

Applied Psychophysics I

Applications in audio quality

Spatial hearing sensations

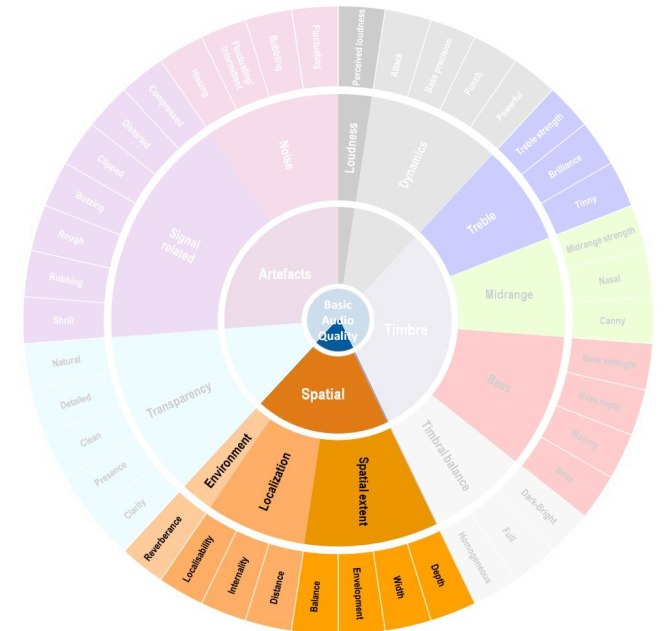
Stephan Töpken, 27.11.2024

Applied Psychophysics I – Lecture schedule

Week	Date	Topic
1	16.10.24	Introduction
2	23.10.24	Auditory physiology and selected auditory models
3	30.10.24	Suprathreshold scaling methods
4	06.11.24	Loudness perception, loudness models and audio metering
5	13.11.24	Selected sensations, sound attributes and measurement of timbre
6	20.11.24	Chocolate results Presentation topics
7	27.11.24	Spatial hearing sensations
8	04.12.24	Audio quality assessment methods and model approaches
9	11.12.24	Calibration of listening setups
10	18.12.24	Ethics and data protection
Winter break		
11	08.01.25	Student presentations
12	15.01.25	Student presentations
13	22.01.25	Student presentations
14	29.01.25	Student presentations

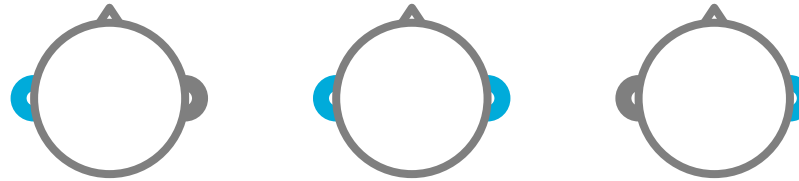
Spatial hearing sensations

- Basic Definitions
- Binaural cues
- Sensitivity to binaural cues
- Localization
- Selected spatial sensations in closed rooms

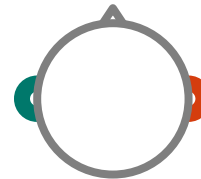


Basic definitions

Monaural hearing



Binaural hearing

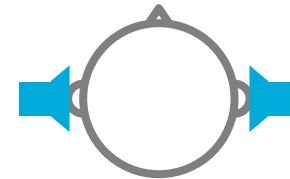


Basic definitions

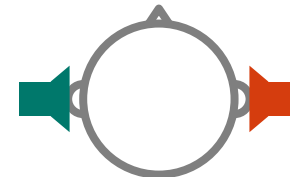
Monotic presentation



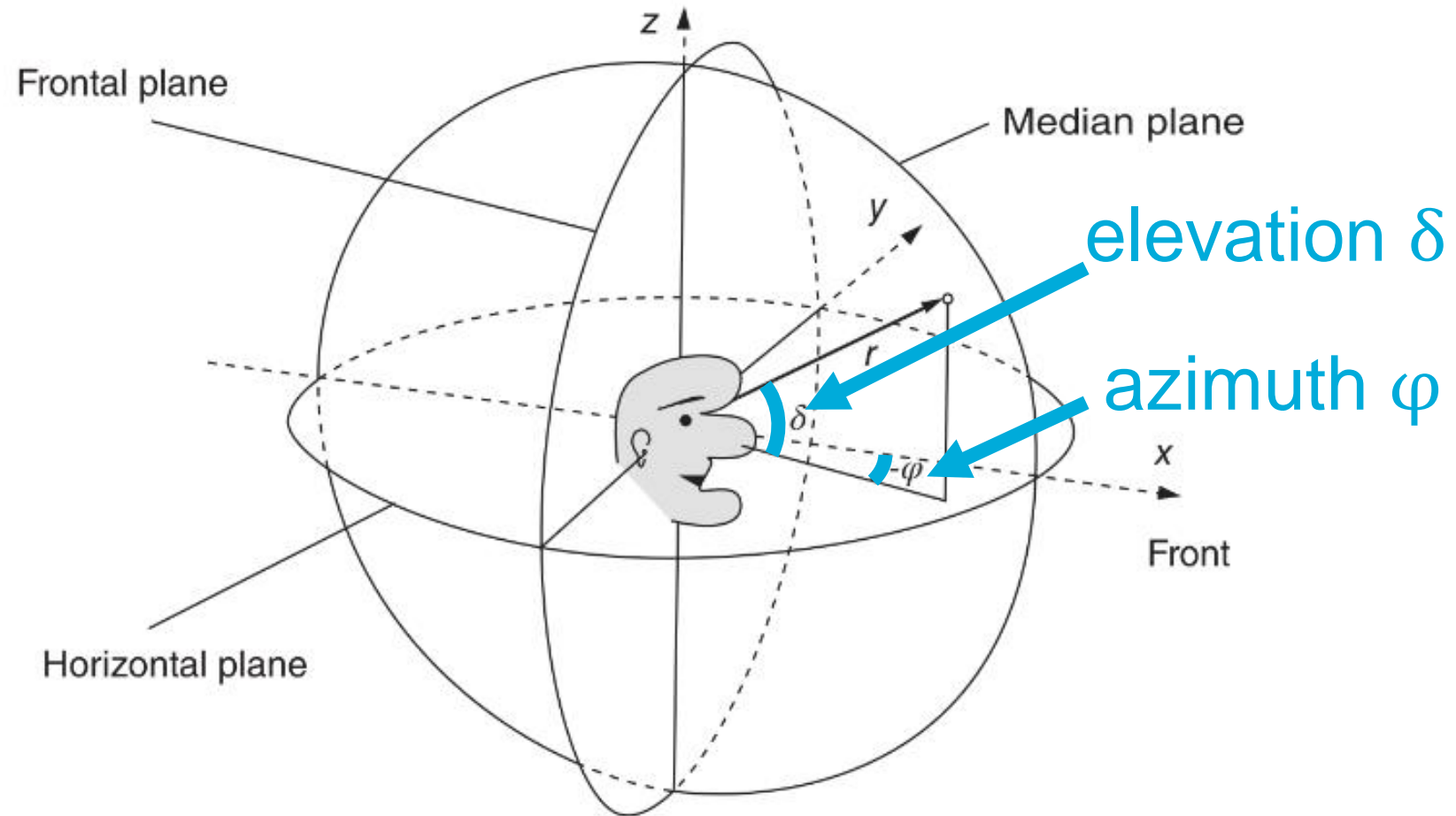
Diotic presentation



Dichotic presentation



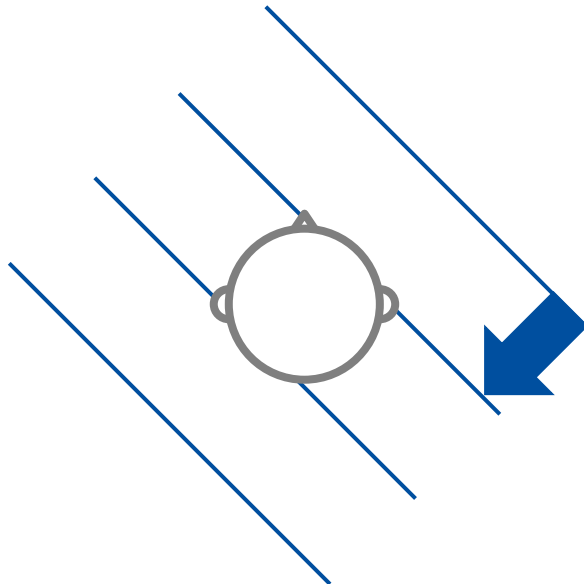
Head centered coordinate system for spatial hearing



Physical wave phenomena

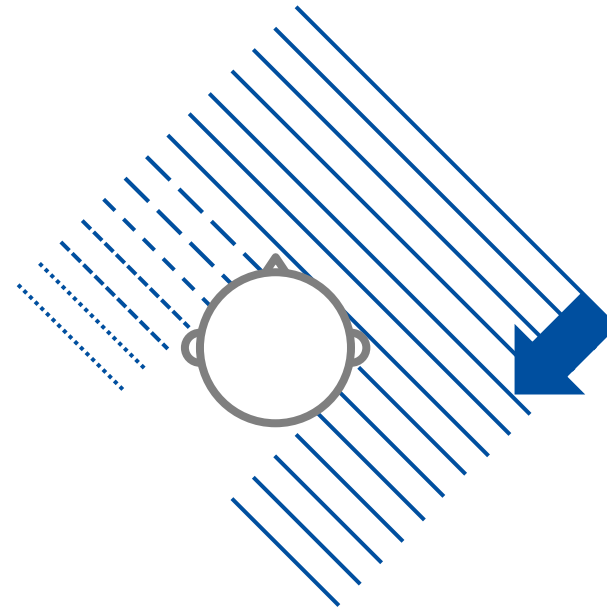
Diffraction around the head

when wave length larger than head



Head shadow and level reduction

when wave length smaller than head



Transition
 $f \approx 1 \dots 2 \text{ kHz}$
 $\lambda \approx 0.17 \dots 0.34 \text{ m}$

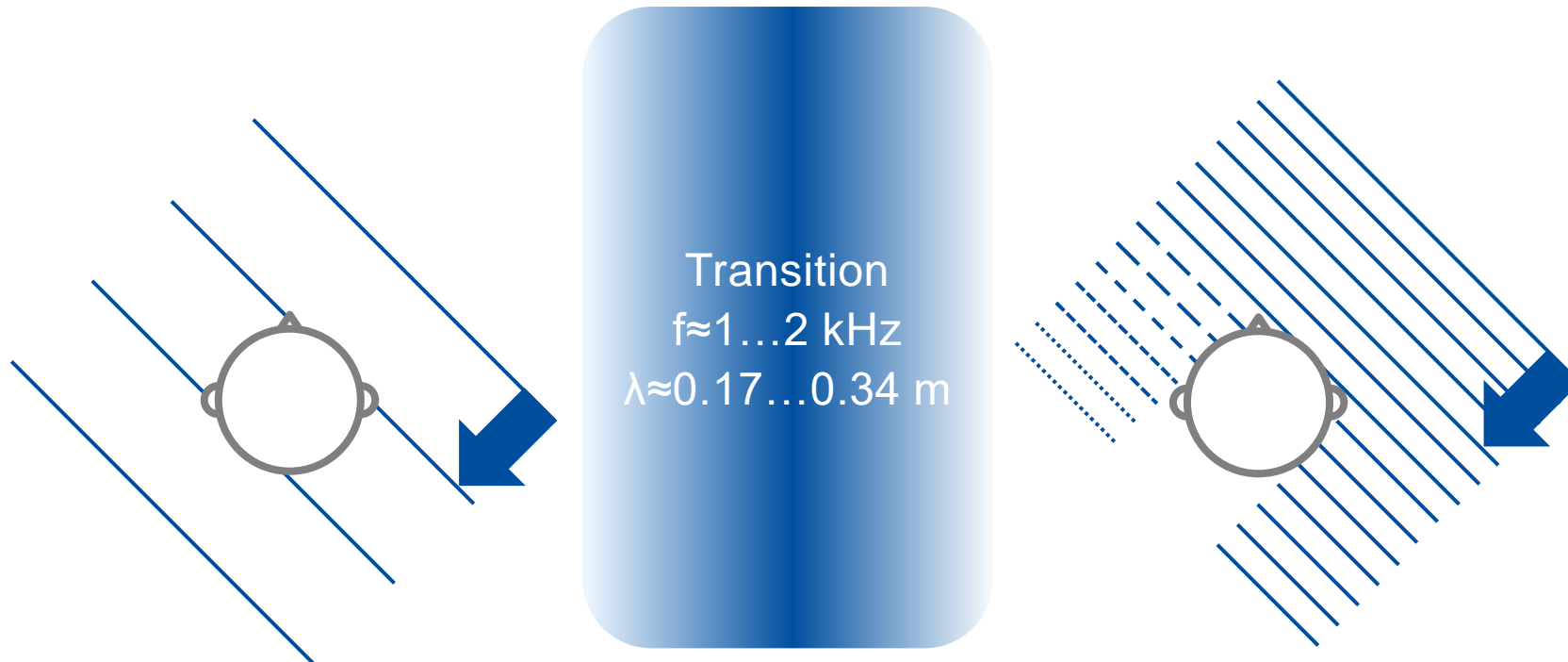
Pulkki, 2015

Available binaural cues

ITDs: Interaural Time Difference

IPD: Interaural Phase Delay

ILD: Interaural Level Differences



Pulkki, 2015

Influence of the head

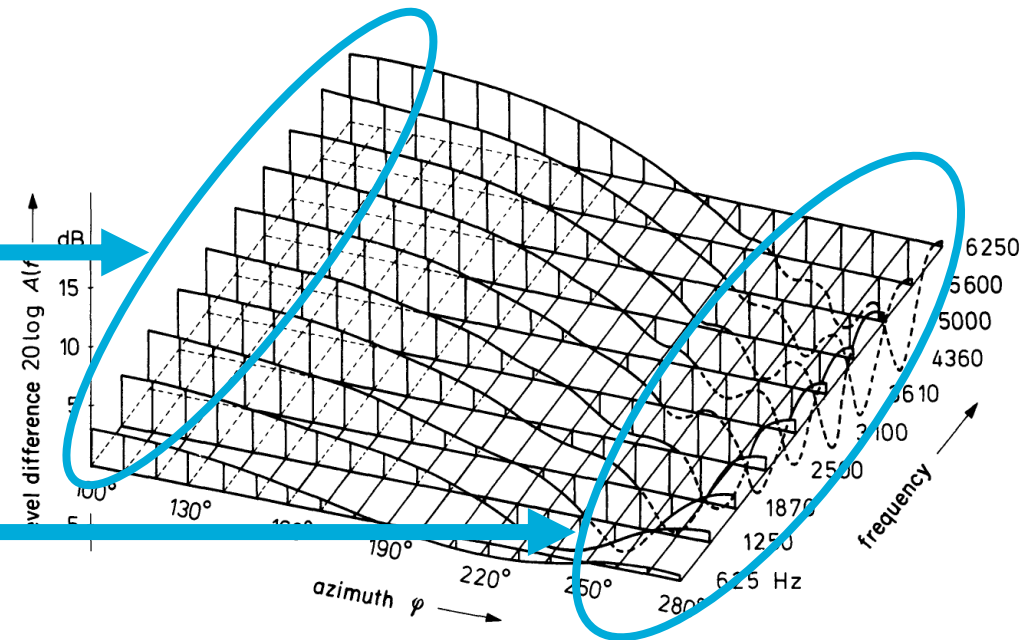
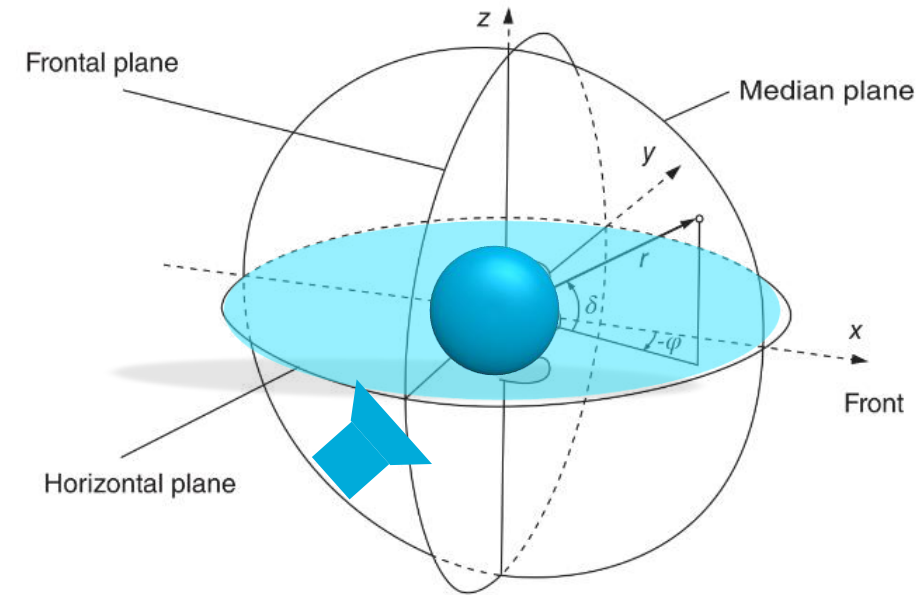
Level at left ear compared to no
sphere present as a function of
source incidence angle

Ear at azimuth $\varphi=100^\circ$,
elevation $\delta=0^\circ$ (middle of head)

Sphere 17.5 cm diameter

Amplification for waves directly
facing the ear

Complex diffraction pattern at
opposite side



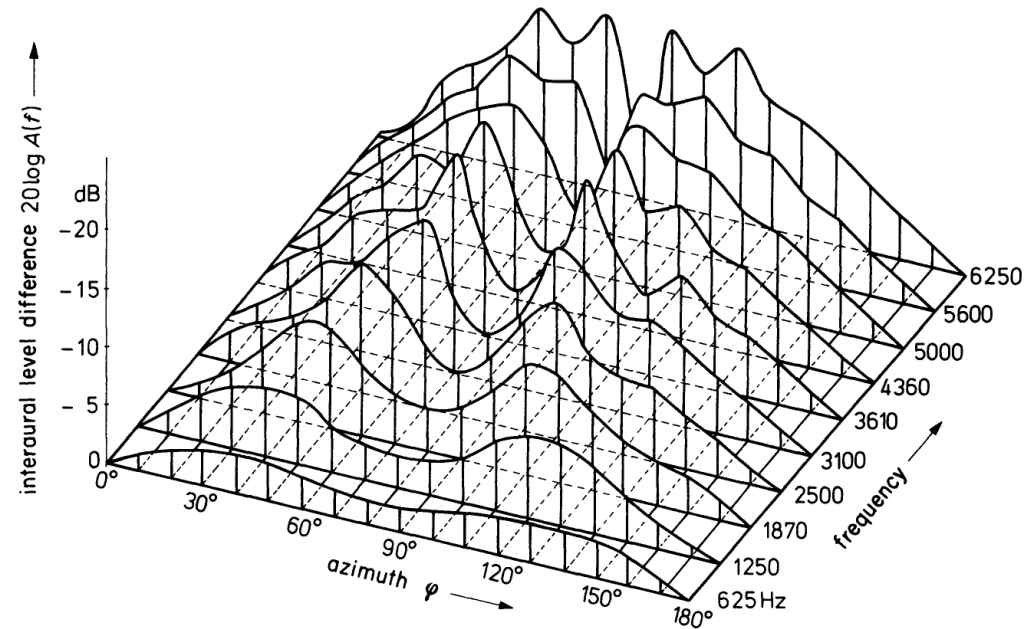
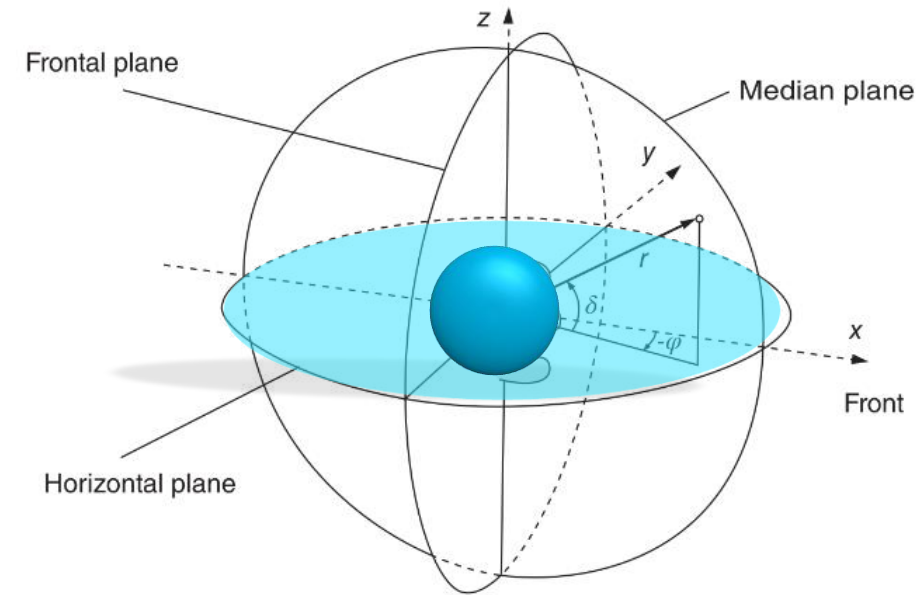
Interaural level differences (ILDs)

Interaural Level Difference (ILD)
as a function of source incidence
angle

Small dependence of ILD on
azimuth at low frequencies.

At high frequencies stronger
dependence.

ILD increases with azimuths up
to about 45° .



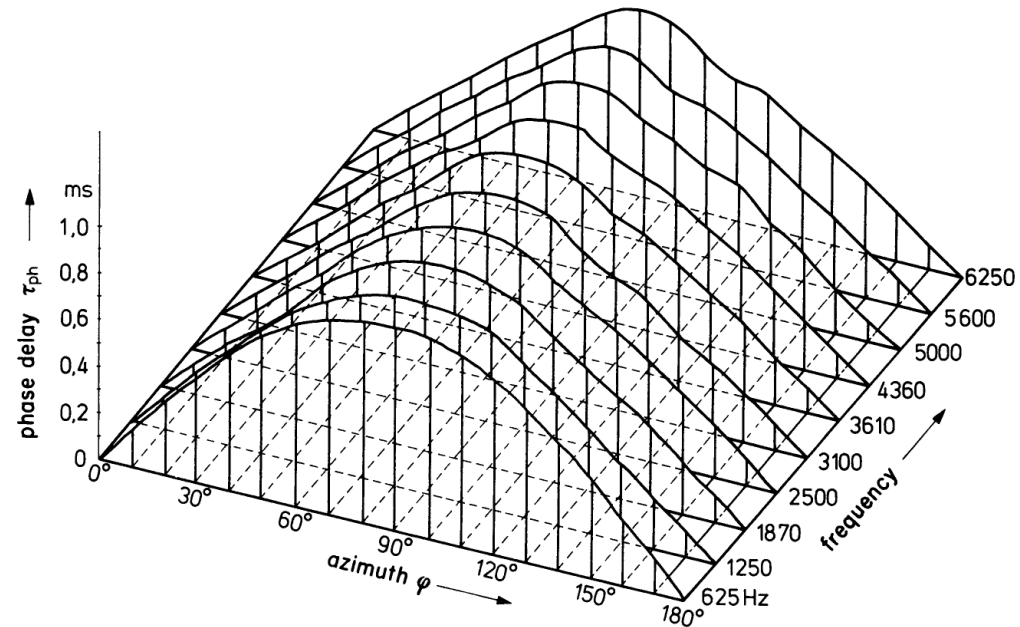
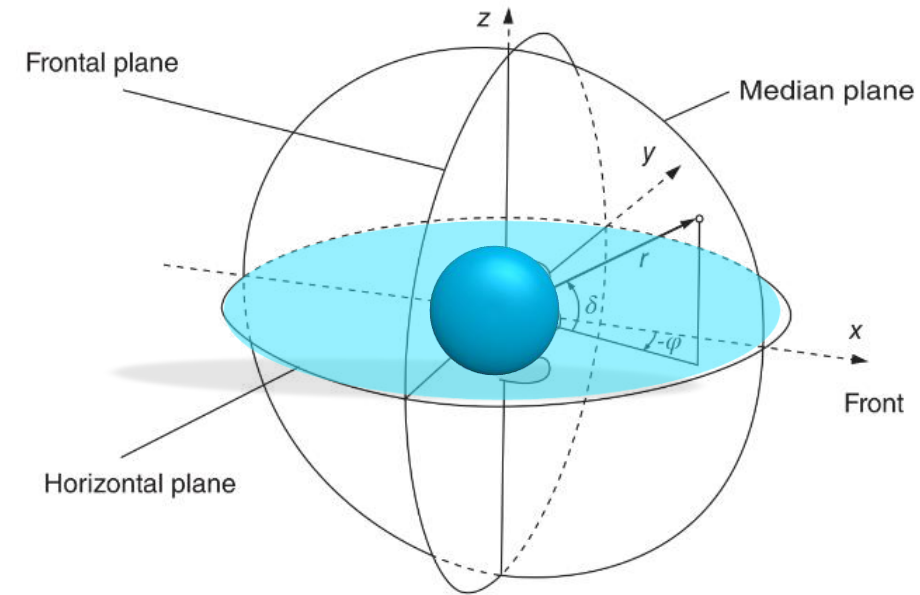
Blauert, 1996
Pulkki, 2015

Interaural phase delay (IPDs)

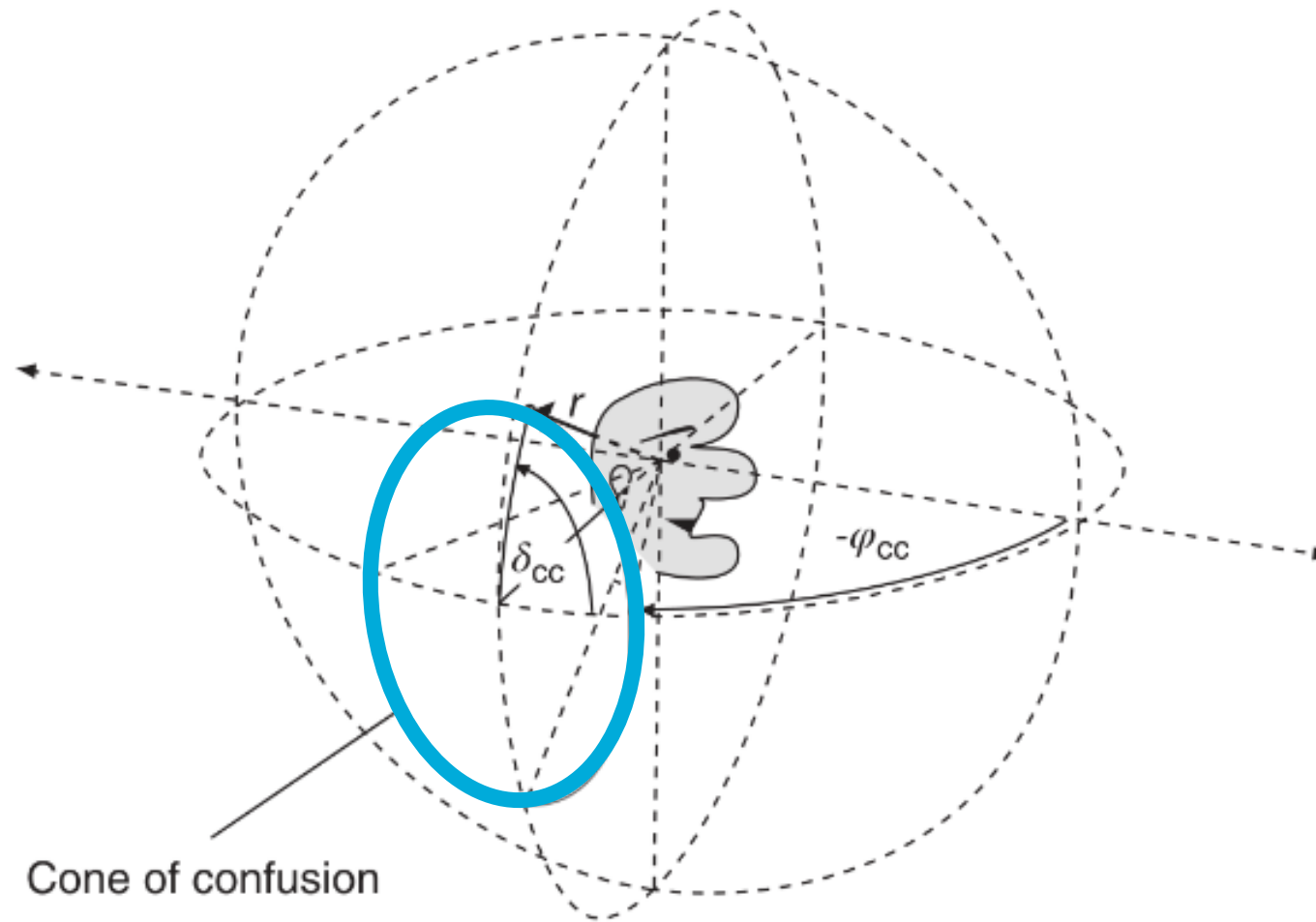
Interaural Phase Delay (IPD) in ms
as a function of source incidence
angle

Nearly uniform dependence of
azimuth on IPD independent of
frequency.

Note that at 5 kHz,
a phase delay of 1 ms
corresponds to 5 periods of a
sinusoid:
One cannot infer direction of
sound from phase then.



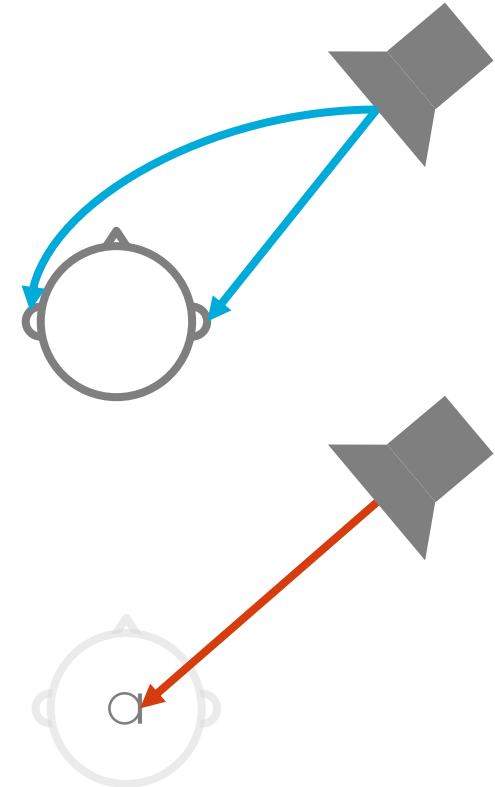
Cone of confusion: ITDs and ILDs are constant



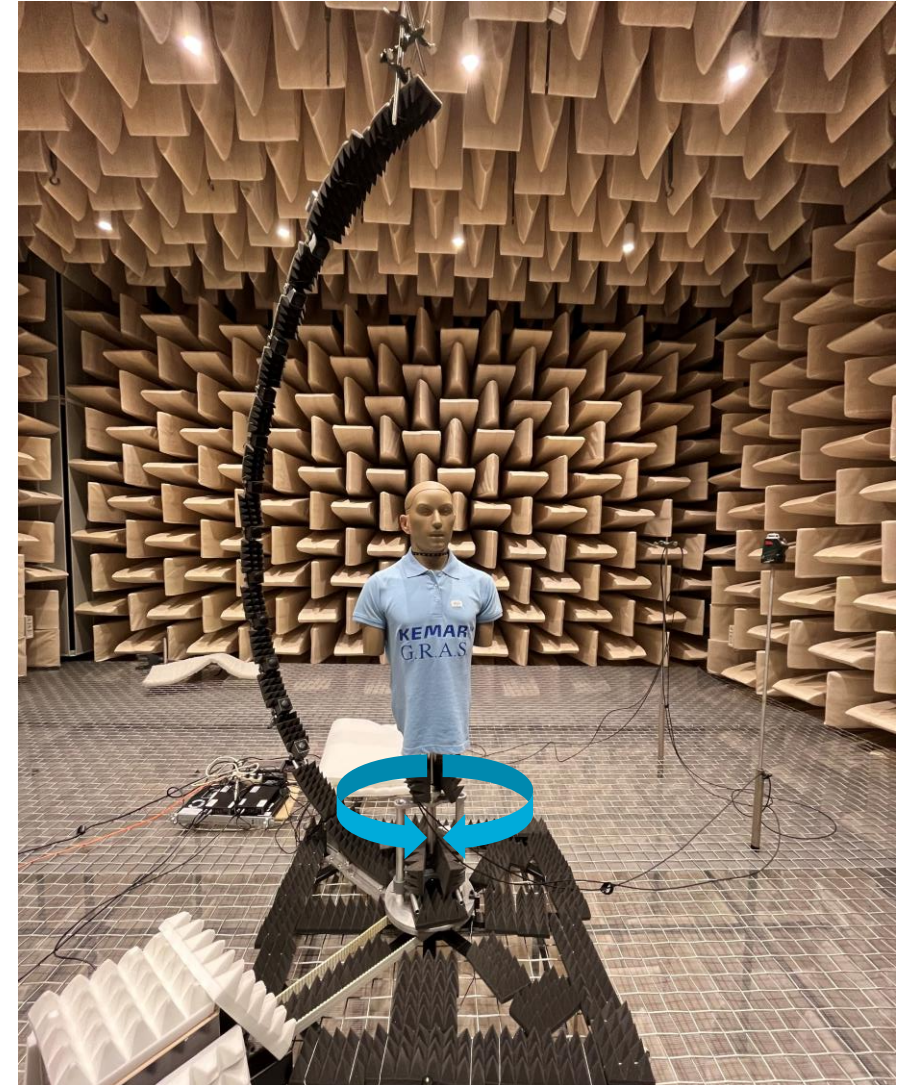
Head related transfer functions (HRTFs)

$$H_R(\varphi, \delta, f) = \frac{H_R(\varphi, \delta, f)}{H_{ref}(f)}$$

$$H_L(\varphi, \delta, f) = \frac{H_L(\varphi, \delta, f)}{H_{Ref}(f)}$$



Measurement of head related transfer functions (HRTFs) here: for an artificial head



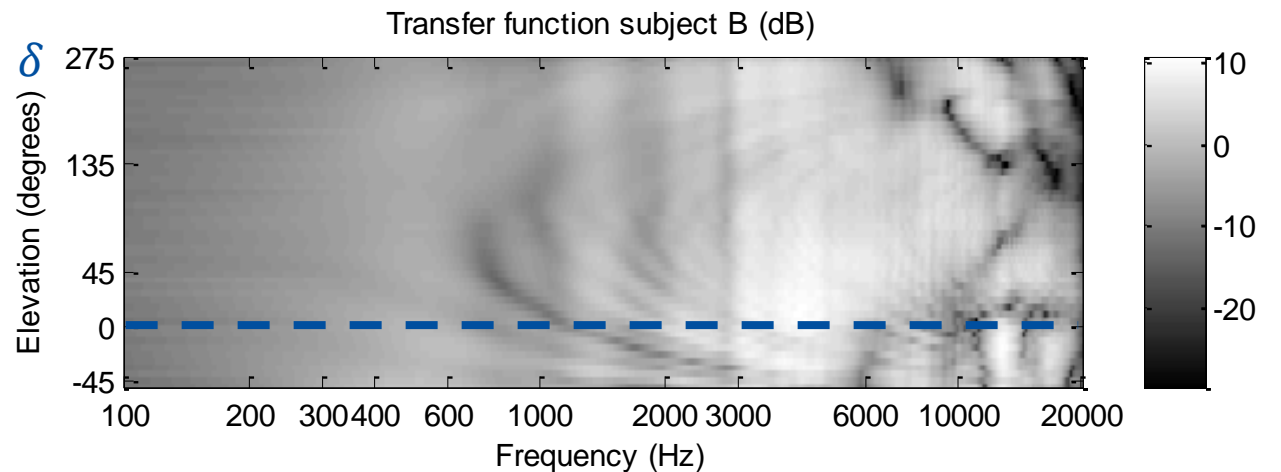
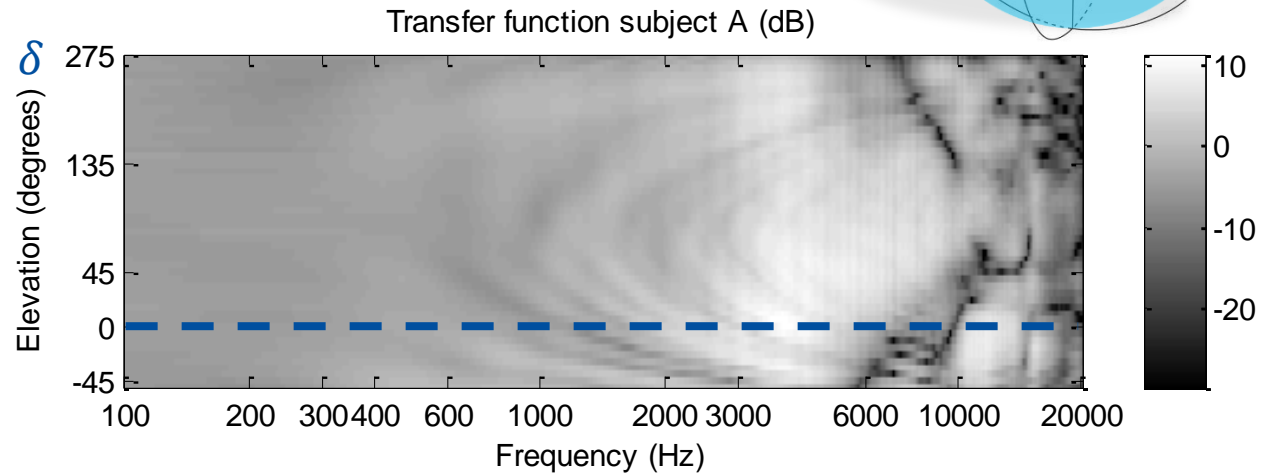
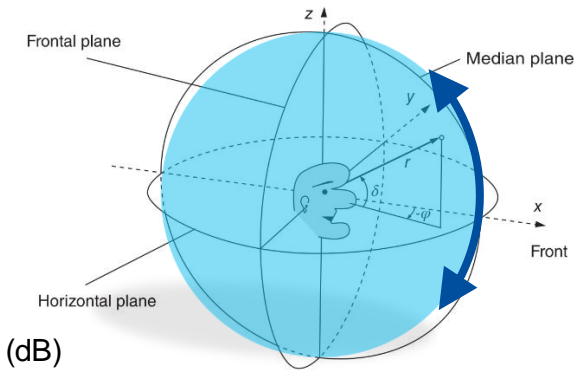
Head related transfer functions (HRTFs)

HRTF on right ear as a
function of **elevation**:

$$20 \cdot \log_{10} |H_R(\varphi = 0, \delta, f)|$$

HRTF reveals a strong
dependence on elevation
above 6 kHz

Position valleys depends on
elevation



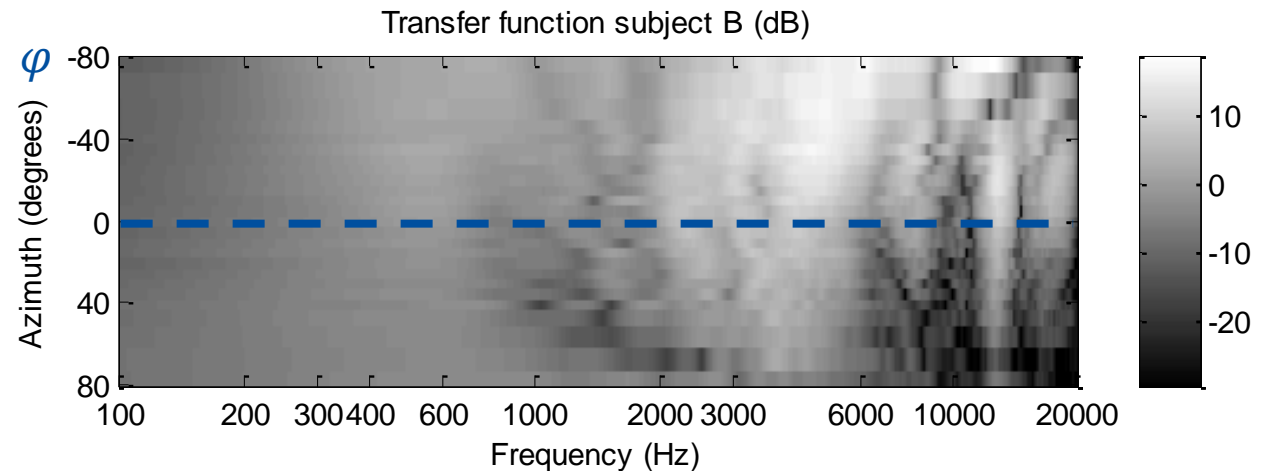
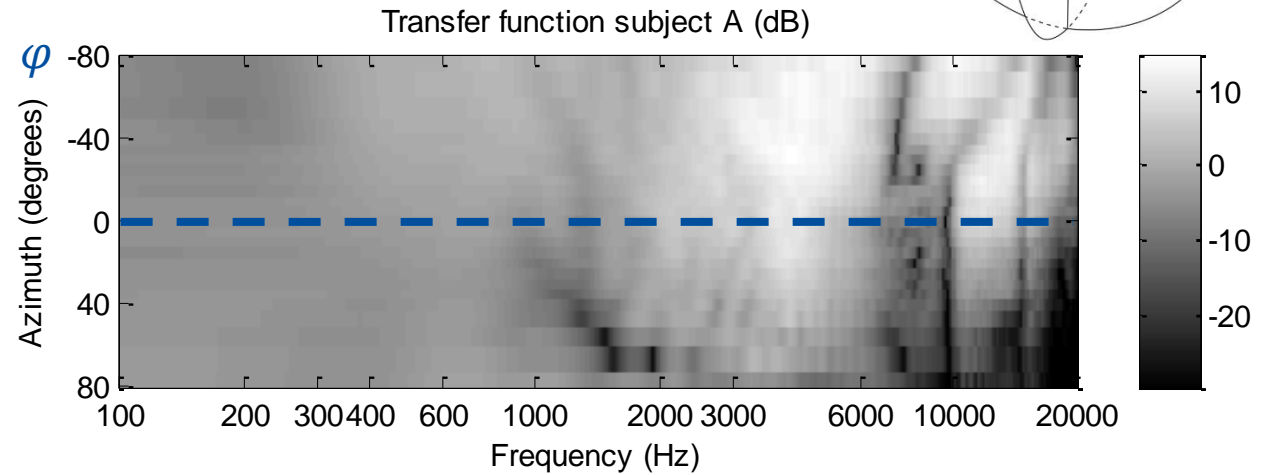
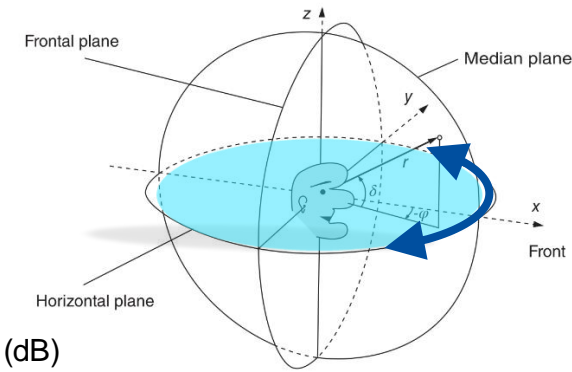
Head related transfer functions (HRTFs)

HRTF on right ear as a
function of azimuth:

$$20 \cdot \log_{10} |H_R(\varphi, \delta = 0, f)|$$

HRTF reveals a strong
dependence on azimuth
above 6 kHz

Position valleys mostly
independent of azimuth

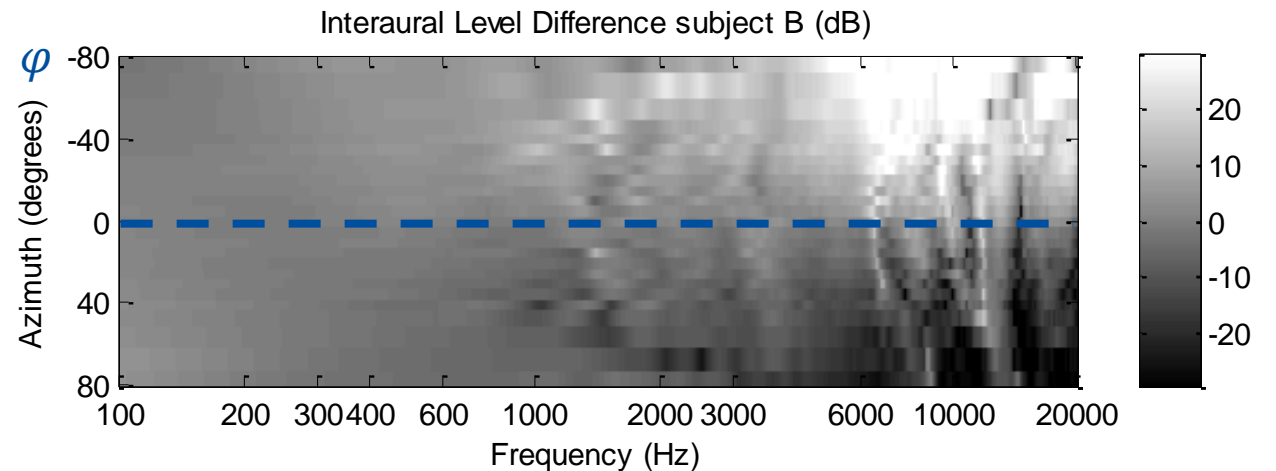
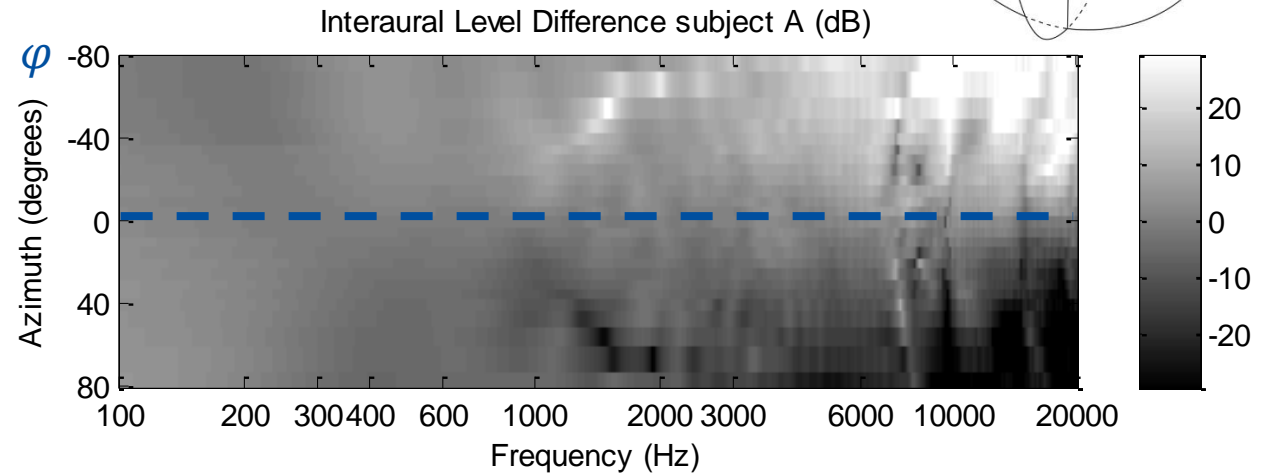
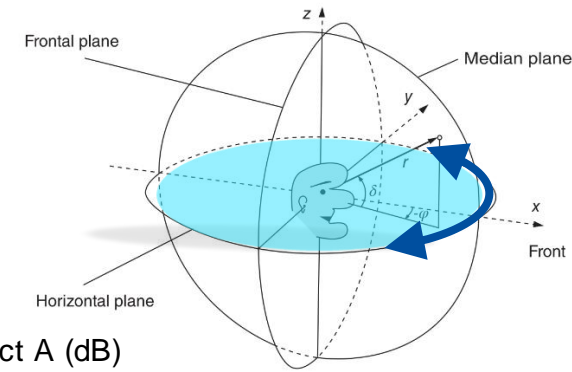


Interaural level differences (ILDs)

Interaural Level Difference
as a function of **azimuth**:

$$20 \cdot \log_{10} \frac{|H_R(\varphi, \delta = 0, f)|}{|H_L(\varphi, \delta = 0, f)|}$$

ILD depends strongly on
azimuth above 6 kHz

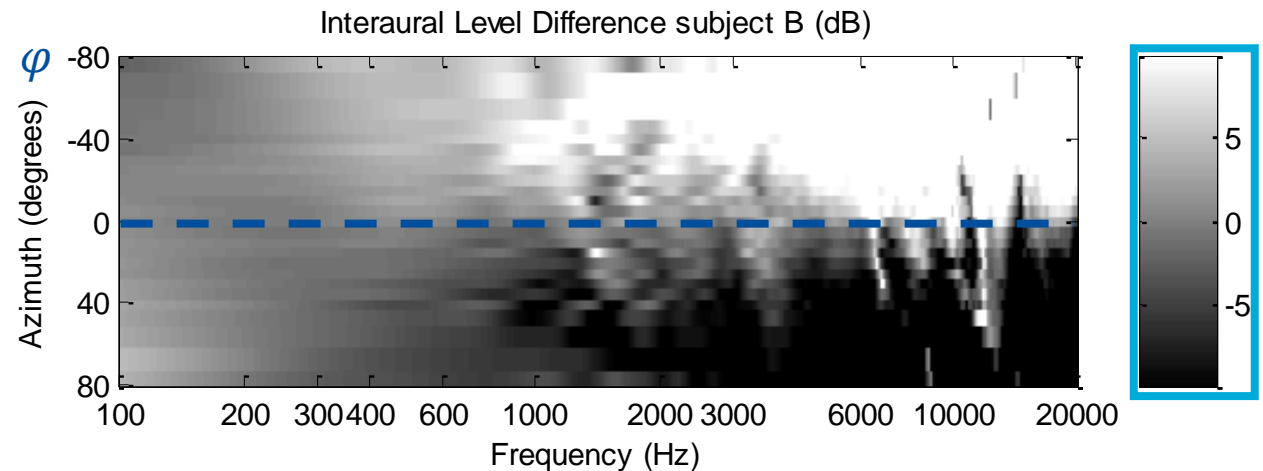
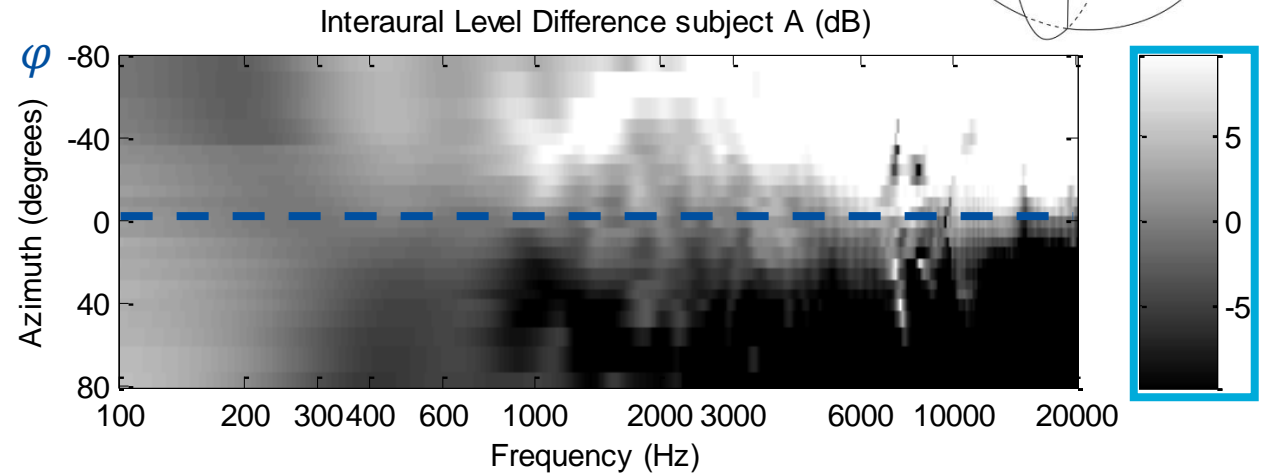
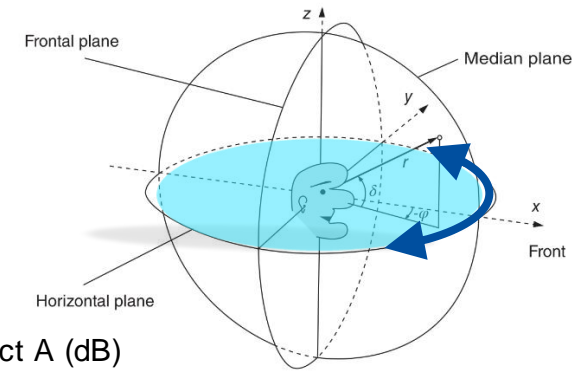


Interaural level differences (ILDs)

Interaural Level Difference
as a function of **azimuth**:

$$20 \cdot \log_{10} \frac{|H_R(\varphi, \delta = 0, f)|}{|H_L(\varphi, \delta = 0, f)|}$$

ILD **already** depends on
azimuth **above about 600 Hz**

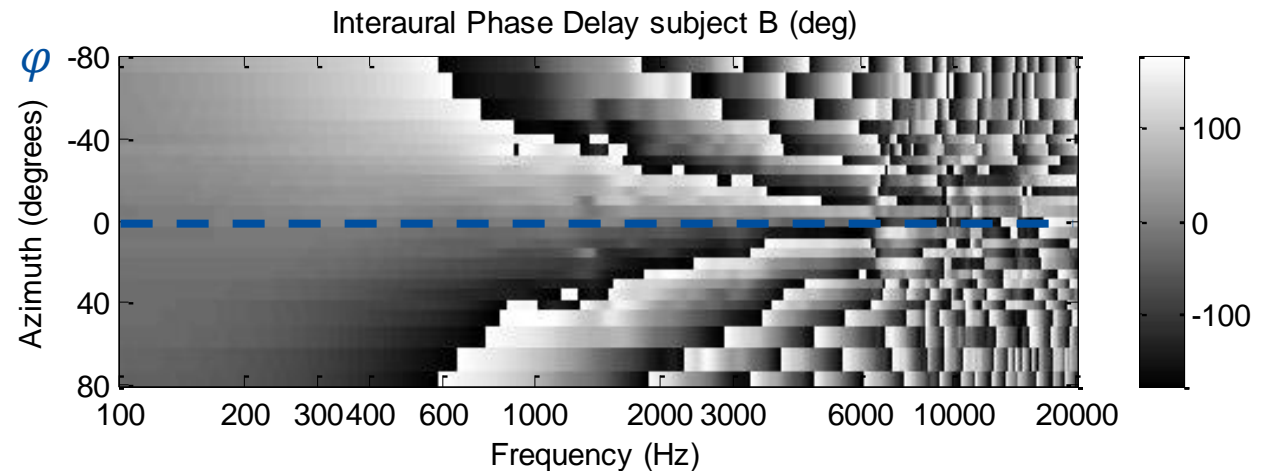
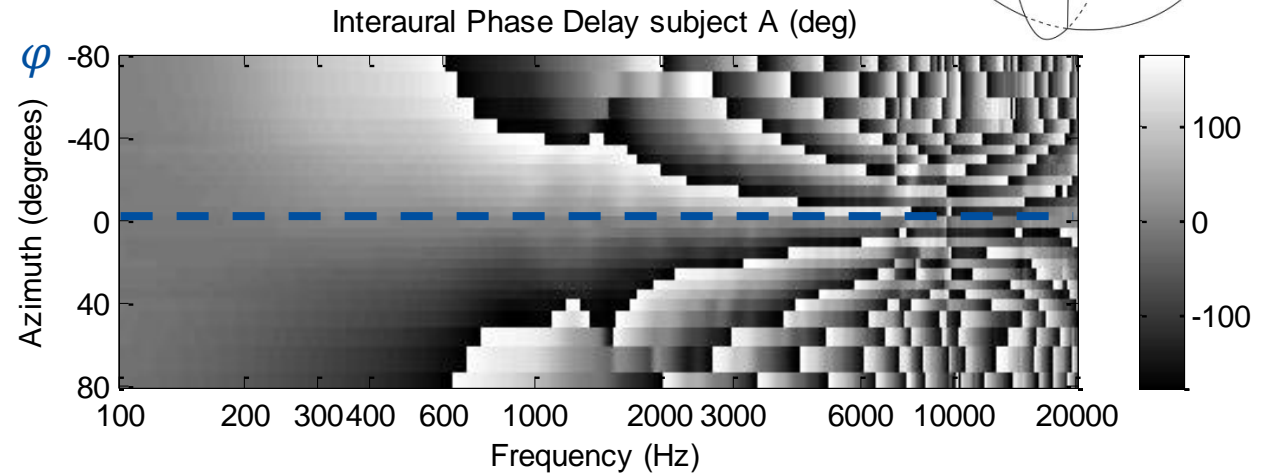
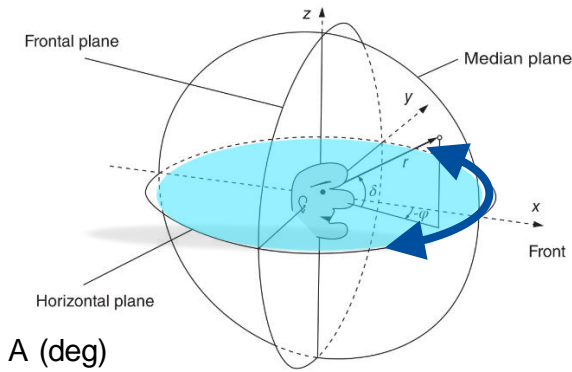


Interaural phase delays (IPDs)

Interaural Phase Delay as a
function of **azimuth**:

$$\arg \frac{|H_R(\varphi, \delta = 0, f)|}{|H_L(\varphi, \delta = 0, f)|}$$

IPD is unambiguous up to
about 600 Hz

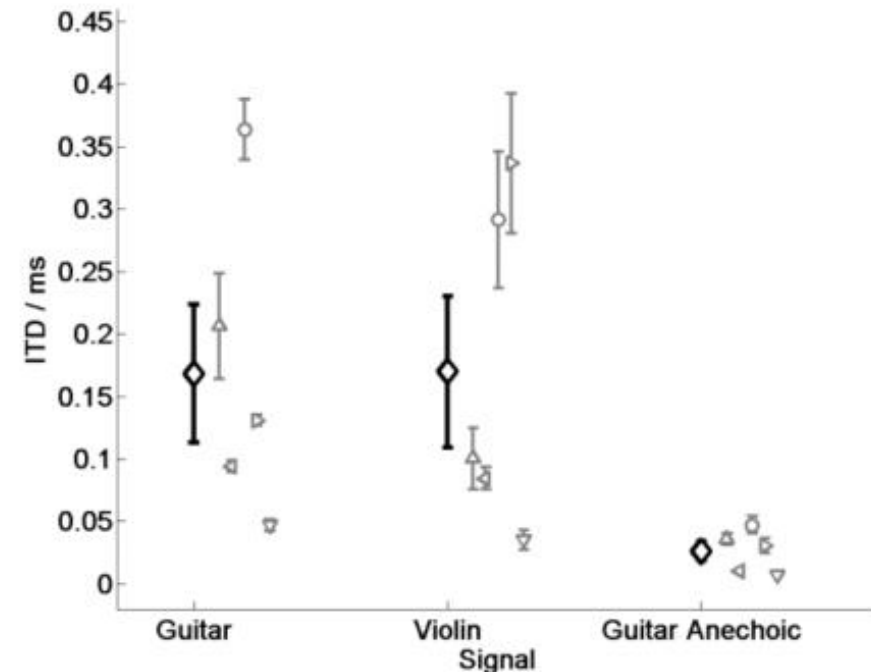


Humans sensitivity for Interaural time differences (ITDs)

- up to frequencies of about 1.5 kHz for sinusoids
- Envelope ITDs at high frequencies
- Perceived as lateral shift
- JNDs of about 20 μ s at 500 Hz

Signal		Threshold in seconds	Number of listeners	Judgments per point per listener
Noise			10	160
			0	80
			9	160
			8	80
	2400–3100 cps	44	10	80
	3056–3344 cps ^b	62	10	80
Clicks (1 millisecond duration)	Single	28	10	120
	Repeated, 30 clicks in a 2 second burst	11	13	160
Tones	90 cps	75	10	10
	125 cps	56	9	80
	250 cps	27	9	80
	500 cps	17	9	80
	1000 cps	11	9	80
	1300 cps	24	10	10
	1500 cps	...	10	10
	1800 cps	...	10	10
	3200 cps	...	10	80

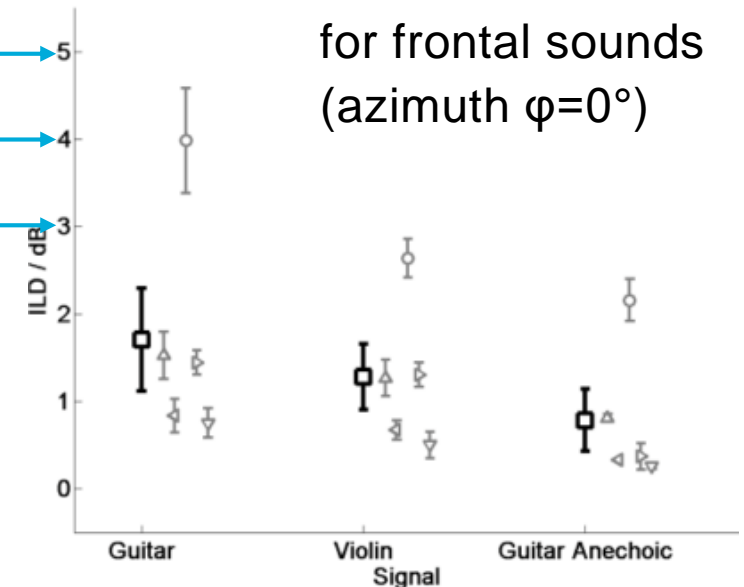
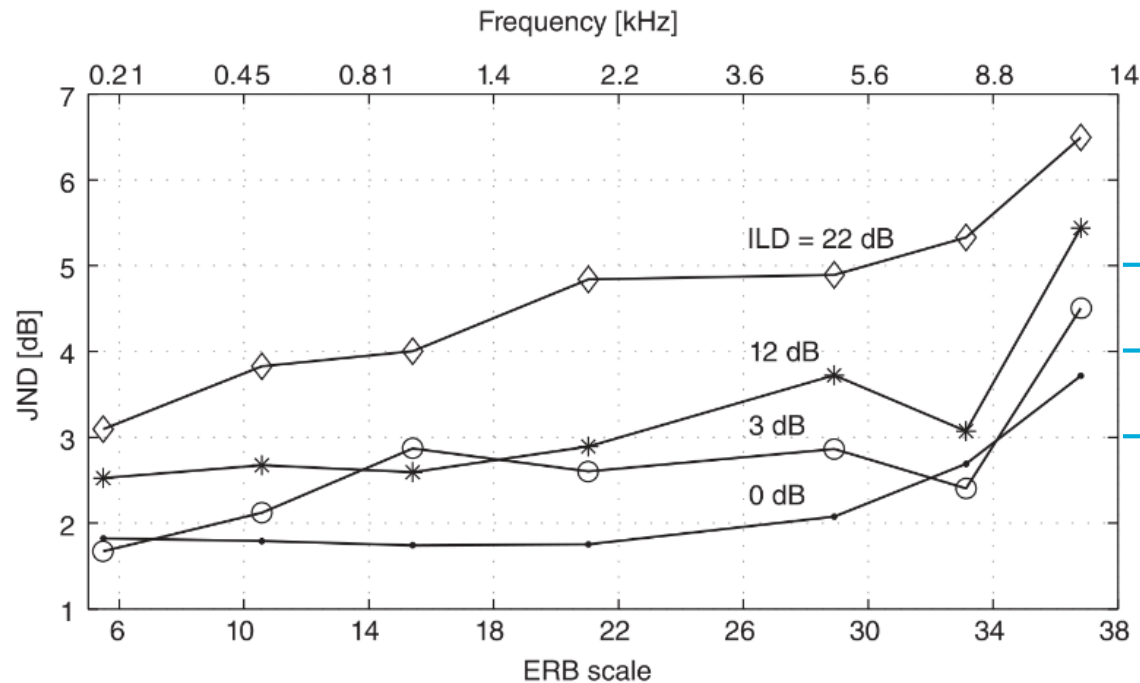
≈ 1 sample at
 $f_s = 44100$ Hz



Klumpp1956
Klockgether 2013

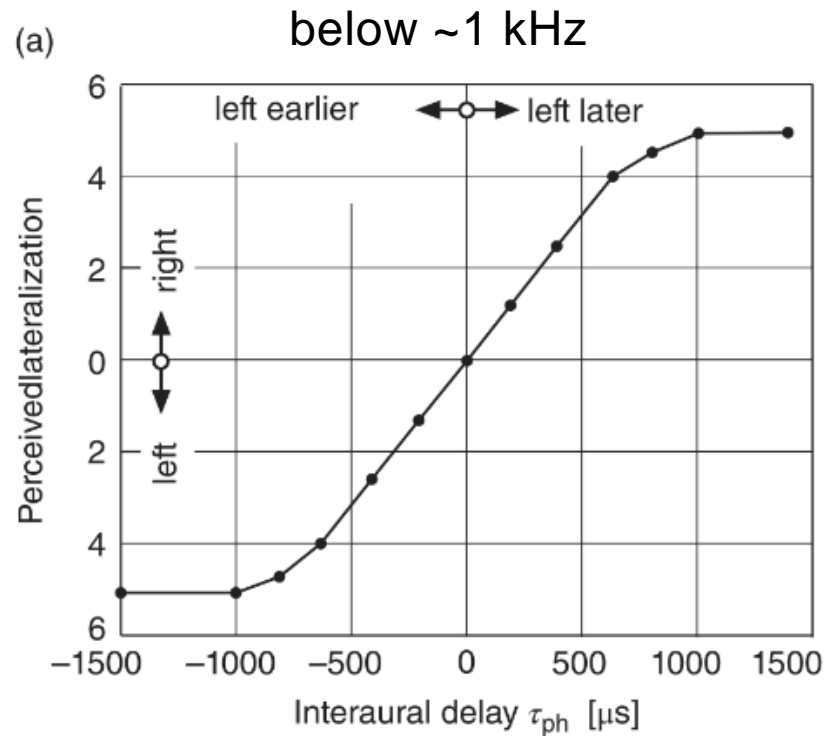
Humans sensitivity for Interaural level differences (ILDs)

- JNDs of about 1-2 dB across all frequencies for ILDs = 0 dB
- Perceived as lateral shift

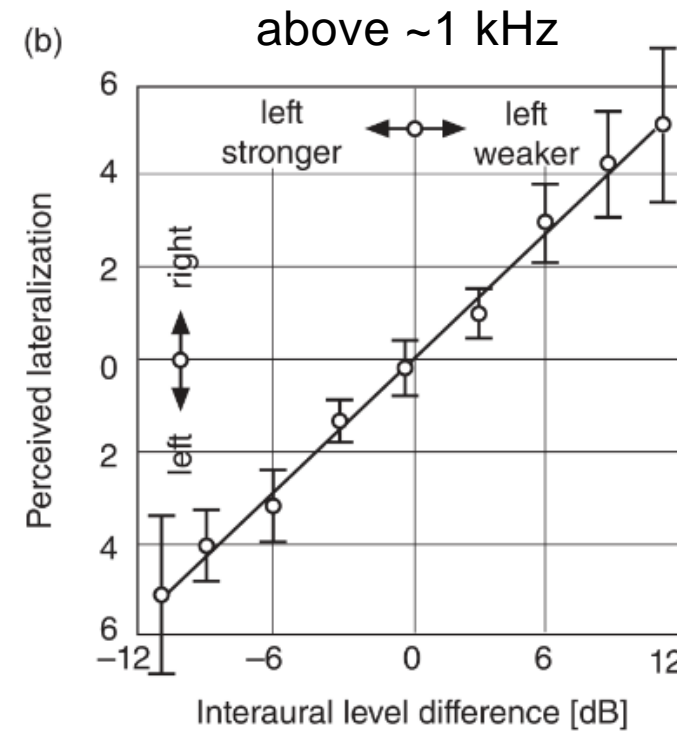


Pulkki, 2015 (adapted from Weiping)
Klockgether 2013

Perceived lateralization as a function of ITDs and ILDs



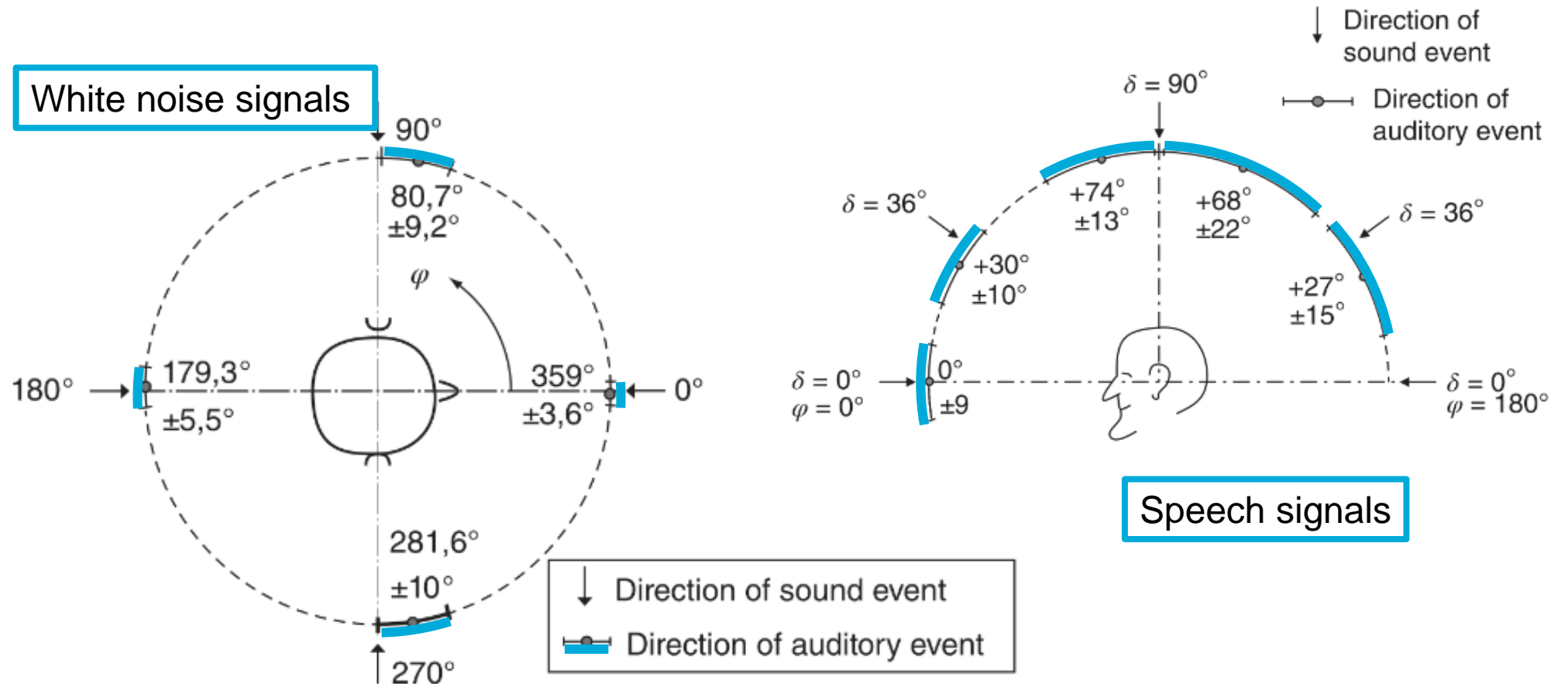
Maximum lateralization for
 $\tau \approx 700 \mu s$



ILD ≈ 12 dB

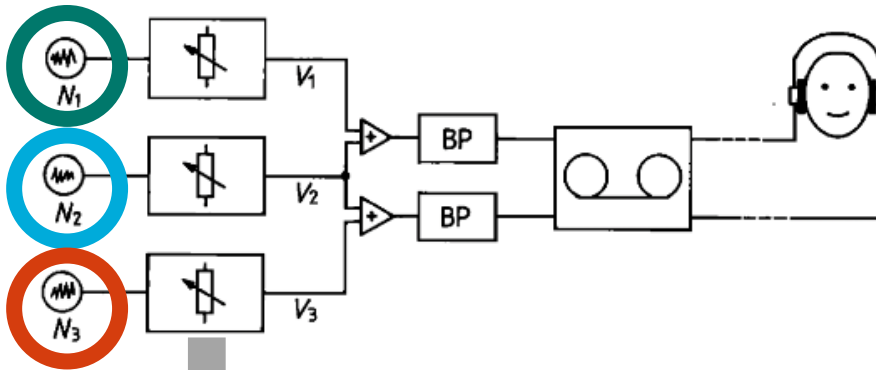
Pulkki, 2015 (adapted from Sayers)

Localization accuracy in azimuth and elevation



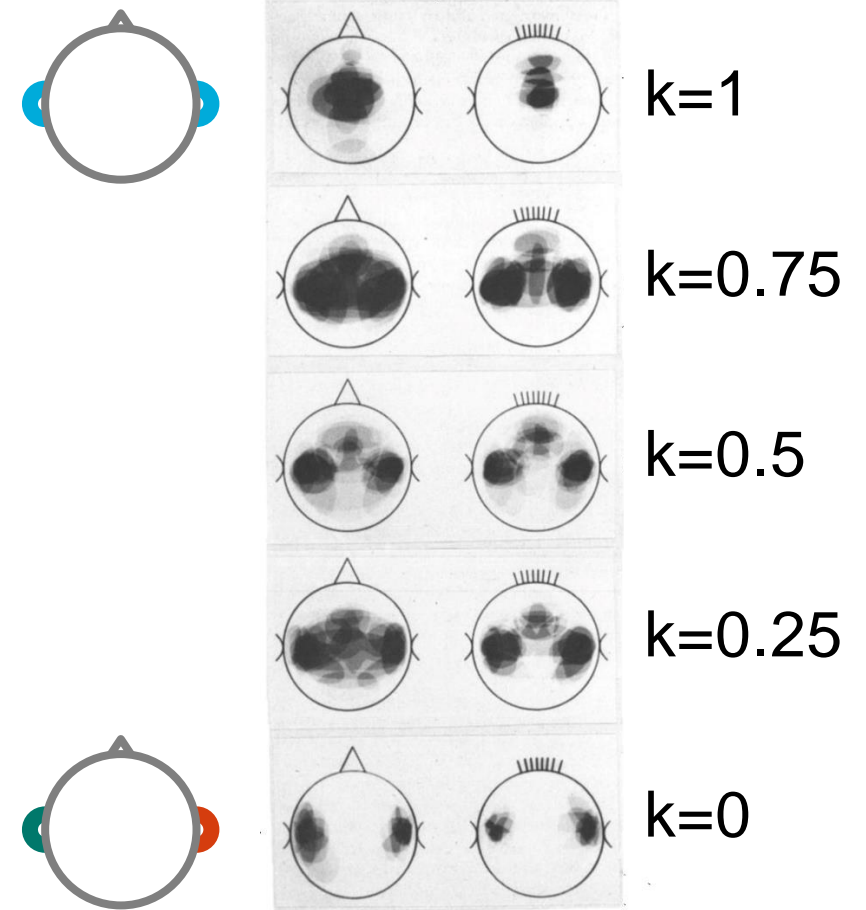
Pulkki, 2015 (adapted from Blauert
/ Damaske and Wagner)

Interaural coherence / cross correlation (IACC) and lateralisation



$$IACC(\tau) = \lim_{T \rightarrow \infty} \left(\frac{1}{2T} \int_{-T}^T l(t + \tau) r(t) dt / l_{rms} r_{rms} \right)$$

$$k = \max |IACC|$$

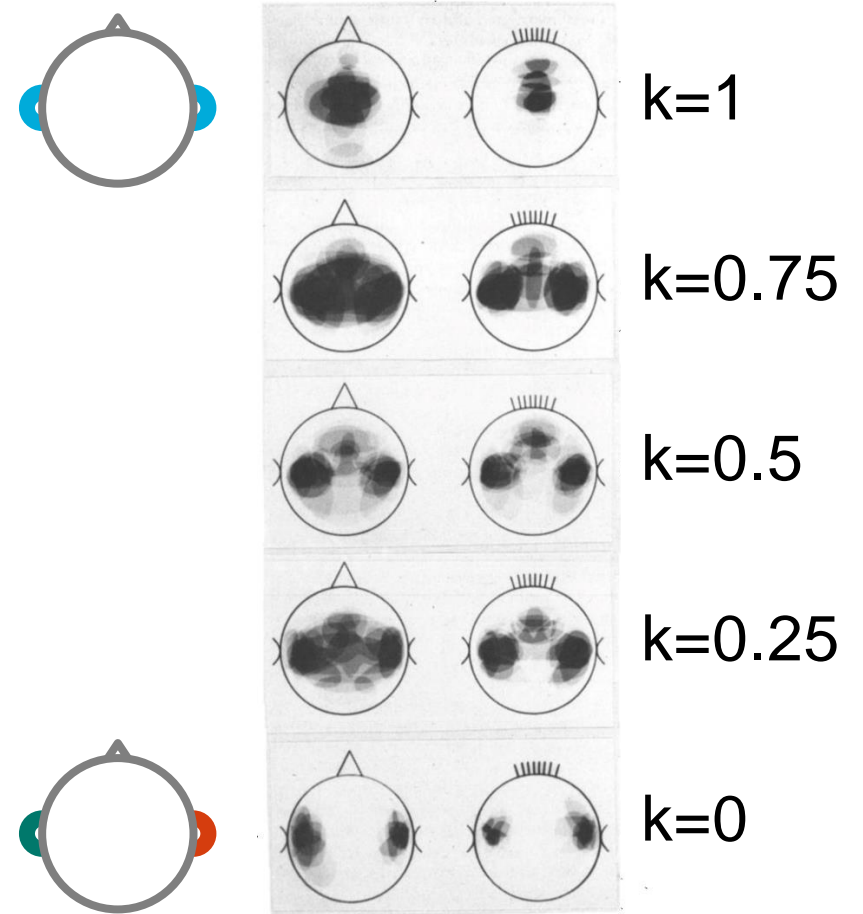


Blauert 1986

Humans sensitivity for IACC

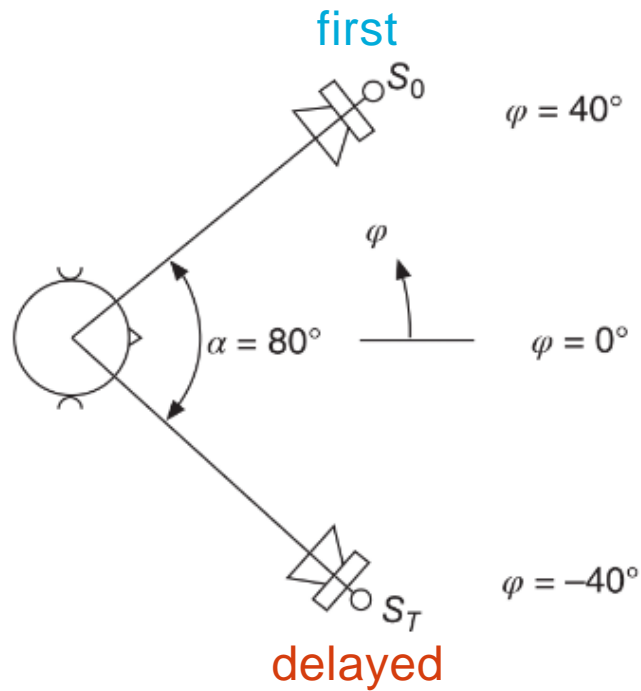
– $k \approx 1$: JNDs of about 2 %

– $k \approx 0$: JNDs of about 30 %

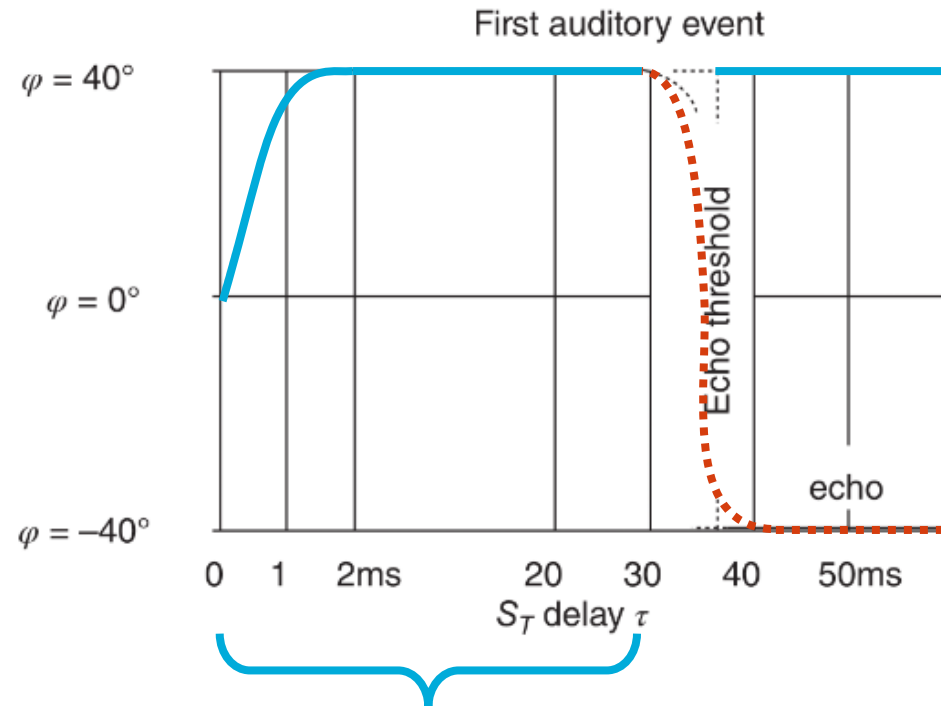


Blauert 1986

Perception first lateral reflections

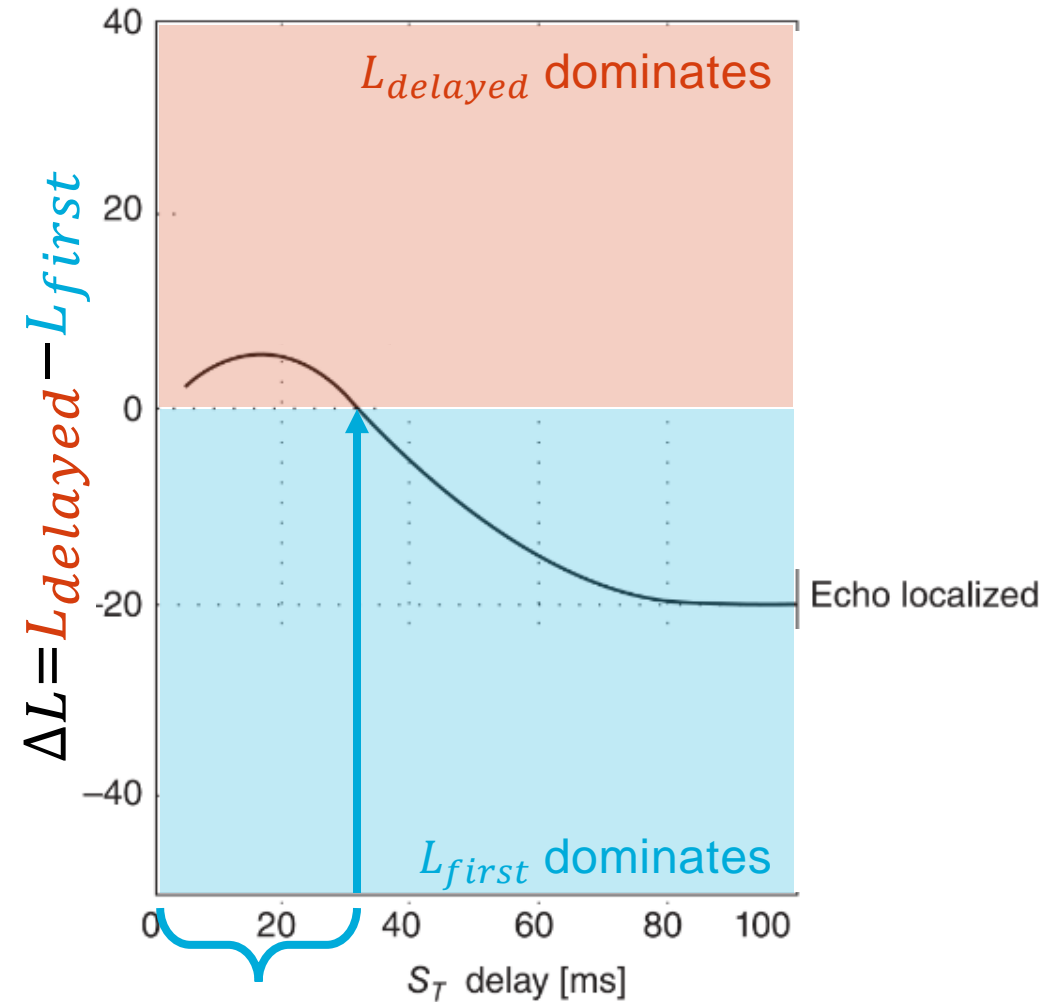
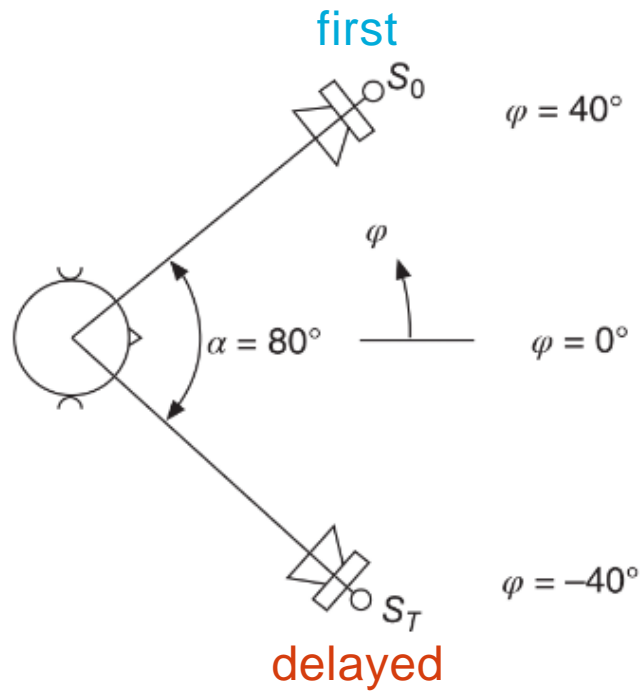


$$L_{\text{delayed}} = L_{\text{first}}$$



~12 m difference

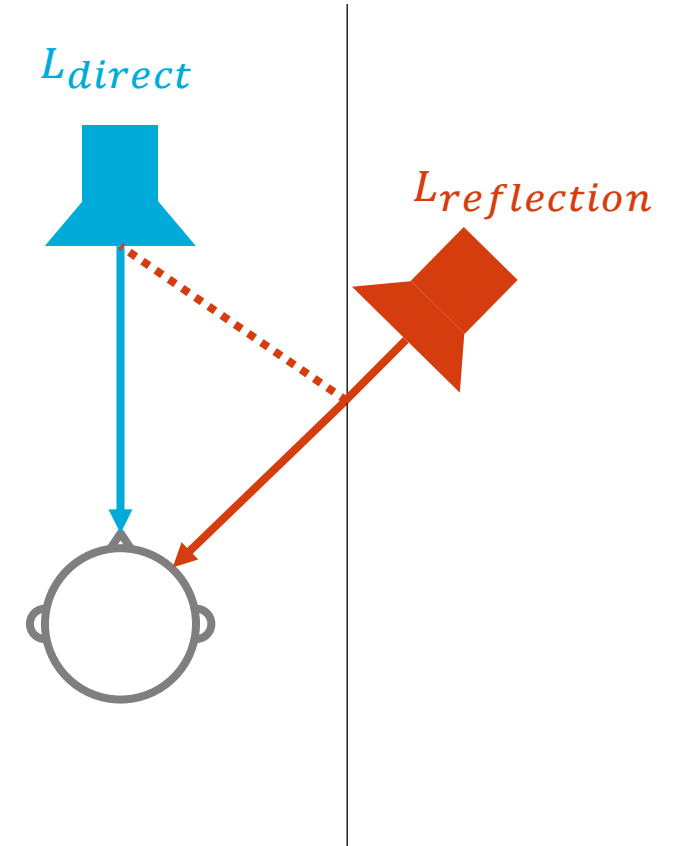
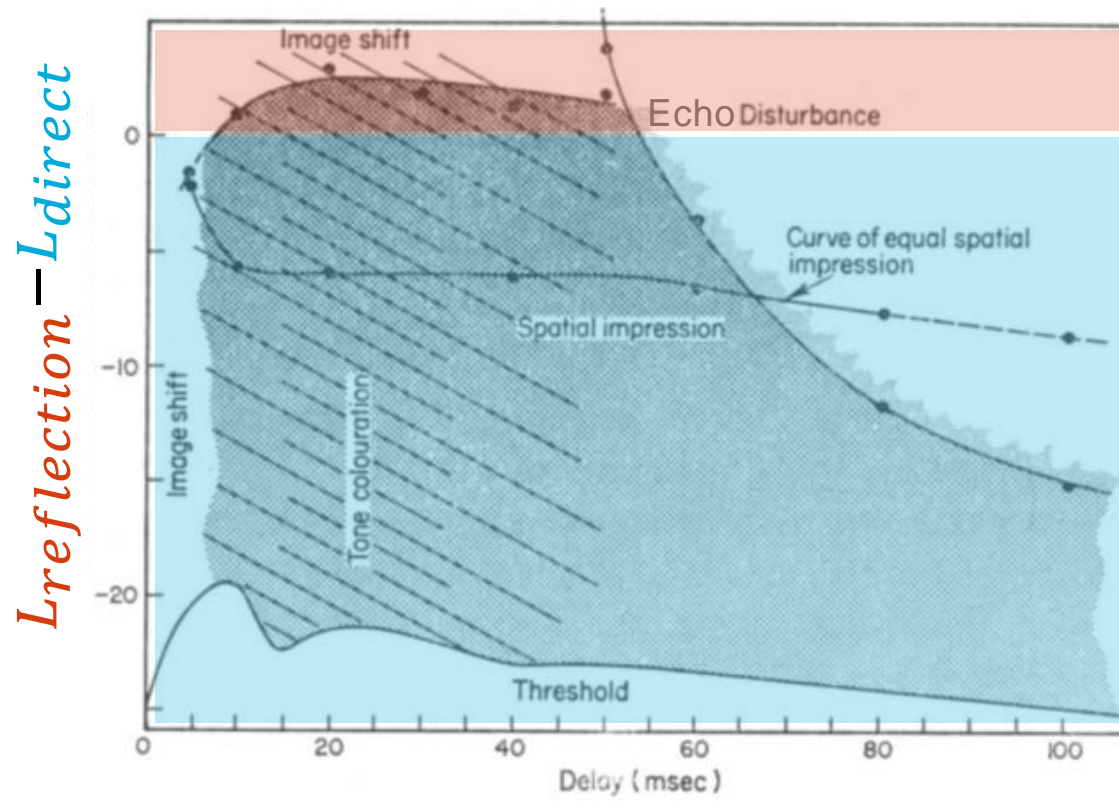
Perception first lateral reflections



~12 m difference

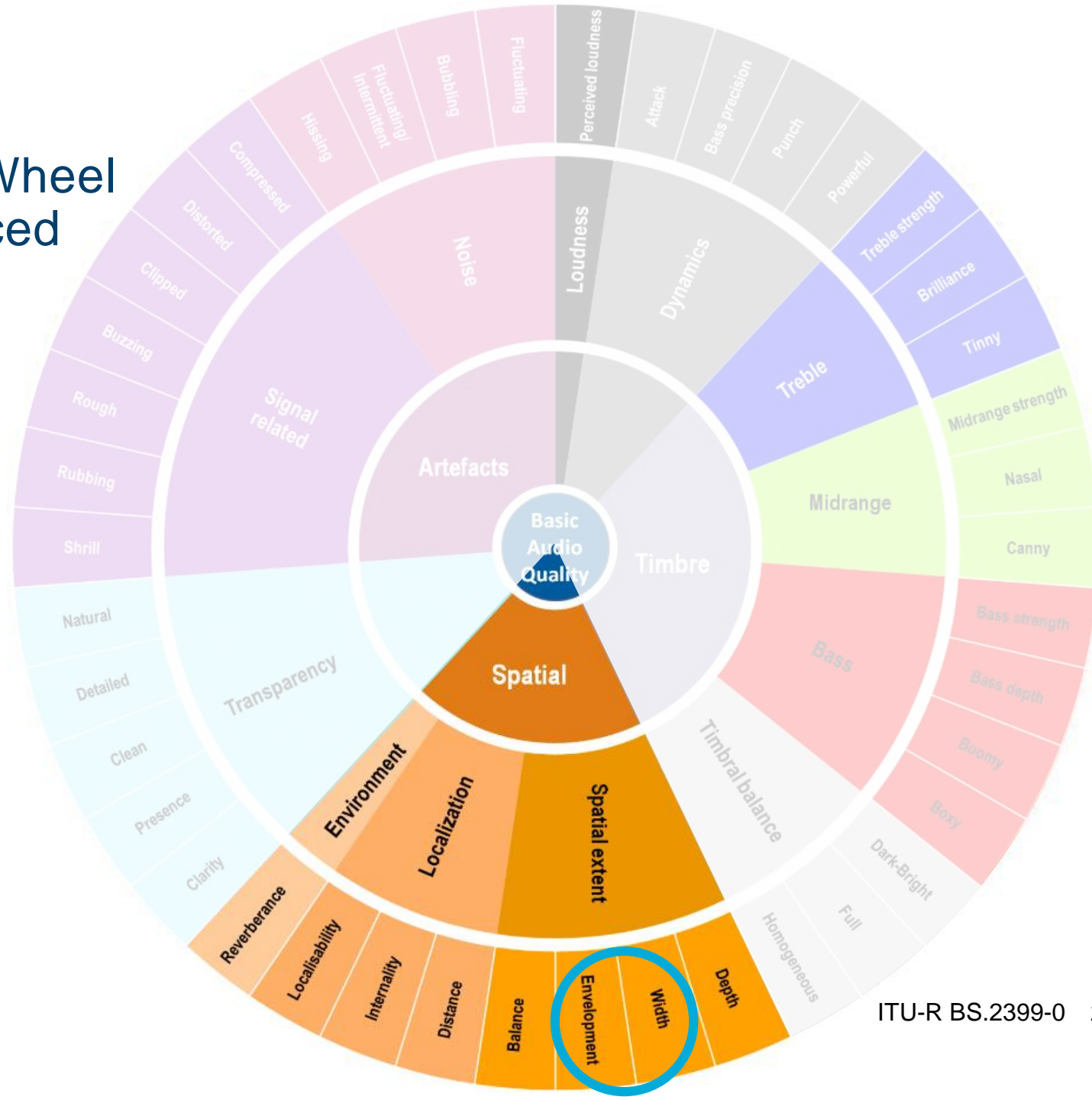
Pulkki, 2015
(adapted from Blauert)

Perception first lateral reflections



Barron 1971

The Audio Wheel for reproduced sound



ITU-R BS.2399-0 2017

Spatial Audio Quality Inventory (SAQI, Lindau 2014)

ACTA ACUSTICA UNITED WITH ACUSTICA
Vol. 100 (2014)

Lindau *et al.*: Spatial audio quality inventory

Table I. Spatial Audio Quality Inventory (SAQI) - English version.
*sound examples may be downloaded from <http://dx.doi.org/10.14279/depositonce-1>

Quality	Circumscription	Scale End Label
Difference	Existence of a noticeable difference.	none – very large
Timbre		
Tone color	Timbral impression which is determined by the ratio of high to low fre-	darker – brighter
Geometry		
Horizontal direction	Direction of a sound source in the horizontal plane.	shifted anticlockwise – shifted clockwise (up to 180°)
Vertical direction	Direction of a sound source in the vertical plane.	shifted down – shifted up (up to 180°)
Front-back position	Refers to the position of a sound source before or behind the listener only. Impression of a position difference of a sound source caused by 'reflecting' its position on the frontal plane going through the listener.	dichotomous scale: not confused/confused
Distance	Perceived distance of a sound source.	closer – more distant
Depth	Perceived extent of a sound source in radial direction.	less deep – deeper
Width	Perceived extent of a sound source in horizontal direction.	less wide – wider
Height	Perceived extent of a sound source in vertical direction.	less high – higher
Externalization	Describes the distinctness with which a sound source is perceived within closed between the phenomena of in-head localization and out-of-head localization. Examples: Poorly/not externalized = perceived position of sound sources at diotic sound presentation via headphones, good/strongly externalized = perceived position of a natural sound in reverberant environment and when allowing for movements of the listener.	less internalized – more externalized
Room		
Level of reverberation	Perception of a strong reverberant sound field, caused by a high ratio of reflected to direct sound energy. Leads to the impression of high diffusivity in case of stationary excitation (in the sense of a low D/R-ratio). Example: The perceived intensity of reverberation differs significantly between rather small and very large spaces, such as living rooms and churches.	less – more
Duration of reverberation	Duration of the reverberant decay. Well audible at the end of signals.	shorter – longer
Envelopment (by reverberation)	Sensation of being spatially surrounded by the reverberation. With more pronounced envelopment of reverberation, it is increasingly difficult to assign a position. Impressions of either low or high reverberation envelopment arise with either diotic or dichotic (i.e., uncorrelated) presentation of reverberant audio material.	less pronounced – more pronounced

Room Acoustic Quality Inventory (RAQI, Weinzierl 2018)

TABLE II. Three possible solutions of the CFA yielding four, six, and nine factors as sub-dimensions of room acoustical impression. The corresponding questionnaires would contain 14, 20, and 29 items, which are given with corresponding poles. Weights (W) and intercepts (I) should be used to measure factors and for structural equation analysis (see Sec. IV). Four additional items with high re-test reliability, which could not be assigned to one of the factors, are given below.

		Factors	Items	Poles	W	I	
9-factor RAQI	6-factor RAQI	4-factor RAQI	Quality	Liking	I like it – I don't like it	1.0	2
				Room acoustic suitability	suitable – not suitable	1.0	56
				Ease of listening	difficult – effortless	0.9	10
				Global balance	balanced – unbalanced	0.8	4
		Strength	Size	small – large	1.0	57	
			Loudness	soft – loud	0.7	64	
			Width	small – large	0.8	57	
			Reverberance	Duration of reverberation	short – long	1.0	47
		Reverberance		dry – reverberant	1.0	54	
		Strength of reverberation		weak – strong	1.0	51	
		Envelopment by reverberation		weak – strong	0.7	48	
		Brilliance	Brilliance	not brilliant – very brilliant	1.0	48	
	Tone Color bright/dark		bright – dark	-0.8	-7		
	Treble range characteristic		attenuated – emphasized	0.7	3		
	Irregular decay		Flutter Echo	none – very strong	1.0	26	
		Echo	none – very strong	0.7	40		
		Coloration	Irregularity in sound decay	none – very strong	0.9	32	
			Boominess	not boomy – very boomy	1.0	37	
	Roughness		not rough – very rough	0.7	31		
	Comb filter coloration		none – very strong	0.8	34		
	Clarity	Temporal clarity	clear – blurred	1.0	10		
		Spatial transparency	blurred – transparent	1.0	1		
		Precision of localization	precise – diffuse	-0.8	-6		
		Liveliness	Liveliness	dead – lively	1.0	11	
Intimacy	Spatial presence	low – high	1.0	63			
	Dynamic range	small – large	0.9	50			
	Intimacy	remote – intimate	1.0	-3			
	Distance	close – distant	-0.8	51			
Single items	Single items	Warmth	cool – warm	0.5	3		
		Metallic tone color	not metallic – very metallic				
		Openness	open – constricted				
		Attack	soft – crisp				
Single items	Single items	Richness of sound	low – high				

classifications compared to an *a priori* probability of 3.2% for most rooms (2.1% for the ten rooms not suitable for orchestra play). Hence, the factors are able to reliably discriminate between different rooms, even if the factor scores alone are, of course, not able to identify a specific room.

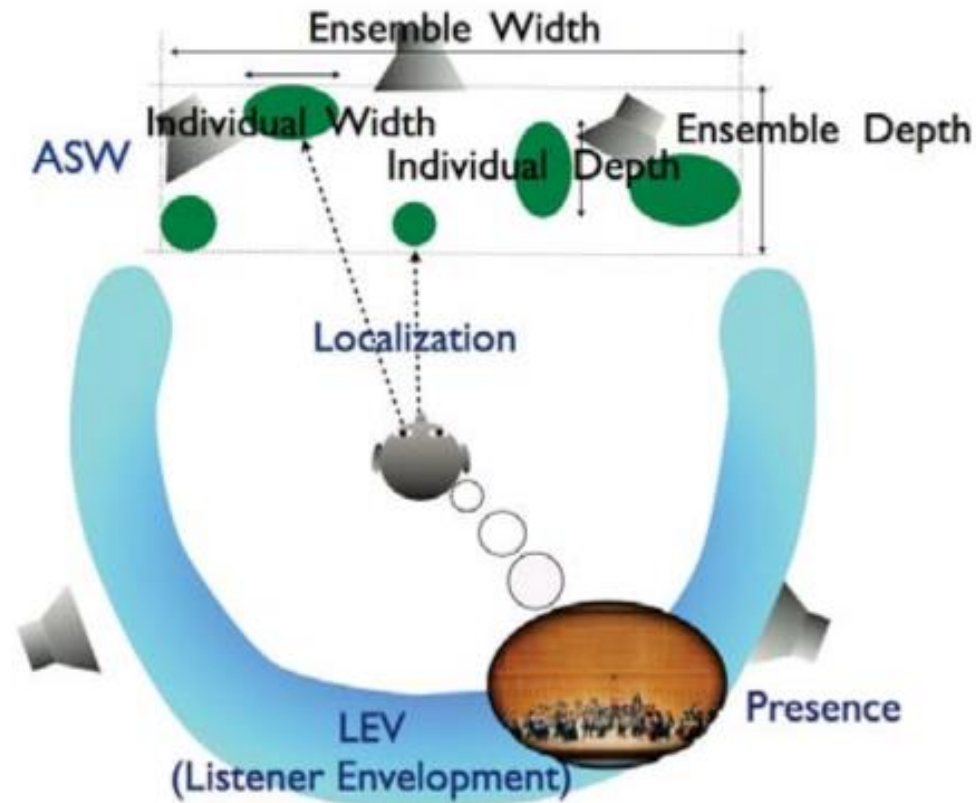
E. Primary research data

All primary research data of the current study are available as a digital publication (Ackermann *et al.*, 2018).

performance venues for music and speech. In a first step, a vocabulary of 50 attributes was generated by an expert focus group, including scholars as well as room acoustical consultants. Even if some of the attributes turned out to be unsuitable in the following stimulus-based evaluation, the item battery itself can be considered as an attempt to standardize the mostly inconsistent terminology to assess the qualities of performance venues for music and speech.

Important spatial attributes for audio/music in enclosed rooms

Apparent Source Width (ASW)



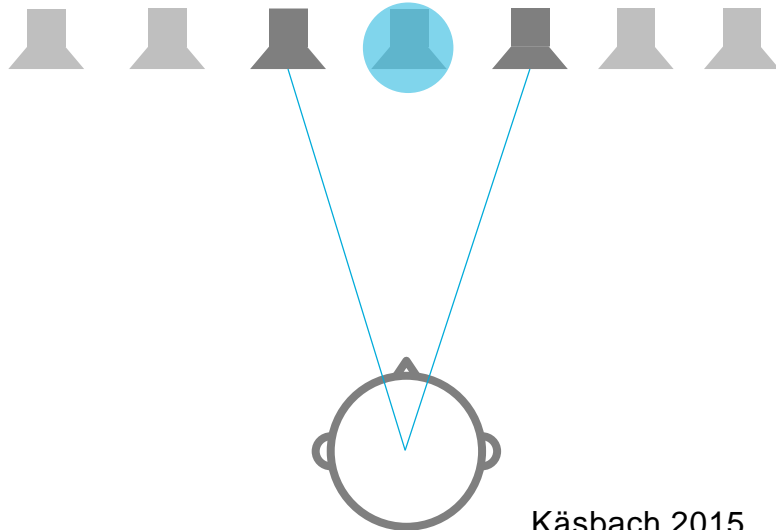
Listener Envelopment (LEV)

Kamekawa 2013

ASW in normal hearing (NH), hearing impaired (HI)

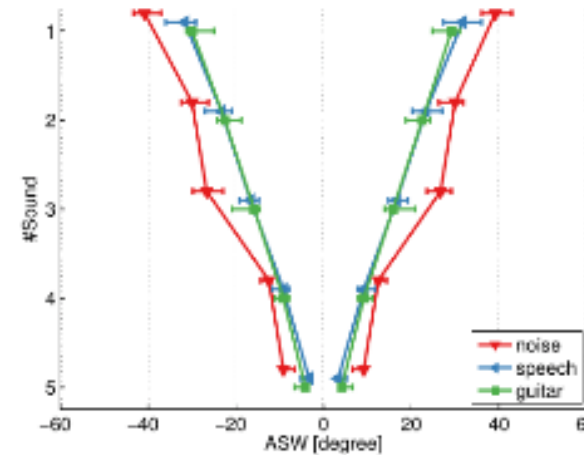
Sound	Parameters			IACC _{E3}		
	LS ang. [°]	dist. [m]	Φ [°]	noise	speech	guitar
#1	± 42.5	2,29	30	0.49	0.43	0.61
#2	± 30	1,95	40	0.47	0.53	0.54
#3	± 30	1,95	20	0.62	0.67	0.63
#4	± 16	1.74	20	0.71	0.73	0.80
#5	± 0	1.69	-	0.82	0.85	0.80

Phantom sound source

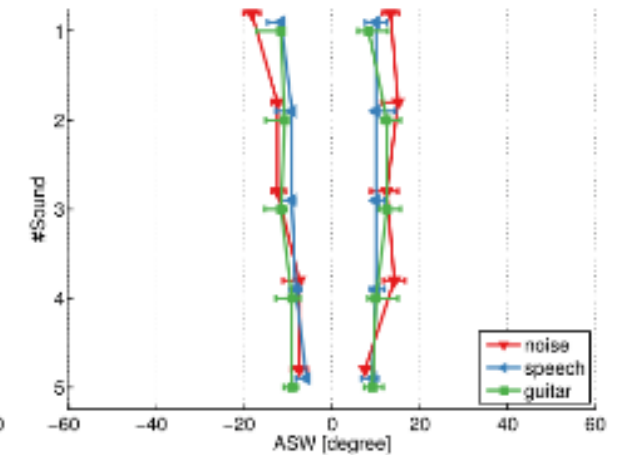


Käsbach 2015

NH



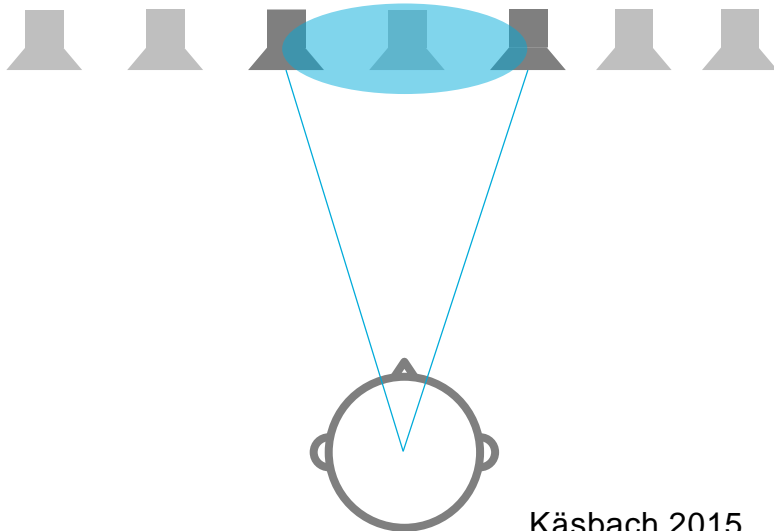
HI



ASW in normal hearing (NH), hearing impaired (HI) and aided listening

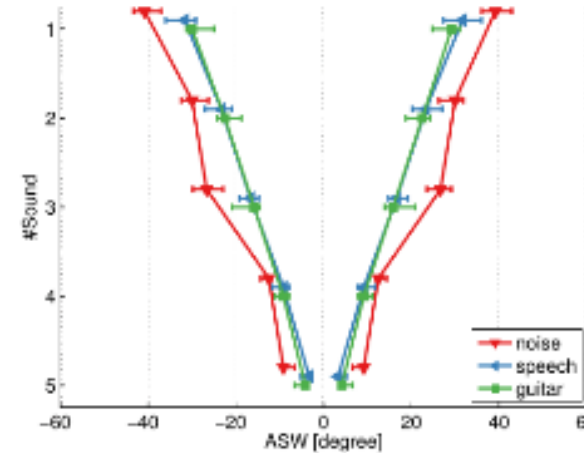
Sound	Parameters			IACC _{E3}		
	LS ang. [°]	dist. [m]	Φ [°]	noise	speech	guitar
#1	± 42.5	2,29	30	0.49	0.43	0.61
#2	± 30	1,95	40	0.47	0.53	0.54
#3	± 30	1,95	20	0.62	0.67	0.63
#4	± 16	1.74	20	0.71	0.73	0.80
#5	± 0	1.69	-	0.82	0.85	0.80

Phantom sound source

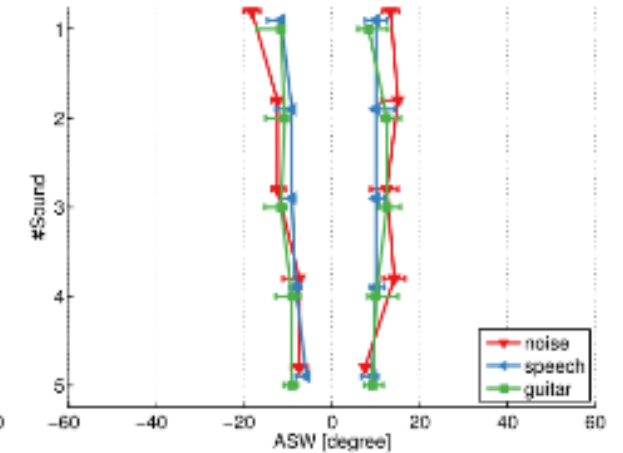


Käsbach 2015

NH

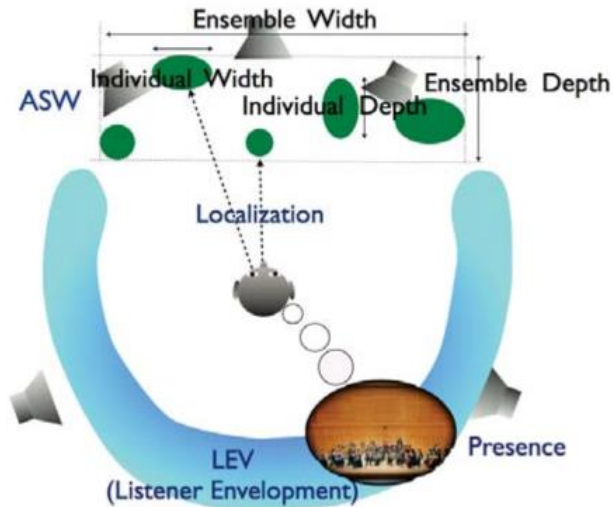


HI

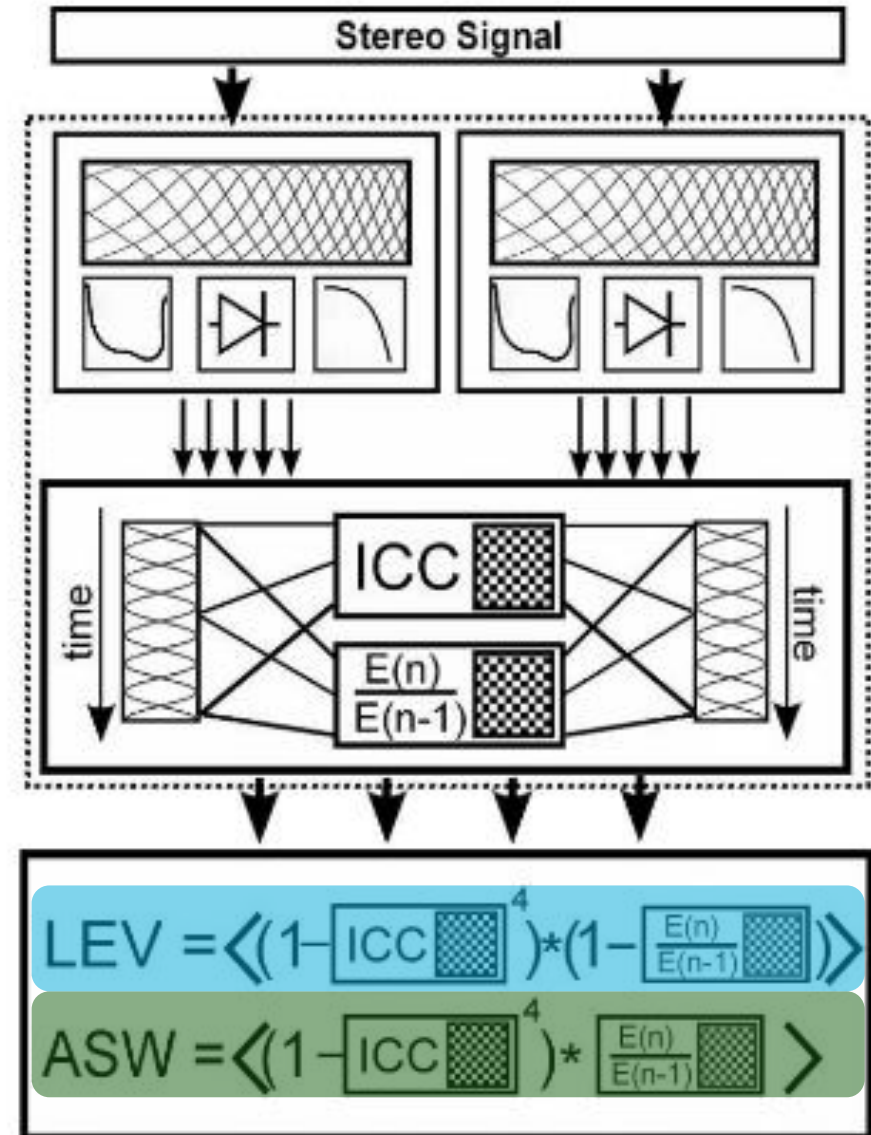


Modelling LEV and ASW

Apparent Source Width (ASW)



Listener Envelopment (LEV)



Energie based measures to characterize concert hall acoustics

- Subjective level of sound

Strength G

$$G = 10 \log_{10} \frac{\int_0^\infty p^2(t) dt}{\int_0^\infty p_A^2(t) dt}$$

p_A : level, 10 m away
from omnidirectional
source

- Perceived reverberance

$$EDC(t) = \int_t^\infty h^2(\tau) d\tau$$

Early decay time EDC(t)

- Perceived clarity of sound

– Clarity C_{80} in dB

Effects of e.g. IACC not reflected!

$$C_{80} = 10 \log_{10} \frac{\int_0^{80 \text{ ms}} p^2(t) dt}{\int_0^\infty p^2(t) dt}$$

- Apparent source width

– Early lateral energy fraction J_{LF}

$$J_{LF} = 10 \log_{10} \frac{\int_{5 \text{ ms}}^{80 \text{ ms}} p_8^2(t) dt}{\int_0^{80 \text{ ms}} p^2(t) dt}$$

p_8 : figure of eight mic

- Listener envelopement

– Late lateral sound level L_J

$$L_J = 10 \log_{10} \frac{\int_{80 \text{ ms}}^\infty p_8^2(t) dt}{\int_0^\infty p_A^2(t) dt}$$

General literature references

Blauert, J. (1996) Spatial hearing: The psychophysics of human sound localization (revised Edition). MIT Press, ISBN 978-0-262-26868-4
<https://doi.org/10.7551/mitpress/6391.001.0001>

Pulkki, V., Karjalainen, M. (2015) Communication acoustics : an introduction to speech, audio, and psychoacoustics. Wiley and Sons, ISBN 978-1-118-86654-2

Applied Psychophysics I – Lecture schedule

Week	Date	Topic
1	16.10.24	Introduction
2	23.10.24	Auditory physiology and selected auditory models
3	30.10.24	Suprathreshold scaling methods
4	06.11.24	Loudness perception, loudness models and audio metering
5	13.11.24	Selected sensations, sound attributes and measurement of timbre
6	20.11.24	Chocolate results Presentation topics
7	27.11.24	Spatial hearing sensations
8	04.12.24	Audio quality assessment methods and model approaches
9	11.12.24	Calibration of listening setups
10	18.12.24	Ethics and data protection
Winter break		
11	08.01.25	Student presentations
12	15.01.25	Student presentations
13	22.01.25	Student presentations
14	29.01.25	Student presentations

Your comments and take-home messages

Papers for presentation

	Date	Topic	Presenter	Date
1	Porysek et al. 2024	Identifying Principal Attributes for Evaluating Audio Quality of Reproduction Systems With Spatially Dynamic Program Material	Pascal	
2	Marozeau 2007	The effect of fundamental frequency on the brightness dimension of timbre	(Patrick)	
3	Jacobsen & Siedenburg 2024	Exploring the relation between fundamental frequency and spectral envelope in the perception of musical instrument sounds	Patrick	
4	Culling et al. 2024	Effect of ambisonic order on spatial release from masking	Annika	
5	Hots et al. 2024	Spatial weights in loudness judgements	Konosuke	
6	Aker et al. 2024	Some, but not all, cochlear implant users prefer music stimuli with congruent haptic stimulation	Florian	
7	Bürgel et al. 2024	Enhanced salience of edge frequencies in auditory pattern recognition		
8	Joseph Benjamin & Siedenburg 2023	Exploring level- and spectrum-based music mixing transforms for hearing-impaired listeners	Porsche Phufah?	
9	Szwarcberg & Lavandier 2024	Third-octave analyses describing two perceptual dimensions of sound reproduction and the resulting overall perceived dissimilarity between loudspeakers or headphones		
10	Madsen & Oxenham 2024	Mistuning perception in music is asymmetric and relies on both beats and inharmonicity		

Papers for presentation

	Date	Topic	Presenter	Date
11	Francombe 2017a,b	Evaluation of Spatial Audio Reproduction Methods Part 1: Elicitation of Perceptual Differences / Part 2: Analysis of Listener Preference	Porsche	
12	Wilson 2016	Perception of Audio Quality in Productions of Popular Music		
13	Joseph Benjamin & Siedenburg 2024	Evaluating audio quality ratings and scene analysis performance of hearing-impaired listeners for multi-track music	Phufah	
14	Aker 2024b	Perceived auditory dynamic range is enhanced with wristbased tactile stimulation		
15				
16				
17				
18				
19				
20				

Applied Psychophysics I – Lecture schedule

Week	Date	Topic
1	16.10.24	Introduction
2	23.10.24	Auditory physiology and selected auditory models
3	30.10.24	Suprathreshold scaling methods
4	06.11.24	Loudness perception, loudness models and audio metering
5	13.11.24	Selected sensations, sound attributes and measurement of timbre
6	20.11.24	Chocolate ratings & Presentation topics
7	27.11.24	Spatial hearing sensations
8	04.12.24	Audio quality assessment methods and model approaches
9	11.12.24	Calibration of listening setups
10	18.12.24	Ethics and data protection
Winter break		
11	08.01.25	Student presentations Annika
12	15.01.25	Student presentations Patrick
	22.01.25	Student presentations
14	29.01.25	Student presentations

Thank you for your attention!

stephan.toepken@uol.de