Aalto University
School of Electrical Engineering

Communication acoustics
Ch 12: Spatial hearing

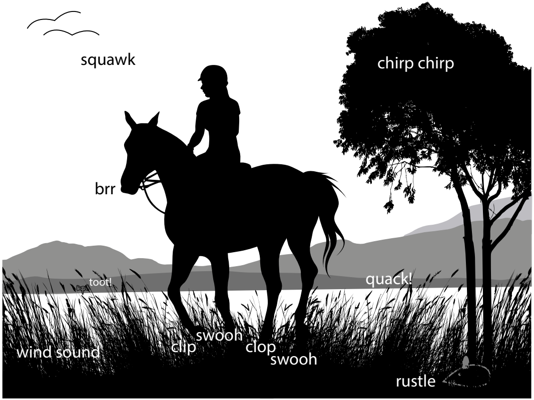
Ville Pulkki

Department of Signal Processing and Acoustics
Aalto University, Finland

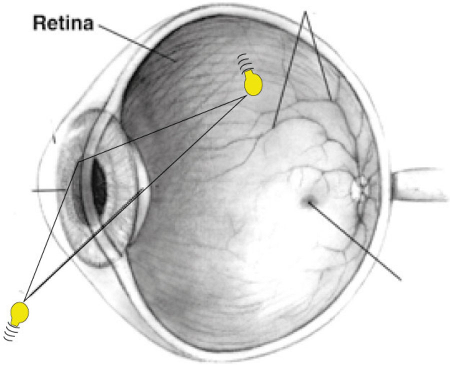
September 20th, 2022

Where and what?

- Localization of sources
- Beam-forming towards different directions

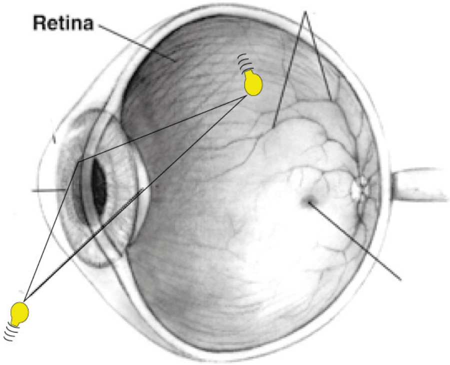


Human eye



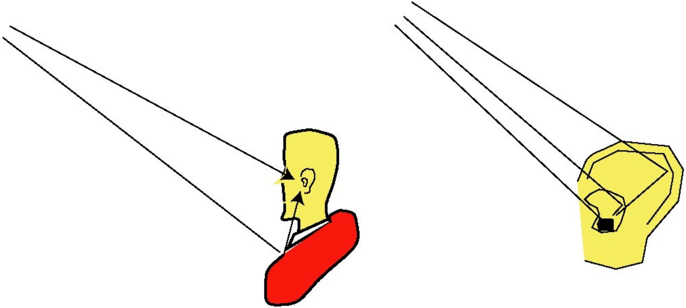
- The cells in the eye are a priori sensitive to direction of light

Human eye



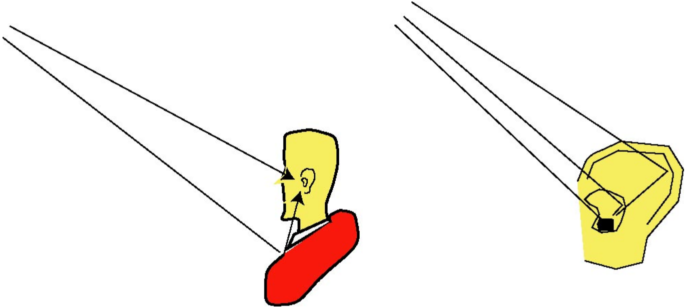
- The cells in the eye are a priori sensitive to direction of light
- Response to quite limited range of wavelengths (380-740nm)

Human spatial hearing



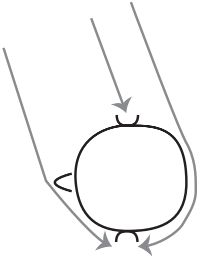
- Response to very large range of wavelengths (2cm–30m)
- Ear canal diameter $<$1cm, sound just bends into the canal

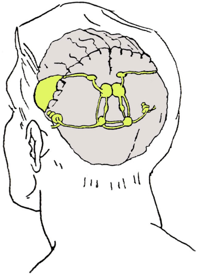
Human spatial hearing



- Response to very large range of wavelengths (2cm–30m)
- Ear canal diameter $<$1cm, sound just bends into the canal
- One ear alone knows quite little of direction

Human spatial hearing

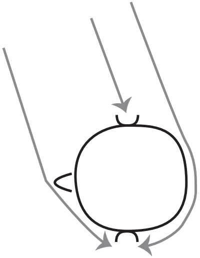


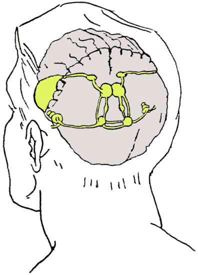


©J. Blauert

- The perception of direction is formed in signal analysis by the brains

Human spatial hearing

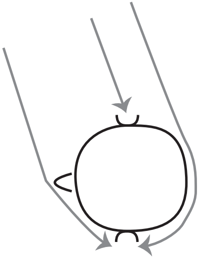


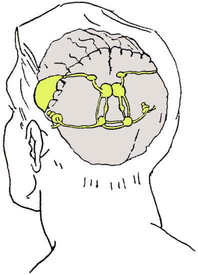


©J. Blauert

- The perception of direction is formed in signal analysis by the brains
- Signal characteristics in one ear / Signal differences between two ears

Human spatial hearing

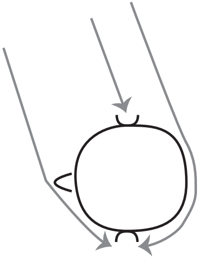


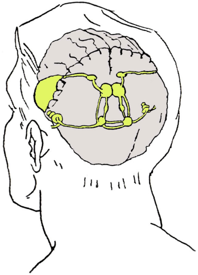


©J. Blauert

- The perception of direction is formed in signal analysis by the brains
- Signal characteristics in one ear / Signal differences between two ears
- Hearing mechanisms estimate the most probable direction for sound

Human spatial hearing





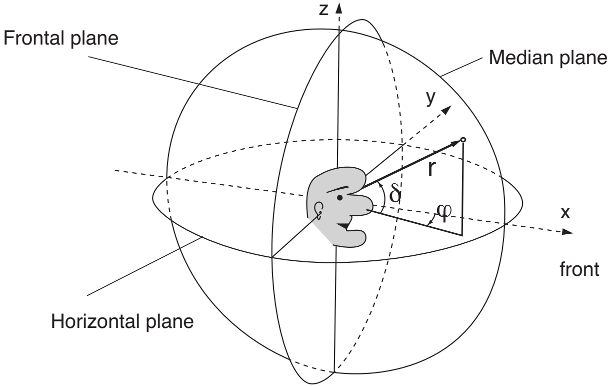
- The perception of direction is formed in signal analysis by the brains
- Signal characteristics in one ear / Signal differences between two ears
- Hearing mechanisms estimate the most probable direction for sound
- Hearing can be fooled easily by audio techniques!

This chapter

- Basic concepts
- Head-related acoustics
- Localization cues
- Localization accuracy
- Perception of direction in presence of echoes
- Ability to listen selectively specific direction
- Distance perception

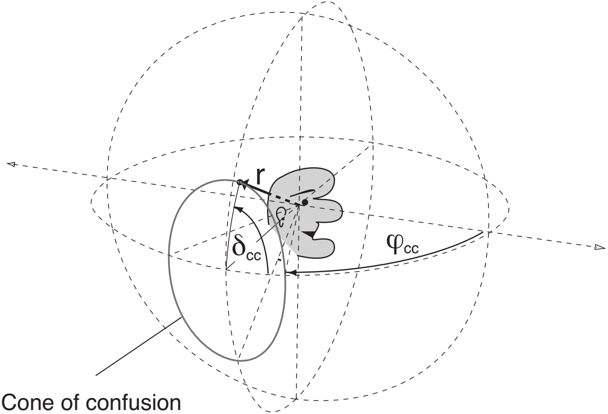
Coordinate system

- Azimuth-elevation / median plane



Coordinate system 2

- Cone of confusion

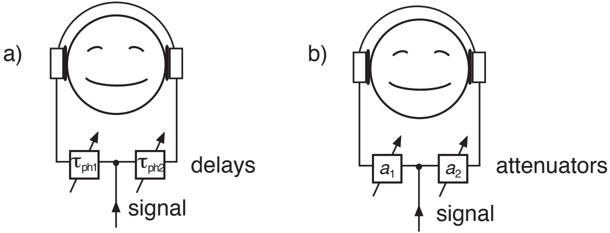


Cone of confusion

Basic concepts

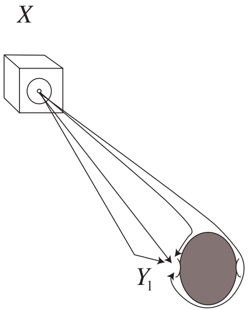
- Listening conditions
  - Monaural hearing (hearing processes and cues that need signal from only one ear)
  - BInaural hearing (differences in ear signals have also an effect)
- Spatial sound reproduction methods
  - Monotic (signal fed to only one ear)
  - Diotic (same signal fed to both ears)
  - Dichotic (different signal fed to ears)

Dichotic listening with headphones



Head-related acoustics

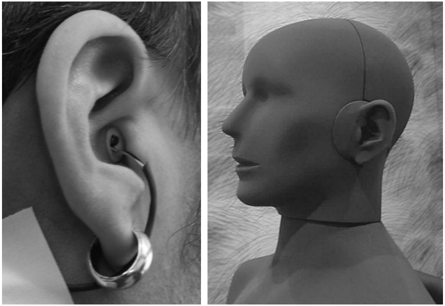
- Let us consider only free field first
- How does incoming sound change as it hits the listener?
- What is the transfer function from source to ear canal?



$Y_{left} = H_{left} X$ $y_{left} = h_{left}(t) * x(t)$

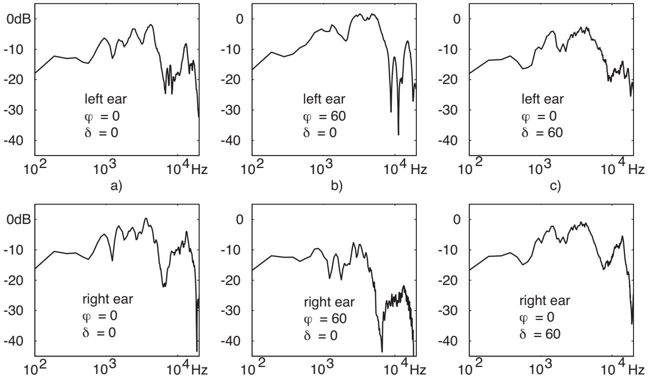
Head-related acoustics measurements

- Transfer function from source to ear canal
- Place a microphone to ear canal or use dummy heads
- Head-related transfer function, head-related impulse response
- Depends heavily on the direction of the source



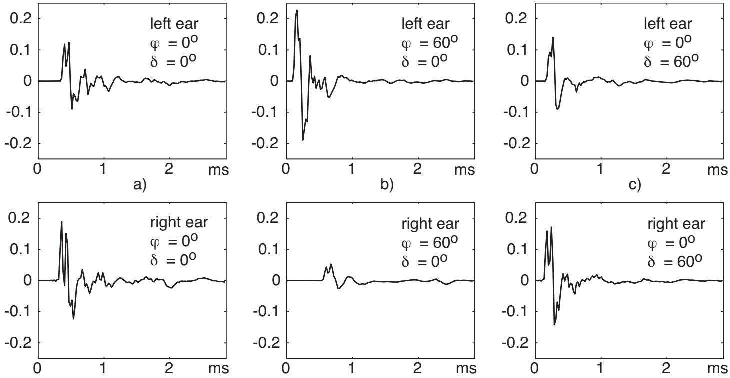
Head related transfer function (HRTF)

- Measured with real subject for three directions
- Magnitude response

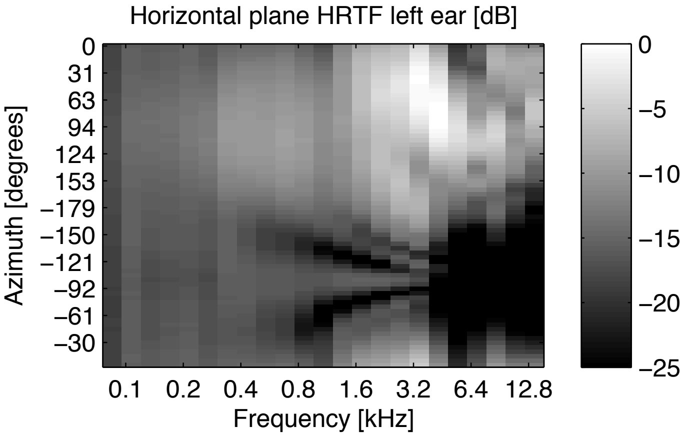


Head related impulse response (HRIR)

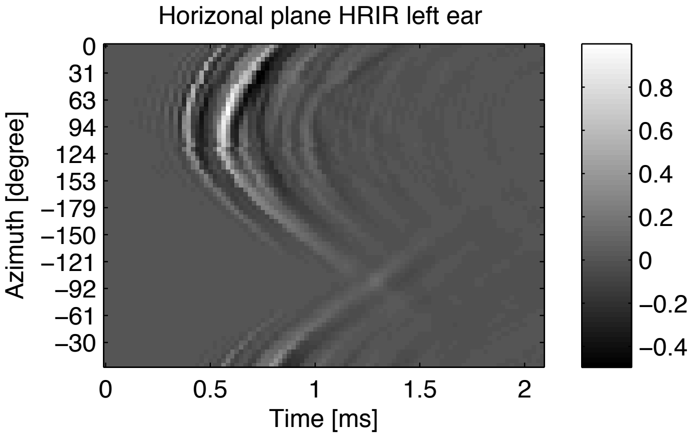
- HRIR == HRTF in time domain, quite often HRIRs are also called HRTFs



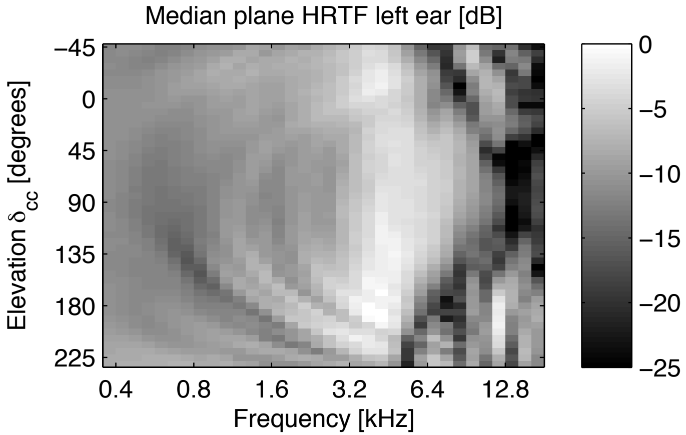
HRTF in horizontal plane



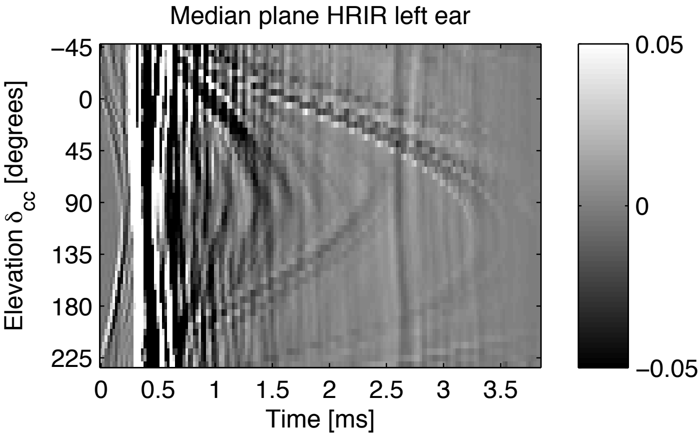
HRIR in horizontal plane



HRTF in median plane



HRIR in median plane



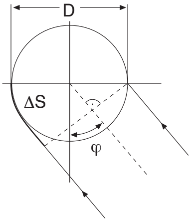
Localization cues

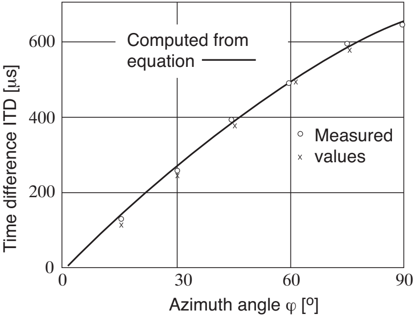
- HRTFs impose some spatial information to ear canal signals
- What information do we dig out from those?
- Binaural localization cues
  - Interaural time difference, ITD
  - Interaural level difference, ILD
- Monaural localization cues: spectral cues
- Dynamic cues

Interaural time difference (ITD)

- Sound arrives a bit earlier to one ear than the other
- ITD varies between -0.7ms and +0.7ms
- JND of ITD is of order of 20μs

$$\tau = \frac{D}{2c}(\varphi + \sin\ \varphi)$$

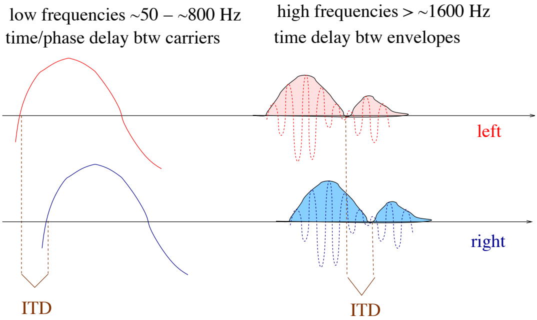




Interaural time difference (ITD)

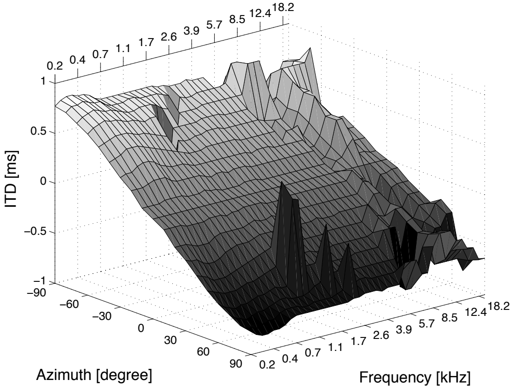
We are sensitive to

- phase differences at low frequencies
- time differences between envelopes at higher frequencies



Interaural time difference

- HRTFs measured from a human, ITD of noise sound source in free field computed with an auditory model

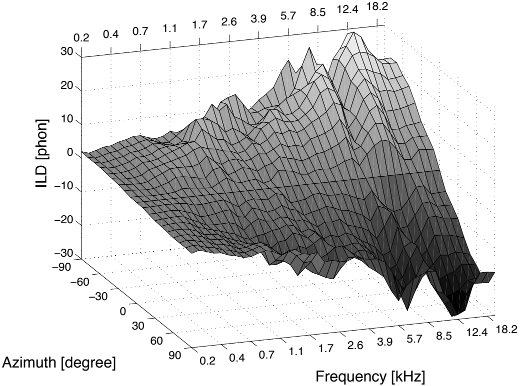


Interaural level difference (ILD)

- Head shadows incoming sound
- Effect depends on
  - wavelength, no shadowing at low frequencies
  - distance, larger ILD at very short distances (< 1m)
- ILD varies between -20dB and 20dB
- JND is of order of 1dB for sources near median plane

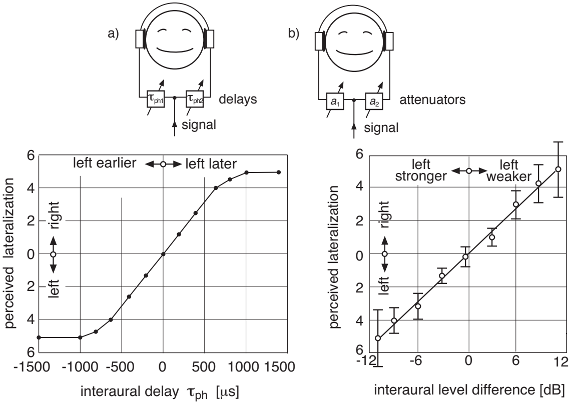
Interaural level difference

- HRTFs measured from a human, ILD of sound source in free field computed with an auditory model



Basic lateralization results

- Subjects report the position of internalized auditory source on interaural axis

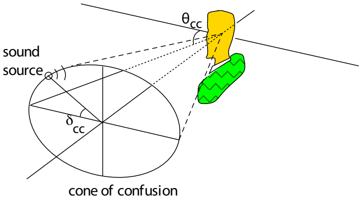


(Toole and Sayers, 1965)

(Sayers, 1964)

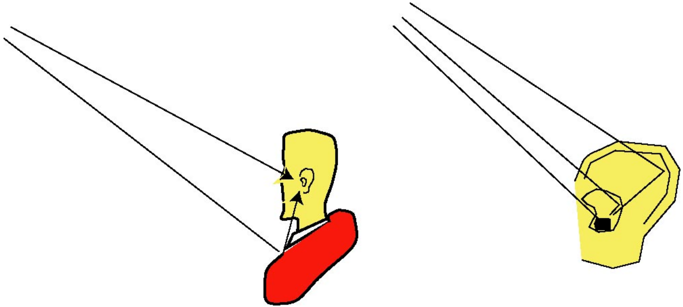
Addtional cues

- Binaural cues (ITD and ILD) in principle resolve only the cone of confusion where the source lies
- Other cues are used to resolve the direction inside the cone
  - monaural spectral cues
  - dynamic cues (effect of head movements to cues)

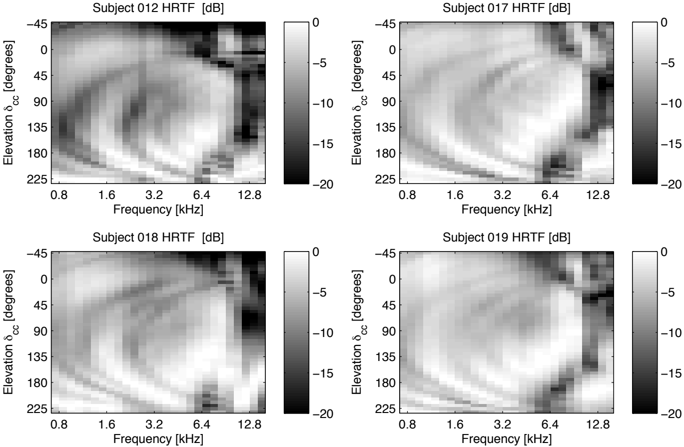


Monaural spectral cues

- Effect of pinna diffraction
- Effect of reflections/diffraction from torso



Monaural spectral cues

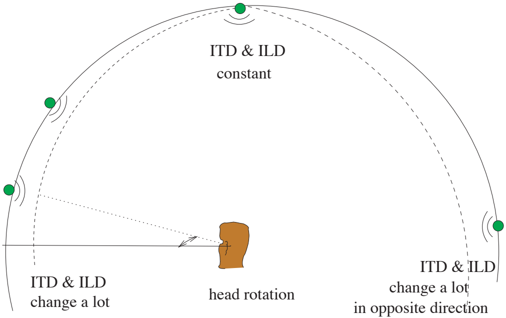


Monaural spectral cues

- Source has to have relatively flat spectrum
- Relatively reliable cue for most natural sounds
- Narrow-band sounds hard to localize (mosquito in dimmed room)
- Humans adapt relatively fast to new HRTFs, especially of visual cues are available
- The ears also grow throughout the life, constant adaptation

Dynamic cues

- When head rotates, ITD and ILD and spectral cues change
- Powerful though coarse cue
- Resolves efficiently front/back/inside-the-head locations, but not fine details



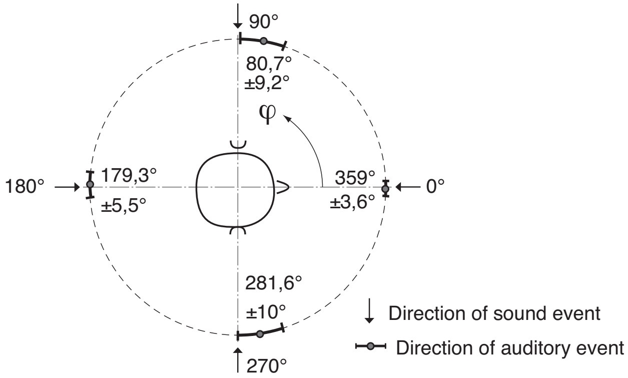
Interaction between spatial hearing and vision

- Both senses try to find out "where" is the source
- If visual cue is clear, vision dominates
- Ventriloquism, within about 30° area visual image captures auditory image
- If visual cue is blurry, or not prominent, auditory cue wins
- In some cases two events occur, visual event and auditory event

Accuracy of localization

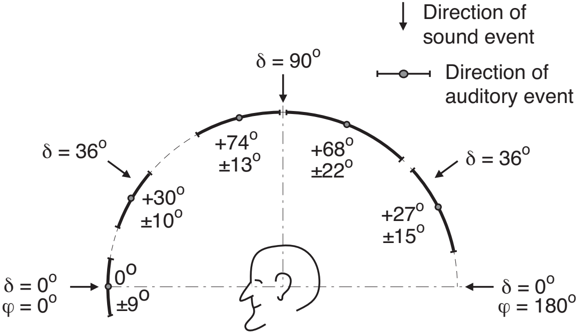
- How well do we localize point-like sources?
- What is the accuracy of directional hearing?
- Azimuth / elevation / 3D
- Accuracy of perception of spatial distribution of wide sound sources

Localization accuracy in horizontal plane



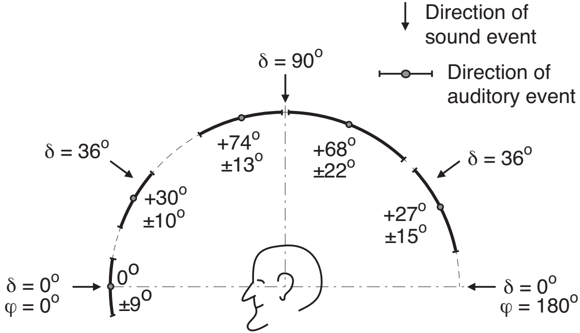
Adapted from Blauert (1996)

Localization accuracy in median plane



Adapted from Damaske and Wagener (1969)

Localization accuracy in median plane

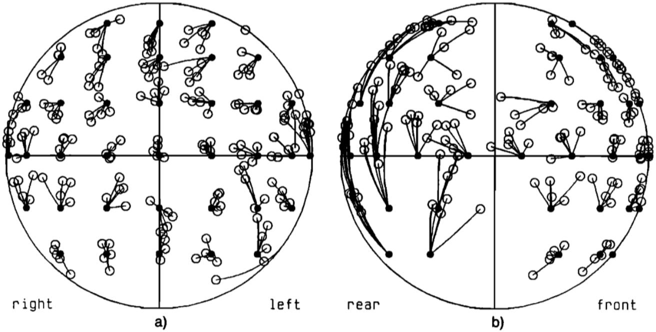


Adapted from Damaske and Wagener (1969)

- Monaural cues are effective only after adaptation with visual reference, thus field-of-vision has better accuracy

Accuracy of 3D localization

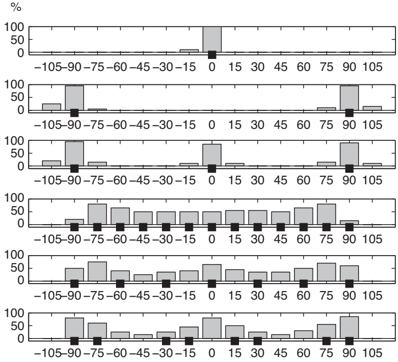
- Task: "Point with nose to the direction of auditory event"



Adapted from Middlebrooks (1992)

Accuracy of perception of spatial distribution of sources

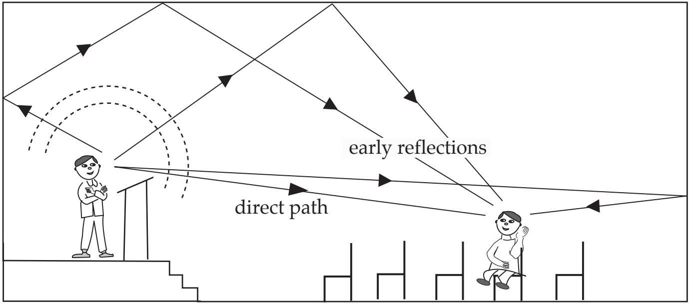
- 13 loudspeakers in free field, black squares denote loudspeakers that produced pink noise
- Task: "Tell which loudspeakers are on"
- Incorrect perception of distribution in complex cases
- Perception of spatial distribution of 1, 2, or 3 loudspeakers correct



Adapted from Santala and Pulkki (2011)

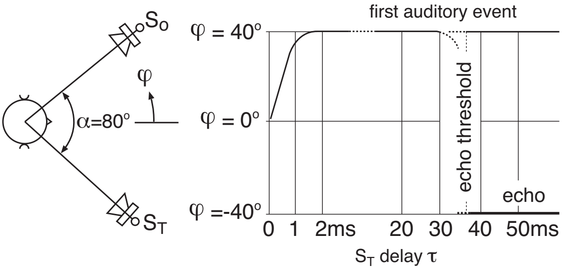
Directional hearing in enclosed spaces

- Previous results in free field
- What about rooms?
- Humans are relatively ok in localization in rooms. How come?



Precedence effect

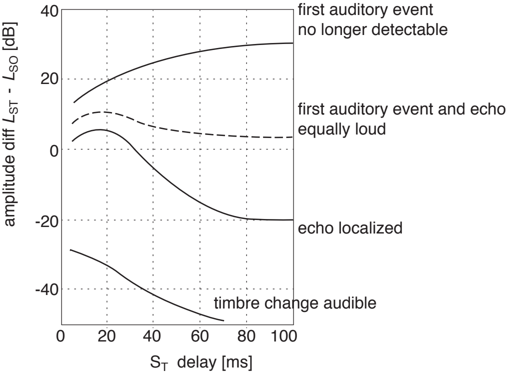
- Free-field exists only for short time 1–10ms after direct sound has arrived, and reflections not yet
- The short time dominates in direction perception
- Reflections (2 - 30ms) are perceived to arrive from the direction of direct sound



Adapted from Blauert (1996)

Precedence effect, experiments with two sounds

- Direct sound S_0: "lead"
- Reflected sound S_T: "lag"



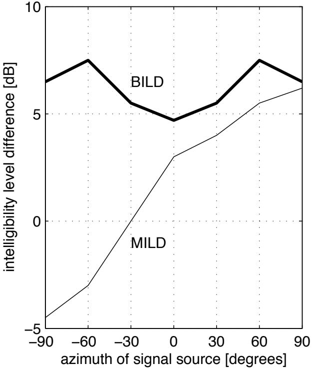
Adapted from Blauert (1996)

Binaural Advantages in Timbre Perception

- Try to listen selectively certain direction
- Binaural detection
- Binaural decolouration

Binaural intelligibility level difference

- Minimum understandable level of speech source in different directions in diffuse noise
- Reference condition: binaural listening with source in back
- Binaural intelligibility level difference (BILD)
- Monaural intelligibility level difference (MILD)
- MILD: best heard with source in same side
- BILD gives few decibels advantage to ipsilateral MILD
- "Better-ear" listening explains most of the effect
- Binaural hearing assists by few dB



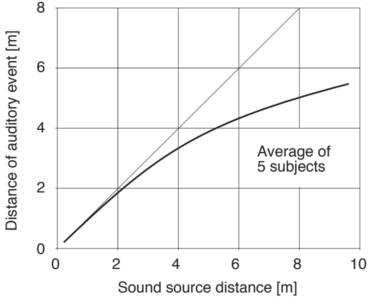
Adapted from Blauert (1996)

Perception of the distance of source

- Cues
  - Loudness, what higher loudness, that closer
  - Effect of room, what less that closer
  - Spectral content, what sharper sound, that closer
  - Binaural cues, excess ILD cues at low frequencies near the head
- Perception depends on signal, listening conditions and source directivity
- Absolute accuracy low
- Relative accuracy good (Is the source approaching?)

Distance cue 1: Loudness

- Amplitude of direct sound decreases with $1/r$ law
- The sound energy carried by the surface of a sphere ($A = 4\pi r^2$) is constant, and pressure is related to square root of energy
- SPL decreases 6dB with every doubling of distance
- If we know the source (human talker, insect, walking sounds), the distance is perceived relatively accurately in short
- In free field with real human speaker, similar effect is found:



Adapted from Békésy (1949)

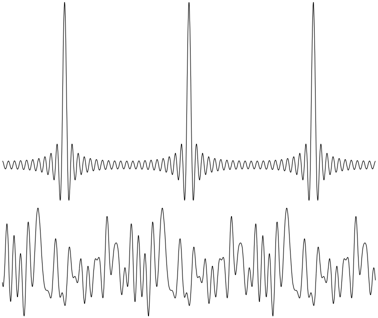
Distance cue 2: Effect of room

- Direct-to-reverberant ratio (DRR) is often mentioned as "perceptual cue", though it is rather a physical measure
- If the listener has heard a source in a specific room, the change in DRR leads into perception of approaching or distancing source
- Effect depends on signal.
- Easily perceivable on harmonic tone complexes

Distance cue 2: Effect of room

- The characteristics in ear canal signals change with DRR = one of distance cues
- The cue is not known well / which modification in the signal produces the perception
- One possibility is the sensitivity of ear to phase

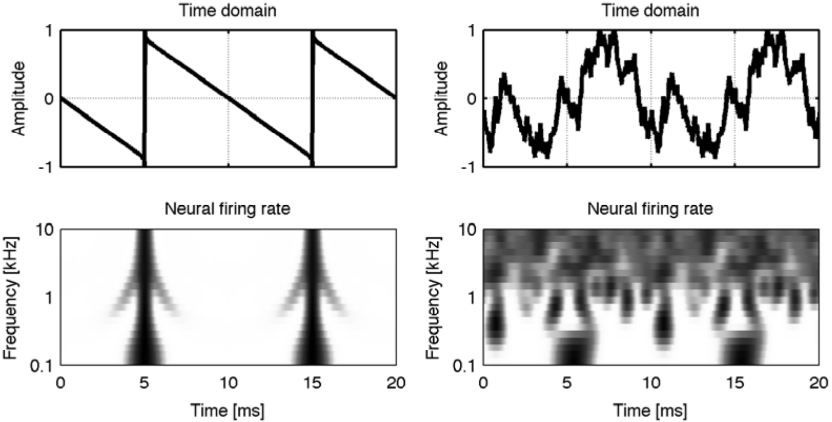
Zero-phase cosine signal



Random-phase (scrambled by the room effect) version

Response of cochlea to 100Hz sawtooth with phase modifications

- Original sound after room response



Adapted from Laitinen et al. 2013

Distance cue 3: Effect of air absorption

- shear viscosity, thermal conductivity or heat dissipation, and molecular relaxation due to oxygen, nitrogen, and water vapor vibrational, rotational, and translational energy
- depends on temperature, humidity, static pressure
- low-pass effect
- can for instance have values [1-100]dB/km
- known sounds are perceived closer if they have more energy at high frequencies
- hissing is perceived closer than normal speech

Distance cue 4: short-range binaural cues

- with plane waves, the ILD at low frequencies is negligible
- e.g., when source is 3cm from one ear, it is 30cm from the other ear
- $1/r$ law makes ILD to have high magnitude
- excess-ILD re ITD brings perceived source closer

Accuracy of distance perception

- In many studies
  - source is perceived too far in short distances
  - source is perceived too close in high distances
- Acoustic horizon, maximum perceived distance of auditory event

References

These slides follow corresponding chapter in: Pulkki, V. and Karjalainen, M. Communication Acoustics: An Introduction to Speech, Audio and Psychoacoustics. John Wiley & Sons, 2015, *where also a more complete list of references can be found.*

References used in figures:

Blauert, J. (1997) Spatial Hearing – Psychophysics of Human Sound Localization. MIT Press.

Damaske, P. and Wagener, B. (1969) Directional hearing tests by the aid of a dummy head. Acta Acustica United with Acustica, 21(1), 30–35.

Middlebrooks, J.C. (1992) Narrow-band sound localization related to external ear acoustics. J. Acoust. Soc. Am., 92, 2607–2624.

Santala, O. and Pulkki, V. (2011) Directional perception of distributed sound sources. J. Acoust. Soc. Am., 129, 1522.

Sayers, B.M. (1964) Acoustic-image lateralization judgments with binaural tones. J. Acoust. Soc. Am., 36(5), 923–926.

Toole, F. and Sayers, B.M. (1965) Lateralization judgments and the nature of binaural acoustic images. J. Acoust. Soc. Am., 37(2), 319–324.

von Békésy, G. (1949) The moon illusion and similar auditory phenomena. Ame. J. Psychol., 62(4), 540–552.