

# **SPECTRAL AND TEMPORAL CONTRIBUTION OF DIFFERENT SIGNALS TO ASW ANALYZED WITH BINAURAL HEARING MODELS.**

43.55.Mc; 43.66.Pn

Becker, Jörg.  
Ford of Europe, Acoustic Centre Cologne ACC  
D-50725 - Köln  
Germany  
Tel: +49 221 9031951  
Fax: +49 221 9035145  
E-mail: jbecke11@ford.com

## **ABSTRACT**

For the characterization of the acoustic quality of concert halls spaciousness is an important term. In the last years the quantities ASW and LEV have been established in modern room acoustics for the judgement of the spatial component of sound fields.

In addition to the spatial information, which could be only binaurally evaluated, monaural attributes like level, spectral composition and temporal composition have an influence on the spatial perception of sound sources. This presentation focuses on the evaluation of ASW in relation to monaural signal attributes with the help of binaural models including a peripheral unit.

## **INTRODUCTION**

Factor analysis on room acoustics showed that spaciousness is one of three important terms for characterizing the acoustic quality of a concert hall. For that reason it is of great interest to find the physical reasons for the amount of spaciousness which a sound field evokes. With regard to human perception the attribute spaciousness can be divided into two parts. One part is directly related to the direct sound and the early reflections and one part is associated with the total sound field or at least with the reverberation part of the sound field. For the judgement of spaciousness connected with the direct sound the quantity auditory source width (ASW) has been established. For the second part which lead to a feeling of being surrounded by the sound the description of listener envelopment (LEV) has been created. A good definition of these quantities can be found at [1].

In order to predict the quantities ASW and LEV several high correlating objective measures have been developed. The search for physically measurable quantities which highly correlate with subjective perceptions bears on the one hand the danger of not noticing all parameters which are responsible for the perception and on the other hand that the sufficiency of the test signals is not obvious and highly correlation quantities may correlate less for another set of test signals. In recent publications a discussion arises about IACC as an objective measure for spaciousness especially for small rooms, although the human sense of hearing uses a system

comparable to the analysis of correlation for localization. Models for this correlation processing already exist for a long time [2]. These models have been developed and modified in order to explain phenomena in binaural hearing like trading experiments [3], or to distinguish between complex stimuli like NoS $\pi$  and NoS0 signals [4]. This progress offers the possibility to use these models as an objective measure for spaciousness. Due to the fact that interaural time delay fluctuations, which can be calculated with these models, are said to be responsible for the perception of ASW [1] good results become more likely.

Several publications showed the influence of different signals on the perception of ASW [5],[6],[7],[8],[9]. In these publications the spectral contribution to the perception of ASW is discussed. Even one of the most basic quantities the SPL of a signal has an influence on the perception of ASW [10]. This presentation focuses on the explanation of spectral, temporal contributions with the help of binaural hearing models. Therefore the influence of the preprocessing in the peripheral unit on the output of correlation mechanisms is discussed and related to the outcome of listening tests made with a virtual acoustic pointer.

### BINAURAL HEARING MODEL

For this research a model consisting of 4 main units was applied. All recordings have been made with an artificial head so that no additional head related transfer function had to be modulated. The first unit of the binaural model was the middle ear transfer function. This time invariant transfer function was modulated as a 3<sup>rd</sup> Order Bessel low pass with a border Frequency of 1200 Hz. This transfer function fits well to the results of [11].

The spectral resolution of the model is given by a filter bank which mirrors the processing of the inner ear. For this purpose 36 Roex filters are calculated with a spacing of 1 filter / ERB [12]. In order to get the filter in the time domain to the magnitude a minimal phase was added and the filters were transformed into time domain. Gammatone filters or even Gammachirp filters are adapted to the waveform of the human middle ear but it seems to be unlikely that in this case the output of a hearing model with these filters would differ to the output of the applied model.

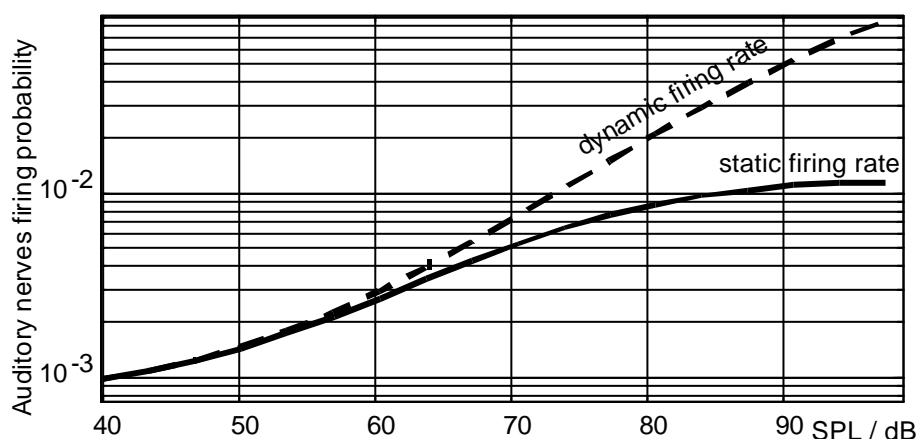


Fig. 1: Output level (firing probability) of the Meddis auditory nerve model [13] for static and dynamic excitation with sinus signal.

For transition of mechanical waves to neural pulses a model of auditory nerves was chosen [13]. The limitation of the dynamic ranges of this model makes it not always applicable. Figure 1

shows the response level of the auditory nerve model on a sinus signal of different levels for the static excitation and an excitation with a hard switched sinus signal (dynamic).

Figure 2 shows the output of the complete hearing model on a hard switched noise signal of 70 dB which was binaurally recorded in the test room. Subplot a) shows the signal after filtering with a Roex filter bank for one filter at 6 ERB (210 Hz). The nonlinear processing of the auditory nerve leads to high amplitudes in the beginning of the signal (Subplot b). This behaviour is essential for the modelling of the temporal processing of the sense of hearing and it is responsible for several masking phenomena. Regarding to the limitation of human sense of hearing noise at threshold level is added. The output of the auditory nerve processor is coupled with two different algorithms for the determination of the interaural latencies. The first algorithm contains an extended correlation method [3] and the second one a subtraction method of the monaural signals [4].

The interaural time delay at the extreme values is plotted in Subplot c. Because of the limited integration in the end of the correlation mechanisms the extreme values of the correlation functions show a fluctuation of interaural latencies. The distribution density function of these extreme values is calculated with the mean value and its standard deviation. Figure 3 shows the distribution density function for the 36 channels of the ERB filter bank. While the mean value indicates the position of the sound source in the horizontal plane the standard deviation could be transformed in an incident angle as a measure for the extension of the sound source. The average of all critical bands was calculated for the evaluation of the incident angle of sound source [14].

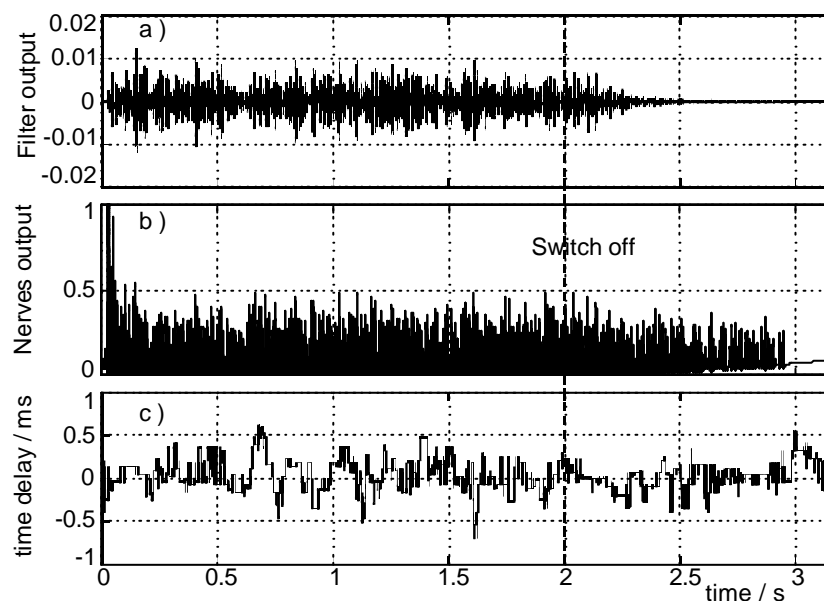


Fig. 2: Output of the different steps of the binaural hearing model on a noise burst of 2 s recorded in a enclosure (Reverberation time 1s). a): signal after band filtering; b): output of the auditory nerve module; c): result of correlation unit.

In Figure 3 it is obvious that the distribution density function shows the most fluctuations for the lowest 3 bands. In the region between 200 Hz and 800 Hz (6-12 ERB) the localisation process of the hearing models seem to be better than in the region between 900 Hz and 2KHz (15-22 ERB). The change of the fluctuations is a result of the phase locking process in the auditory nerve model. While for the low frequencies the correlation mechanism can use the phase

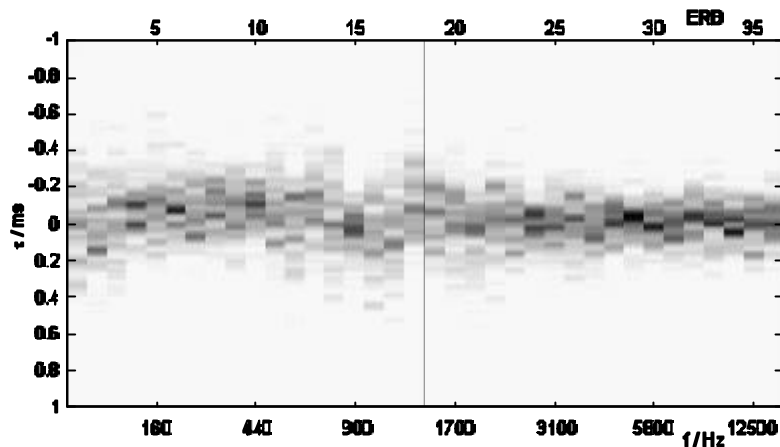


Fig. 3: Distribution density function for 36 critical Bands (ERB Scale), Linear gray scale; Dark high values. Top: ERB-Scale; bottom corresponding frequency scale .

information of the signal, for the higher frequencies this information is not longer available because the inertia of the inner ear and the signal detection has to be done on the envelope of the signal. This causes more fluctuations especially in the transition region between phase and envelope detection.

## LISTENING TEST

For the listening test a recording was made in a room with rectangular shape (20m length, 12m width, 5 m height). The room has a broadband reverberation time of 0.75 s. The reverberation time reaches for lower frequencies a value of 1.4 s. The recordings have been made for two positions in this room. The broadband IACC for the first positions has a value of 0.4 and for the second position a value of 0.65.

A listening study with these sound fields was made for different signals. For the examination of the frequency contribution a white noise signals was high passed filtered with different cut off frequencies (30, 100, 250, 500, 1000, 2000 Hz). For the evaluation of the temporal structure of the signal the broadband noise was pulsed with different durations (0.25s, 0.5s, 1s, 2s). The non signal duration was the length of the signal duration with exception for the 2 s burst were the non signal period was 1 s . The signal level was chosen to 70dB SPL for the broadband noise. The power spectral density for all noises was constant which lead to a lower signal level for the high pass filtered signals.

For the tests 8-10 test persons were instructed in the meaning of ASW. With the help of a virtual acoustic pointer the test persons had to mark the left and right edge of the sound source. The pointer consists of an endless repeated two tone click signal which could be moved with a DSP-System in real time in the horizontal plane. The angle between the pointer positions was noticed as the incident angle of the sound sources [15].

## RESULTS

Figure 4 shows the results of the incident angle for the different signals in combination with the outcome of two hearing models with different localisation methods. For the results of the listening test the mean value and the standard deviation is plotted with the solid lines. The

output of the first hearing model using an cross correlation method according to [3] is added with dashed line. The outcome of the second model, which uses a subtraction method for the calculation of the signal position [4], is displayed with the dash dotted line. Similar to the publications mentioned in the introduction [5]..[9] auditory source width decreases with an increasing cut off frequency of the high pass filter. Signals with low frequency components seem to be broader than signals without. The perception can also be evaluated with both hearing models. This behaviour could be explained with the high fluctuations especially for the first three bands of the models. For signals with cut off frequencies in the region between 100Hz and 1kHz the hearing models especially for the left measurement position (Figure 4a) seem to overestimate the ASW. In order to align the outcome of the models with human perception different weighting function as mentioned in [16] should be discussed. Even the weighting of different interaural time delays is discussed in literature [17] and could be applied for this purpose.

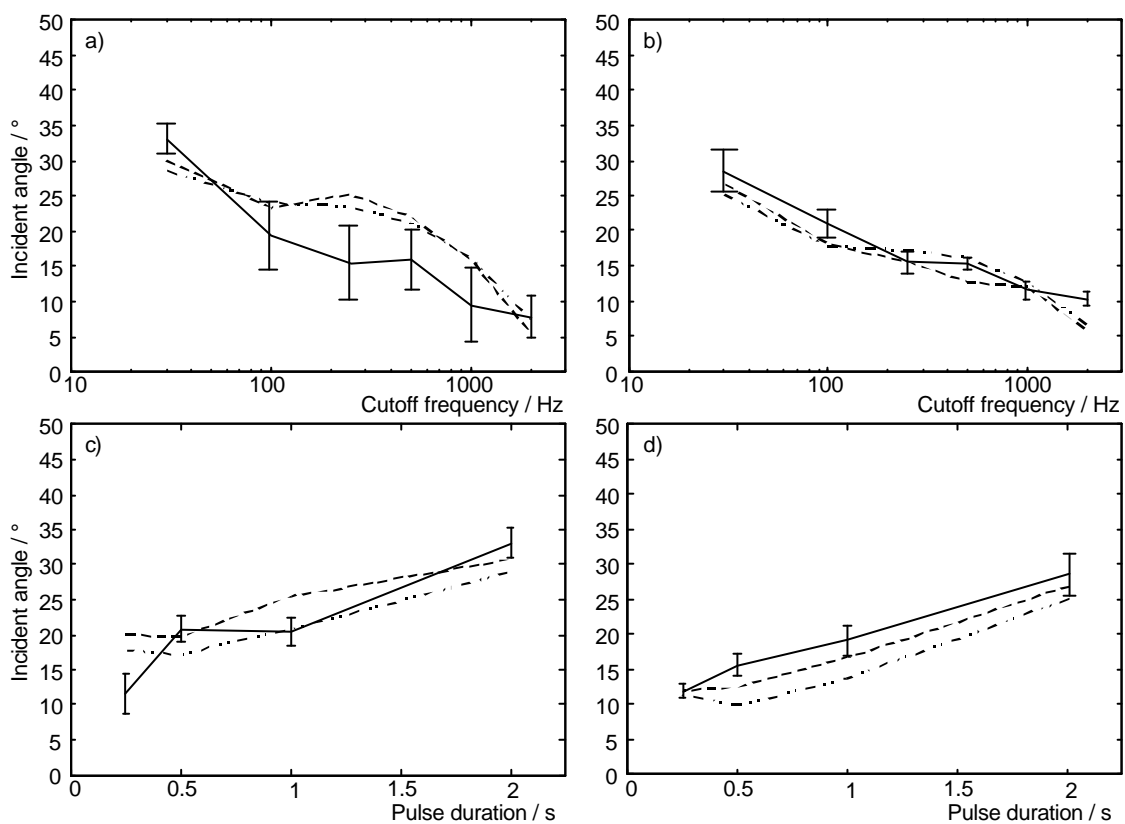


Fig. 4: a),b) Listening test result for noise signals high pass filtered with different cut off frequencies for two measurement positions (solid line; mean values with standard deviations.). Dashed line hearing model output with correlation mechanism according to [3]. Dash dotted line correlation mechanism according to [4]. c), d) listening test and hearing model results for different noise pulse durations for two measurement positions.

Figure 4 c) and d) show the influence of different pulse durations on the perception of ASW. It seems that short signal durations increase the ability of localisation and in the same way it decreases the perception of ASW. This could also be explained with both hearing models. The better ability to localize a signal in the beginning is shown with the help of the binaural model in figure [2]. In the region of the overshoot of the auditory nerves (Figure 2: b) the correlation fluctuations are less (Figure 2: c). The fact that the direct sound of the signal in the onset is only disturbed by the early reflections and not by the reverberation is also a reason for the better localisation possibility.

## CONCLUSIONS

The spectral and temporal composition of the signal has an influence on the perception of ASW. Like the influence of different signal levels [14] these contributions can be explained with the help of binaural hearing models. Especially the nonlinear processing of the mechanical to neural transduction in the auditory nerve seems to be responsible for the temporal and level dependence of ASW. Since this behavior hardly can be explained with physically measurable linear quantities ASW should be played back with a fixed signal at a certain level. Although it is unlikely that the enclosure delivers non linear contributions to the sound field for listening test it has to be considered that the enclosure interacts with the non linear human sense of hearing.

## BIBLIOGRAPHICAL REFERENCES

- [1] Griesinger, D., The Psychoacoustics of Apparent Source Width, Spaciousness and Envelopment in Performance Spaces, *Acustica*, vol. 83, 1997, p. 721-731.
- [2] Jeffress, L. A., A Place Theory of Sound Localization, *Journal Comp. Physiol. Psych.*, 1948, p. 35-39.
- [3] Lindemann, W., Die Erweiterung eines Kreuzkorrelationsmodells der binauralen Signalverarbeitung durch kontralaterale Inhibitionsmechanismen, PhD. thesis, Ruhr-Universität Bochum, 1985. (in German).
- [4] Breebart, J., and van der Par, S. and Kohlrausch, A., A New Model for Binaural Signal Detection, *Proc. 16th Int. Congress on Acoustics, ICA98*, vol. 2, 1998, p. 851-852.
- [5] Blauert, J. and Lindemann, W., Auditory Spaciousness: Some further Psychoacoustic Analysis, *J. Acoust. Soc. Am.*, vol.80, no.2, 1986, p.533-542.
- [6] Blauert, J. and Möbius, U. and Lindemann, W., Supplementary Psychoacoustic Results on Auditory Spaciousness, *Acustica*, Vol. 59, 1986, p. 292-293.
- [7] Morimoto, M. and Maekawa, Z., Effects of Low Frequency Components on Auditory Spaciousness, *Acustica*, Vol. 66, 1988, p. 190-196.
- [8] Morimoto, M., The Relation between Spatial and Cross-Correlation Measures, *Proc. on the 15<sup>th</sup> Int. Congress on Acoustics, ICA 95*, Vol. 2, 1995, p. 581-584.
- [9] Morimoto, M. and Yamanaka, T. and Sakagami, K. and Iida, K., Contribution of Low Frequency Components of Direct Sound and Lateral Reflections to Auditory Source Width, *Proc. on the 15<sup>th</sup> Int. Congress on Acoustics, ICA 95*, Vol. 2, 1995, p. 585-588.
- [10] Keet, W., The Influence of Early Lateral Reflections on Spatial Impression, *Proc. 6th Int. Congress on Acoustics, Tokyo, Japan, 1968*, p. E53-E56.
- [11] Hudde, H. and Engel, E. Measuring and Modeling Basic Properties of the Human Middle Ear and Ear Canal. Part III: Eardrum Impedances, Transfer Functions and Model Calculations, *Acustica*, vol. 84, 1998, p.1092-1109.
- [12] Moore, B.C.J. and Glasberg, B.R., Formulae Describing Frequency Selectivity as a Function Frequency and Level, and their Use in Calculating Excitation Patterns, *Hearing Research*, vol. 28, 1987, p. 209-225.
- [13] Meddis, R. Simulation of Auditory-Neural Transduction: Further studies, *J. Acous. Soc. Am.* vol 83, no. 3, 1988, p. 1056-1063.
- [14] Becker, J., Binaural Hearing Models for the rating of ASW in Room Acoustics, *Proc. on the 17<sup>th</sup> Int. Congress on Acoustics, ICA 01*, 2001.
- [15] Becker, J. and Sapp, M., Synthetic Soundfields for the rating of Spatial Perception, *Applied acoustics*, Vol.62, 2001, p. 217-228.
- [16] Akeroyd, M.A. and Summerfield, A.Q., A fully temporal account of the perception of dichotic pitches, *Br. J. Audiol.*, Vol. 33, no. 2, 1999, p. 106-107.
- [17] Colburn, H.S., Theory of binaural interaction based on auditory-nerve data. II. Detection of tones in noise, *J. Acoustic. Soc. Am.*, vol. 61, 1977, p. 525-533.