring to auditory perception of the precedence effect. It is subdevided into at least two components: auditory source width (ASW) and listener envelopment (LEV)[2]. Bradley [5] introduced the term LEV to evaluate concert hall acoustics which was later adapted to laboratory experiments by Morimoto [6] and others.

[1]A list of German literature dealing with 'Räumlichkeit' is given in [2], page 75.

[2]A list of English literure dealing with ASI can be found in [4]. Please note that ASI does not translate directly to 'Räumlichkeit' – it is better described as looking at the spacial sensations evolving from the perception of the precedence effect from another point of view.

It is assumed that the precedence effect is closely related to cognition[1]. Pattern recognition enables our brains to group similar sensational events together and to perceive them as an extended attribute of one source. At an even higher level of cognition, patterns are grouped into streams. Many different perceptual streams coexist at the same time resulting from the various elicitation of our senses. In a cocktailparty situation e.g. one perceptual stream may be assigned to background noises and one stream may be assigned to the actual conversation. Other streams may also be assigned to smelling, tasting, etc. Switching attention to one of those streams is triggered either intentionally or by events that do not fit into the current pattern or that contain new information about the environment.

## 2. Hypothesis on the Relation of ESIS and LSIS through ER

It is assumed that two auditory perceptional streams evolve from total spatial impression and reverberation: First, an early spatial impression stream (ESIS) may be assigned to information concerning the sound source. This is loudness, directivity, minimum and maximum elevation and azimuth. Second, a late spatial impression stream (LSIS) may be assigned to the perception of room attributes: geometry[2] and absorbtion of walls.

These two perceptional streams are related through the listener's remembered perceptual experience with absobtion and room geometry on one hand and remembered subjective ASI of the sound sources on the other hand. It is therefore assumed that level and spectral attributes of ER contribute to the interconnexion of ESIS and LSIS through knowledge [7]. Level and spectral attributes of ER relate to directivity of music instruments. This interconnexion may be an important factor for a natural sounding room simulation with five or more loudspeakers.

[1]See also [3], page 49ff
[2]Geometry includes volume, length, highth and width as well as balconies, columns and sound-scattering surfaces.

## Transition Period of ESIS to LSIS

In a transition period ER will contribute to both ESIS and LSIS. The transition time may be defined as the crossing point of the LPE with the level of the physical sound field as shown in the following sketch. This may be the time when the diffuse sound field becomes just audible as a seperate auditory event. Blauert mentions this relationship as a "characteristic blurr of auditory events in the timedomain caused by late reflexions and reverberation" [1].

Lehmann [8] stated that a transition time when it makes no longer sense to analyse descrete reflexions in a sound field - which will then be discribed by statistical terms - is approximately given by $t_{gr} = 2\sqrt{V}$ .

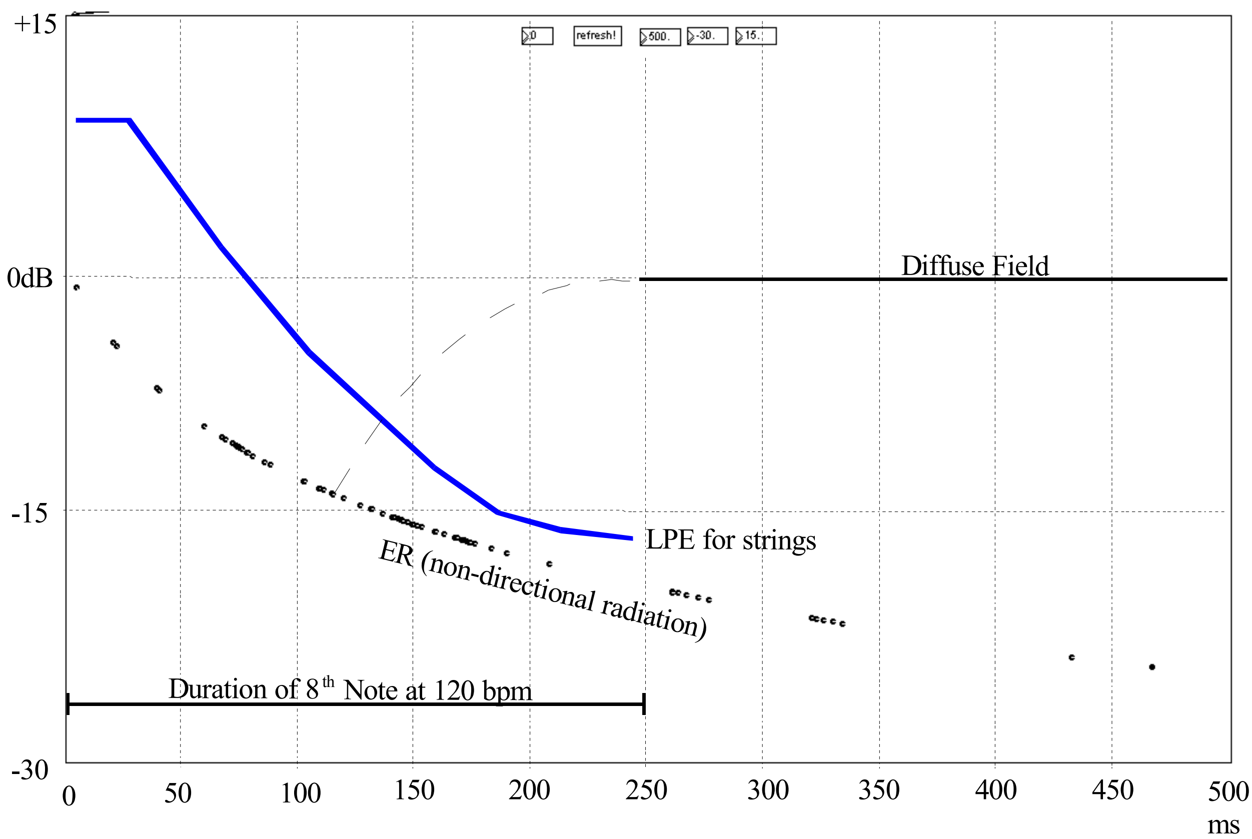


*Figure 1 shows calculated ER for the Vienna Musikverein Hall at a Distance of 10m from the sound source, which is approximately the hall distance for the directed radiation of strings. At the hall distance the diffuse field has the same level as the direct sound. An estimate of the LPE is shown by the blue line. The directivity of violins and violas contains level differences of up to 10 dB in the 2kHz and 4kHz band for the direction of direct sound compared to the reflexion of the ceiling, contributing to LEV in this room [11].*

[1]In German: "charakteristisches Verschleifen der Hörereignisse infolge von späten Rückwürfen und Nachhall" [2], page.75

## 3. LPE in Relation of sound source Characteristics

Morimoto [6] showed in 2003 that reflexions at LPE contribute to a subjective feeling of LEV. A high level of early reflexions in diffuse sound fields is generally preferred in subjective listening tests [9][1]. However, the perception of echos must be avoided. Therefore controlling ER at LPE is a key in reproduced ASI.

### 3.1 Review of Investigations on LPE

Fig. 3 shows the results of different investigations on LPE in anechoic environments. When the level of the delayed signal approaches LPE, at first, a sensation of image split occurs. Then a second auditory event becomes audible. Finally, the echo is described as 'disturbing'.

**Limit of the precedence effect to 'image split' or 'echo perception' with one single delayed signal**

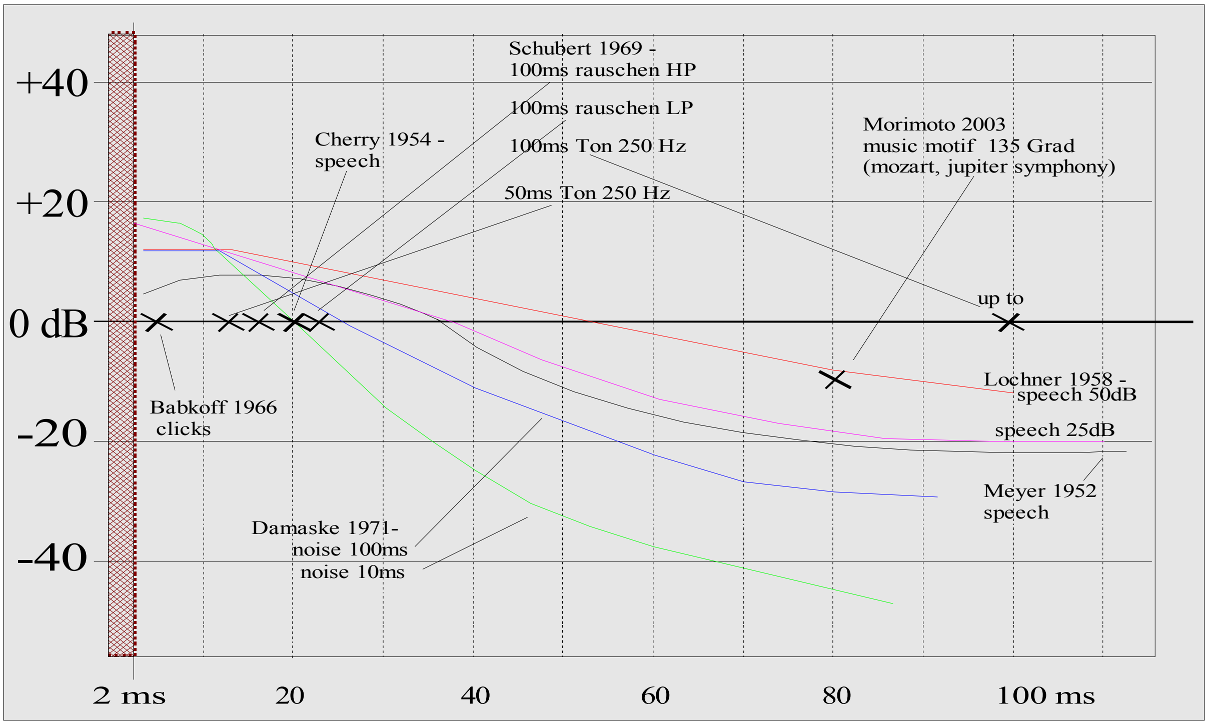


*Figure 2: If not noted otherwise the direct sound is frontal and the angle of the reflexion is set to 40 degrees*

From the preceeding studies, it can be seen that LPE varies greatly with source charcteristics. An incomplete list of some relations may be described as followes:

[1]A survey of measurements of subjective spatial attributes can be found in [10].

| **Condition** | | | | **Investigation** |
|---|---|---|---|---|
| The greater | the number of reflexions | the more unlikely | echos are audible. | Ebata, Sone, Nimura, 1968 [1,p.219] |
| The greater | the spectral incoherence | the more likely | echos are audible. | See 3.2.1 |
| The greater | the transients | the more likely | echos are audible. | Schubert, Wernick, 1969 [1,p. 185] |
| The greater | the time gap between direct and reflexion | the more likely | echos are audible. | [All investigations] |
| The higher | the level of ER | the more likely | echos are audible. | [All investigations] |
| The higher | the overall loudness | the more likely | echos are audible. | See 3.2.2 |
| The shorter | the note durations | the more likely | echos are audible. | Damaske 1971[1,p. 182] |
| The higher | the frequency contents of ER | the more likely | echos are audible. | Schubert and Wernick, 1969[1,p. 185] |
| The greater | the angle of occurence (0...180 degrees) | the more likely | echos are audible. | Boerger 1965 [1,p. 183] |

*Table1: List of qualitative relations of LPE to different source charcteristics*

## 3.2 Experiments with anechoic Drum Sounds

Because it can not be predicted from the preceeding investigations which values of the LPE can be expected if music instruments in a studio are used for simulated reflexions, LPE was measured with three anechoically recorded instruments in a preliminary study. Additionally a -3dB highshelf filter was applied to the delayed signal.

### 3.2.1 Bass Drum, Hihat and Snare at 50 dB SPL

Fig. 3 shows the results of the first measurement.

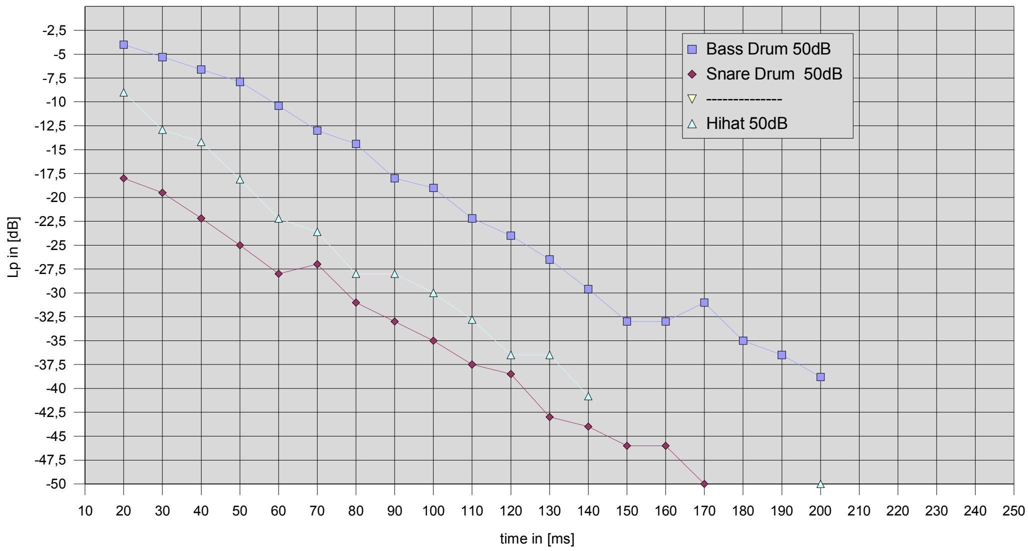


*Figure 3: LPE measured with a single bassdrum hit, a snaredrum hit and a hihat hit at 50 dB SPL. The direct sound was played back from the center and the delayed signal from left surround according to 5.1-ITU in a small dumped studio.(1 subject)*

#### 3.2.2 Bass Drum and Snare Drum at 35 dB SPL

Fig. 4 shows the results of the second measurement.

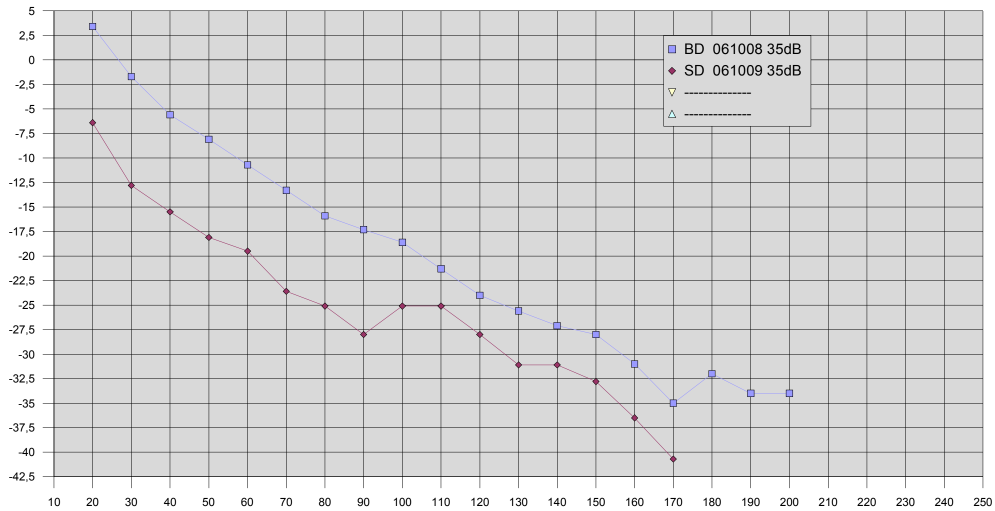


*Figure 4: LPE measured with a single bassdrum-hit and a snaredrum hit at 35 dB SPL.*

#### 3.2.3 Discussion

The measurements correspond qualitatively to the findings in Table 1. LPE decreases with total loudness for high frequency components, meaning that ASI increases with total loudness. This corresponds to findings by Marschall, Keet and Barron as described in [2]. Therefore the contradicting findings as depicted by Blauert in [1] referring to investigations by Lochner and Burger from 1958 may be wrong.

In this study, LPE for low frequency components stays mostly constant for different total loudness exept for short delaytimes smaller than 30ms. LPE for the snare drums is about 15 dB under LPE for the bass drum at a total loudness of 50 dB SPL. This difference results from a wider spectrum, a shorter note duration and more high frequency components compared to the bass drum. The Hihat has a narrower spectrum than the snaredrum but contains more high frequency components than the bass drum and therefore lies between the two other measurements. LPE for comparable frontal reflexions as shown in Fig. 2 is about 10 dB higher than the measured values for 110 degrees.

### 3.3 Summary on Source Characteristics

The precedence effect is continously retriggered by new sound events in a concert. This means that reflexions arriving at the onset of notes are closer to LPE than in the sustain phase. Therefore long sustained notes do not contribute to ASI as much as a series of short notes.

Directivity affects LPE in the following ways:

- The spectral contents of the reflexion differs from the direct sound.
- The overal loudness of the reflexion differs from the direct sound.

This means if ER result from the directivity that lower LPE (e.g. flute, strings, horn), ASI of these music instruments is perceived as high. If, on the other hand, ER result from the directivity that highten LPE (e.g. trombone, trumpet), ASI of these music instruments is perceived as low.

# 4. Experiments with Reflexions of Higher Order

## 4.1 Enlargement of the Listening Area through ER

### 4.1.1 Experimental Setup

For this test, a recording of flute solo[1] (J.B. Bach, BWV 1013) was used, a mono mix from two microfones with a reverberation without ER. The directivity of the flute was considered and transcribed by a filterbank for each of the reflexions, calculated by the urban-reflexions-program, as described in [12]. The pattern was animated through the simulation of a constantly moving sound source.

The following ER pattern was applied:

| | dly time | |
|---|---|---|
| **1. Order** | **1.247 ms** | |
| **1. Order** | **14.76 ms** | |
| | 20.32 ms | 2. Order |
| | 26.32 ms | 2. Order |
| | 32.83 ms | 3. Order |
| **1. Order** | **39.38 ms** | |
| | 48.98 ms | 2. Order |
| **1. Order** | **49.07 ms** | |
| | 49.76 ms | 2. Order |
| | 52.88 ms | 3. Order |
| | 57.24 ms | 3. Order |
| | 57.89 ms | 2. Order |
| | 59.02 ms | 3. Order |
| | 61.50 ms | 3. Order |
| **1. Order** | **97.43 ms** | |
| | 97.91 ms | 2. Order |
| | 103.6 ms | 2. Order |
| | 106.3 ms | 3. Order |
| | 106.9 ms | 2. Order |
| | 109.3 ms | 3. Order |
| **1. Order** | **114.8 ms** | |

*Table2: calculated delaytimes and increased density through reflexions of 2nd and 3rd order*

[1]Studios der MDW, Birgit Herrmannseder, Flöte, Tonmeister: U. Schlemmer, 2003

The angle of appearance was resolved considering the speaker setup shown in Fig.9. In addition, a standart five-speaker-ITU setup was employed.

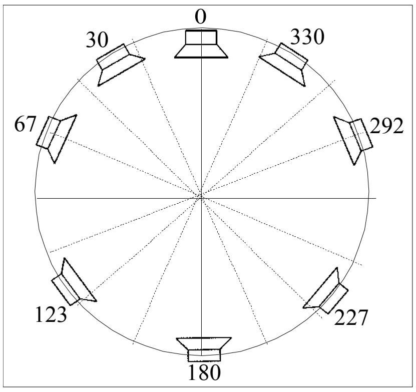


Fig. 9: *eight speaker setup used for experiment 4.1*

#### 4.1.2 First impressions

- Although the latest simulated reflexions had delay times of 160 ms and distances between reflexions appeared of 25 ms, no echos could be heard.
- The effect is subtle and does not cause unwanted coloration. It is hardly consciously perceived.
- Single loudspeakers are not localized, even if one moves far outside the sweetspot.
- Less reverberation is needed in the mix.
- Eight channels sound much better than five. Particularly speakers at the sides seem to be important for the representation of lateral reflexions.
- The feeling of envelopment seems realistic and location-independent.
- The surround-centre seems to be important as you move around and turn around in the soundfield.
- If you turn around, the feeling of beeing 'misplaced' disappears. One enjoys an enlarged listening area - no need to stay in the sweetspot. You can walk around.
- It works best for one solo-instrument. Stereo-techniques can be tricky to integrate.

Most importantly these reflexions or delayed signals are not being allowed to have the same frequency response. If none or the same filtering is used for all delays, then:

- the image of the instrument is blurred
- ASI sounds artificial
- when moving around combfilter-effects are audible.

#### 4.1.3 Discussion

As Franssen showed in 1960 [13] after a successful localization, the recognized angle of incidence is assumed as long as a new evaluable localization-stimulus is given. This may be an explanation for the fact that localization remains stable even if one turns around 180 degrees. This requires no disturbing reflexions to occur when moving one's head. This would is the case when the ER-pattern can be assigned to a real room or if the ER-pattern has the same characteristics as in a real room, because in this case the occurrance of reflexions is predictable. [14] - [16] When a new localizition stimulus occurs the acoustical situation in the new position is approved even if a localization is not possible due to the head being in an infavourable position. The head movement therefore acts as a validation of the reproduced ER-pattern: when the reflexions or delayed signals can be comprehensibly assigned to a natural room or a remembered environment[1], the previously occurring ASI may be affirmed.

Another experience is, that with a non-animated ER-pattern one can omit the ER-pattern after a while and nothing seems to be missing. This is also because the pattern is being remembered even when no new stimulus is given. Only when the listener moves around, a validation of the remembered ER pattern takes place. This corresponds to the above, as when you move, the previously stored pattern is compared to the actual situation for orientation purposes [2][14].

Following Mackensen, spontaneous head movements improve localization. [7] With two speakers, the evaluation of the ER-pattern through head movements may not be possible and the validation will fail. In contrast, it feels pleasent if the localization stays stable in spite of head movements and walking around.

### 4.2 Recognition of Source Directivity

#### 4.2.1 Experiment I

The ER-pattern of Table 2 was calculated with two different directivity patterns, violin and piano. It was then applied to an anechoic recording of the 'pizzicato polka' by R.Strauß, played solely with pizzicato strings[2].

#### 4.2.2 Experiment II

The ER-pattern of Table 2 was calculated with two different directivity patterns, flute and trombone. It was then applyed to the flute recording described in section 4.1.1.

[1]Mackensen mentions 'knowledge, memory and learning effects' as additional cues in localization [7].
[2]Taken from the Dennon CD.

#### 4.2.3 Results and Discussion

- five out of six subject could determine which directivity correspondeds to the played back instruments in experiment 1.
- three out of three subjects could determine which directivity correspondeds to the played back instrument in experiment 2.

Directivity of music instruments seems to be an important factor for the perceived ASI. Auditory level of reflexions and colour differs greatly from the direct sound. Perceptionally a typical coloration is achieved which can be partly reproduced with five or eight loudspeakers. The direct-to-reflexion ratio can be evaluated in subjective ASI. The loudness of the first reflexion from the frontwall of a horn e.g. will be higher than the direct sound. This is desired and intended. The 'indirect' sound of a horn belongs to that instrument, just like a strong reflexion from the ceiling belongs to a string section. Lateral reflexions from a flute may have more high frequency components than the direct sound and therefore may sound more spacious than an oboe [11].

For an desired ASI to occur, reflexions must either lower interaural crosscorelation (which is provided by strong lateral reflexions, e.g. flute, first violins of an orchestra) or be perceptually near LPE. Those reflexions coming from the medial plane are provided by the directivity of violins and violas (ceiling), horn (front wall), trompet and trombone (back wall) to mention just a few examples.

## 5. Conclusion

Knowledge of the directivity of music instruments seems to be an additional cue in ASI aswell as in localisation. Head movements may 'uncover' reflexions as unexpected – which then become audible as disturbing echoes - if the ER-pattern can not be assigned to a real room. An enlarged listening area results from the ability to move around in the reproduced sound field without changing subjective ASI. This is achieved if the simulated ER pattern is so dense that moving towards one loudspeaker keeps reflexions under LPE. In addition, delayed and filtered signals support a decorrelated sound field contributing to ASW and a feeling of envelopment.

## 6. Acknowledgements

This work was supported by ideas of Prof. Ulrich Vette. The Author wants to thank Dr.-Ing. Klaus Wogram and his team from PTB in Braunschweig for support with the directivity data of music instruments.

# 7. References


[1] Blauert, J., “Räumliches Hören” *Hirzel-Verlag, Stuttgart, 1972*

[2] Blauert, J., “Räumliches Hören - 1. Nachschrift - Neue Ergebnisse und Trends seit 1972”, *Hirzel-Verlag, Stuttgart, 1985*

[3] Blauert, J., “Räumliches Hören - 2. Nachschrift - Neue Ergebnisse und Trends seit 1982”, *Hirzel-Verlag, Stuttgart, 1997*

[4] Griesinger, D., “The psychoacoustics of apparent source width, spaceousness and envelopment in performance spaces”, *Acta Acoustica 83/721ff, 1997*

[5] Bradley, J.S. and Souloudre G.A., “Objective measures of Listener Envelopment”, *J. Acoust Soc. Am. 98/2590ff , 1995*

[6] Morimoto, M., “The Relation Between Spacial Impression And The Precedence Effect”, *15th ICA Norway, 1995*

[7] Mackensen, P., “Auditive Localization”, *Dissertation TU Berlin, 2004*

[8] Lehmann, P., “Über die Ermittlung raumakustischer Kriterien und deren Zusammen hang mit subjektiven Beurteilungen der Hörsamkeit” *Dissertation, TU Berlin, 1976*

[9] Barron, M., “Spatial Impression due to Early Lateral Reflections in Concert Halls”, *J. Sound and Vibration 77/211ff, 1981*

[10] Mason, R., “Elicitation and measurement of auditory spatial attributes in reproduced sound”, *Dissertation Univ. of Surrey, 2002*

[11] Mayer, J., “Akustik und musikalische Aufführungspraxis”, *Hirzel Stuttgart, 1975*

[12] Schlemmer, U., “Representation of Spatial impression through the simulation of source image models with 5 or 8 loudspeakers”, *MA Thesis, MDW Vienna, 2006*

[13] Franssen, N., “Some considerations of the mechanism of direction hearing”, *Dissertation Techn. Hochschule Delft, 1960*

[14] Blauert, J., Col, J.-P., “A study of temporal aspects in spacial hearing”, *Auditory psychology, and perception, 531-538, Pergamon Press, Oxford, 1992*

[15] Clifton, R. K., “Breakdown of echo suppression in the precedence effect”, *JASA 82, 1834ff, 1987*

[16] Blauert, J., Col, J.-P., “Etude de quelques aspects temporels de l'audition spaciale”, *CNRS Marseilles, 1989*