Computational Perception

15-485/785

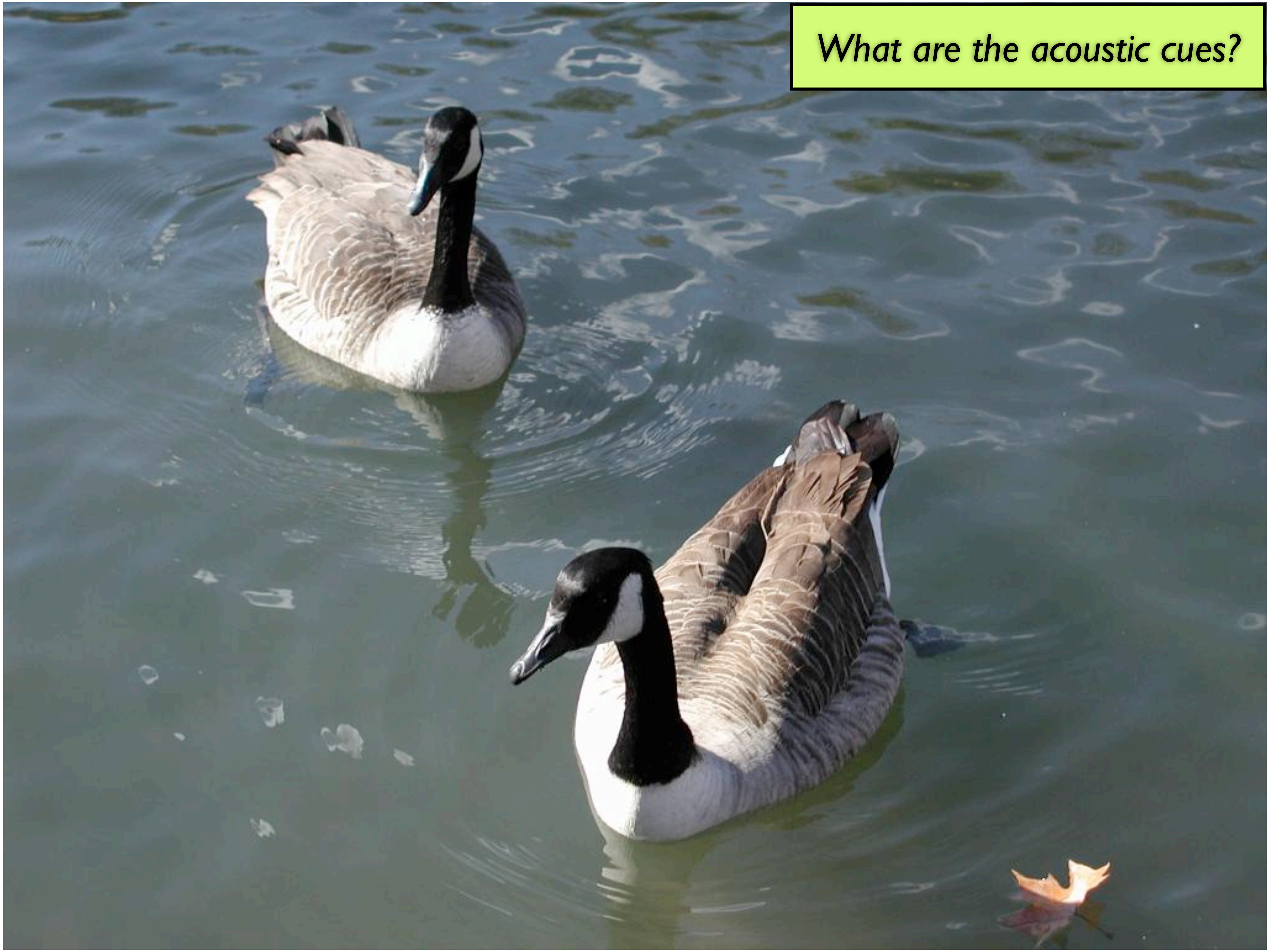
January 17, 2008

Sound Localization 1

Orienting

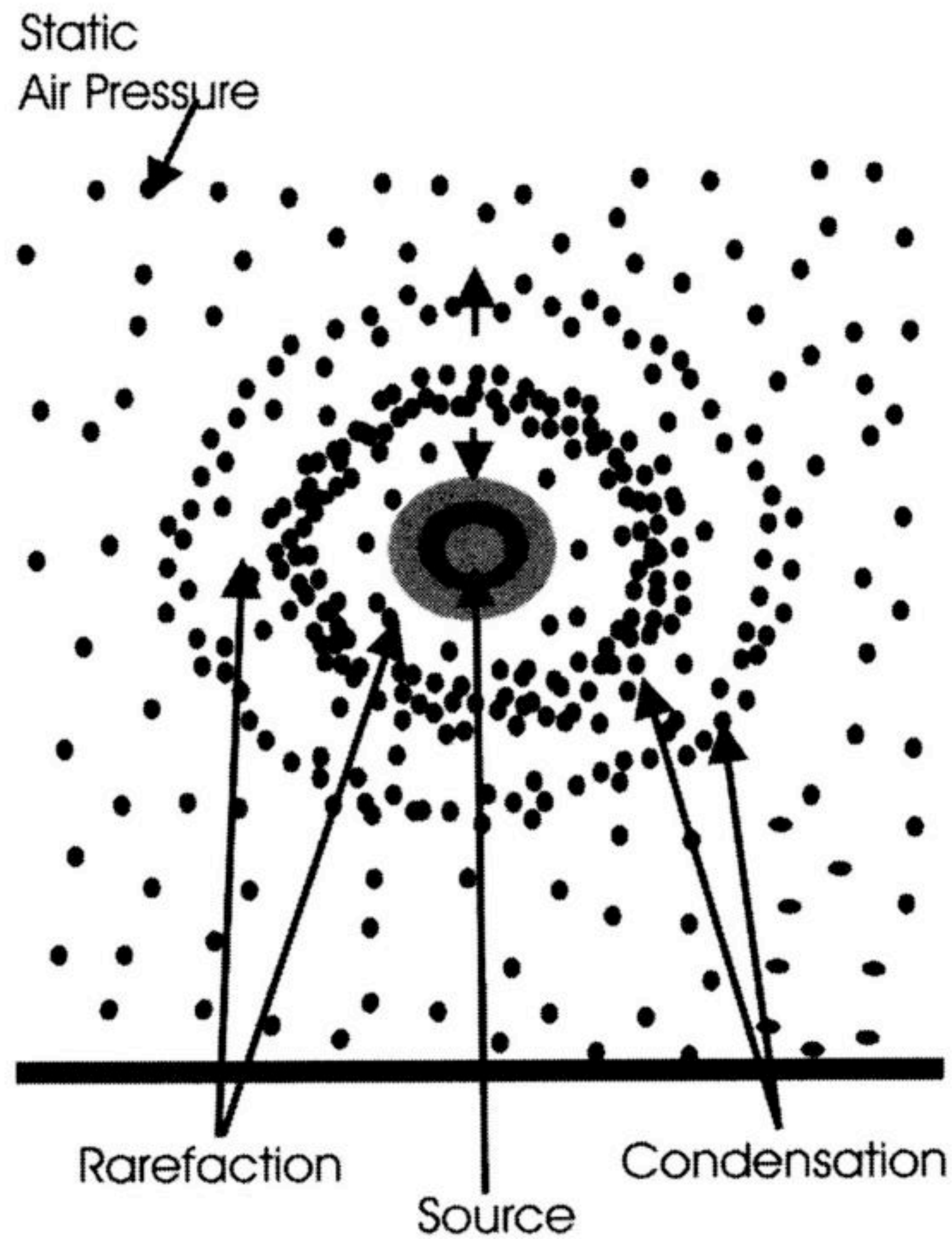
- sound localization
- visual pop-out
- eye/body movements
- attentional shift

The Problem of Sound Localization



What are the acoustic cues?

Sound propagation



from Yost, 2000

Units of sound intensity: deciBels (dB)

- dB is always *relative*.
- A standard in acoustics is:

$$
\begin{aligned}
\text{dB SPL} &\equiv 10 \log(I/I_0) \\
I_0 &= 20\mu \text{ Pascals} \\
&\approx 10^{-12} W/m^2
\end{aligned}
$$

- dB SPL for common sounds

sound level	typical example
140	close range gunshot
100	close shouting
70	normal conversation
30	soft whispering
6.5	human threshold (at 1 kHz)
-10	threshold for some animals

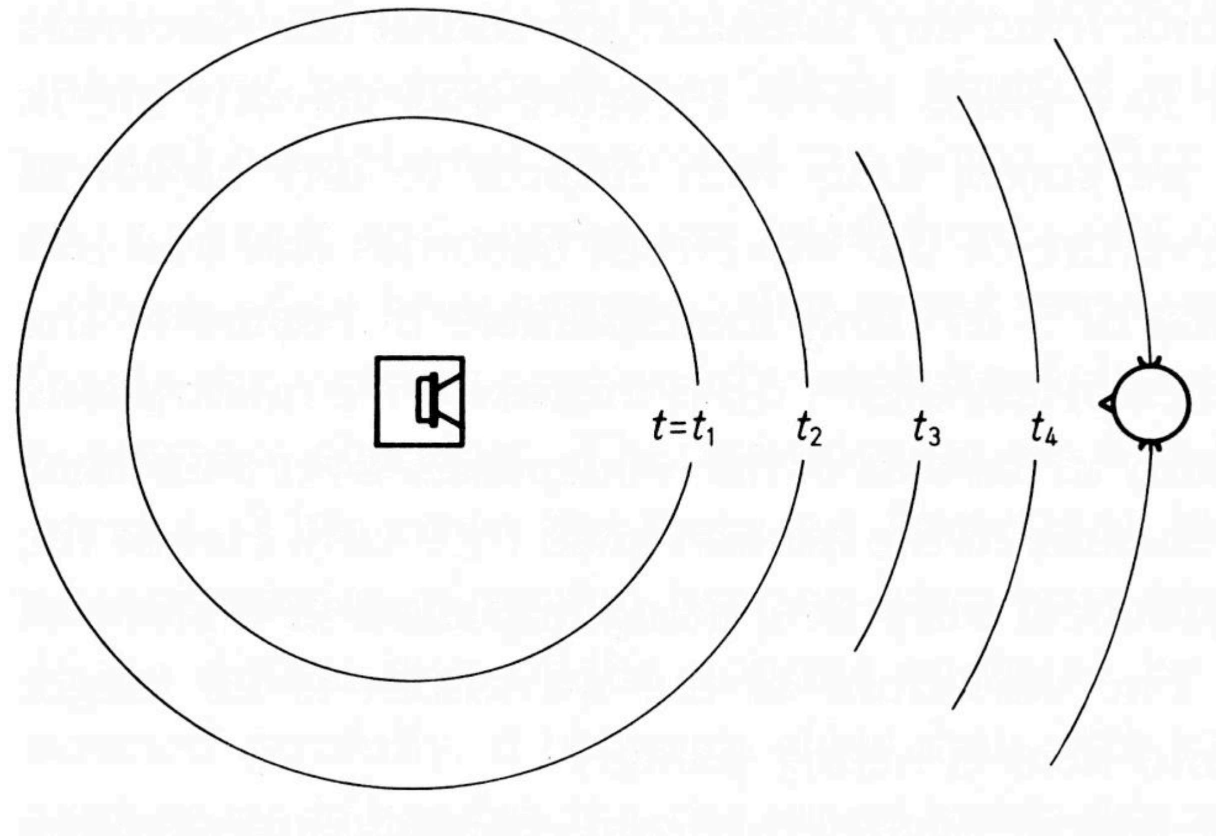
The just noticeable level difference for humans is ~ 1dB.

Perceived sound level is freq. dependent. dB ≠ loudness.

How does sound level attenuate?

$$
\begin{aligned}
\text{intensity} &= \text{power / unit area} \\
I &= P/A \\
&= P/(4\pi r^2) \\
&\Rightarrow 6 \text{ dB per } 2r
\end{aligned}
$$

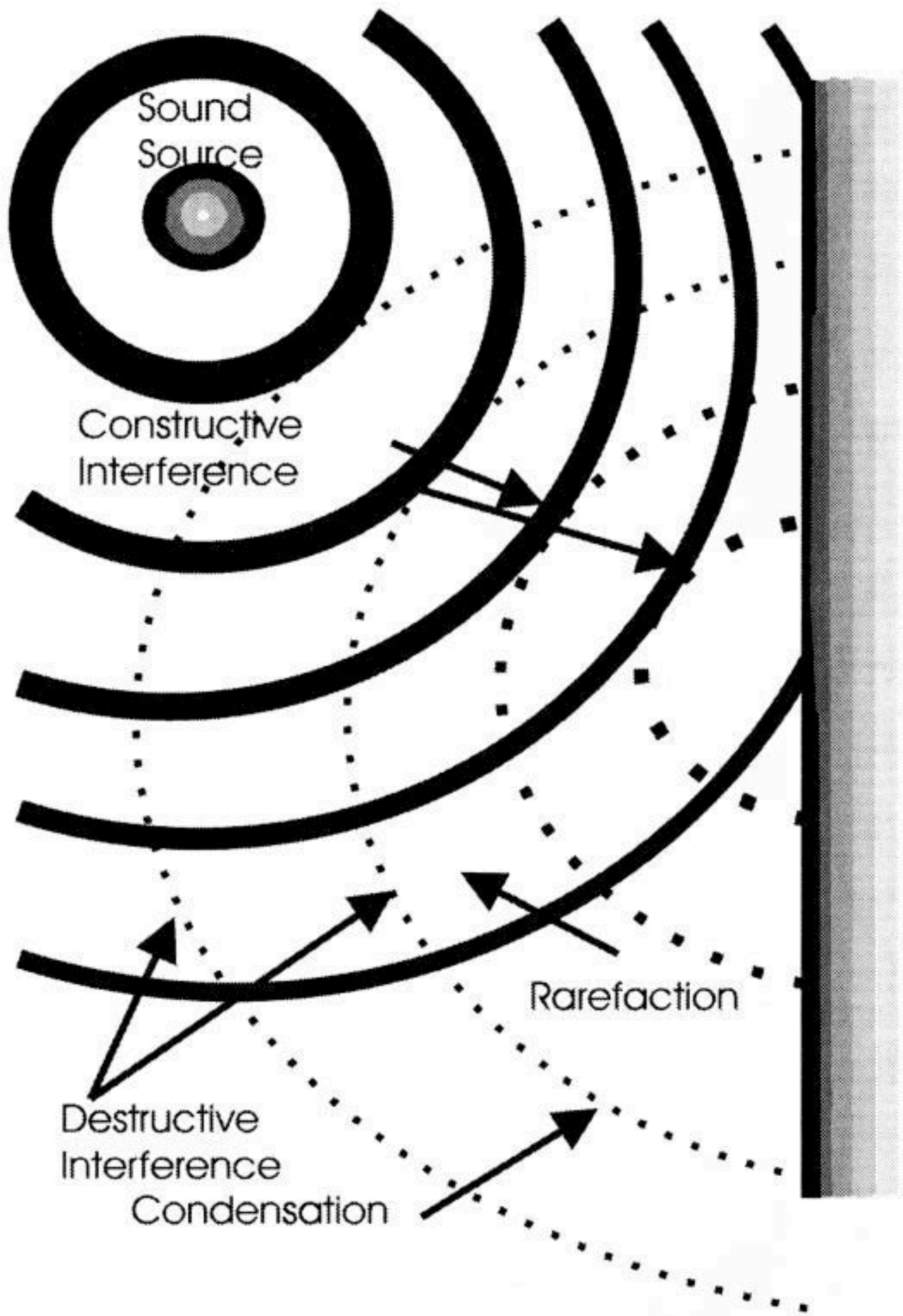
This equation holds for a *free field*.



from Blauert, 1997

How does this fail?

Reflection of sound energy



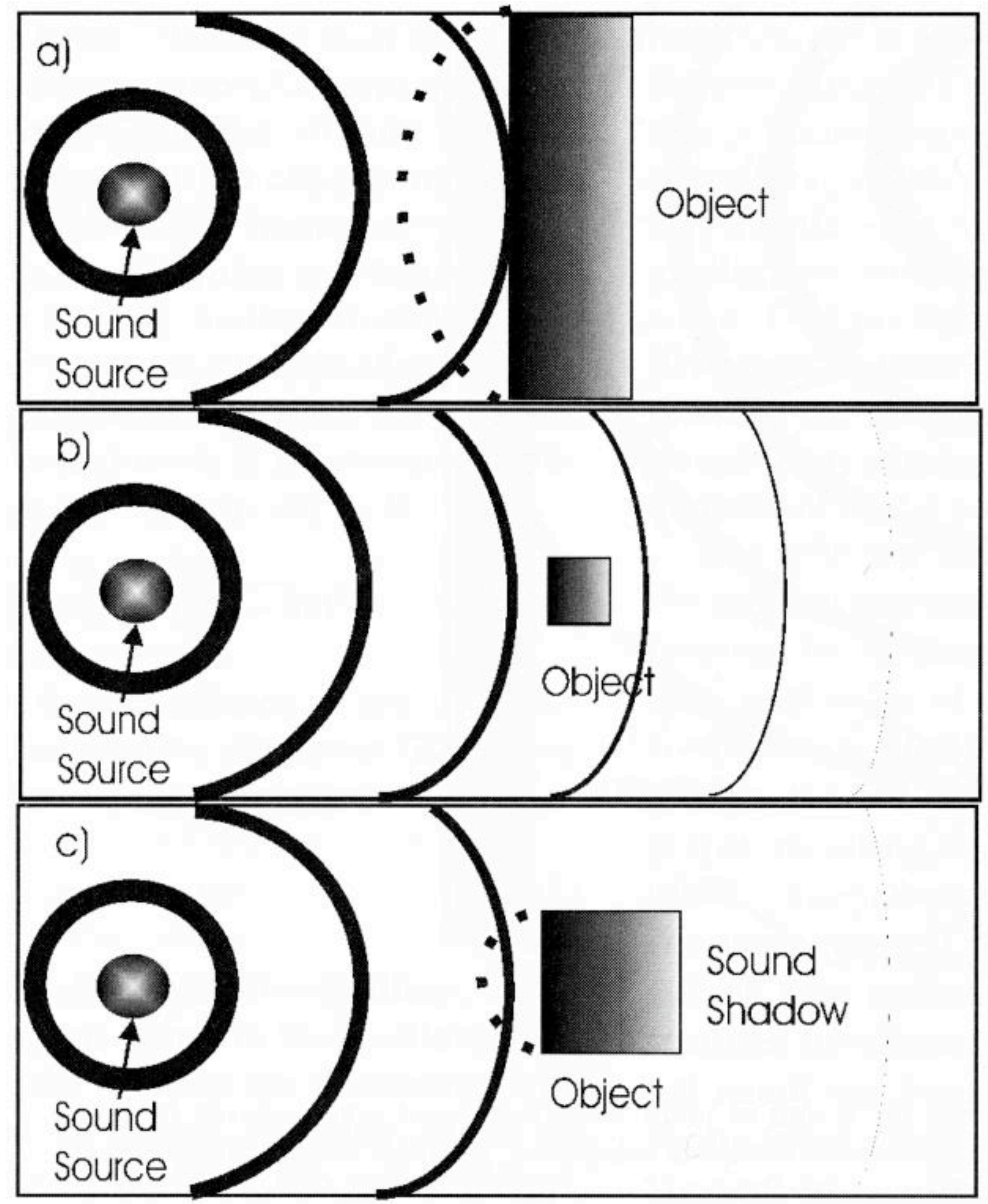
from Yost, 2000

Sound propagation and objects

- Reflection depends on the relative size of the wavelength and the object:
  - "small" wavelengths are reflected and create *interference*
  - "large" wavelengths pass by an object
  - "intermediate wavelengths cast an *acoustic shadow*

- Relevant wavelengths for a typical human head:
  - 345 m/s / 0.18 m = 1916 Hz

- What are factors are there?



from Yost, 2000

Solving the computational problem

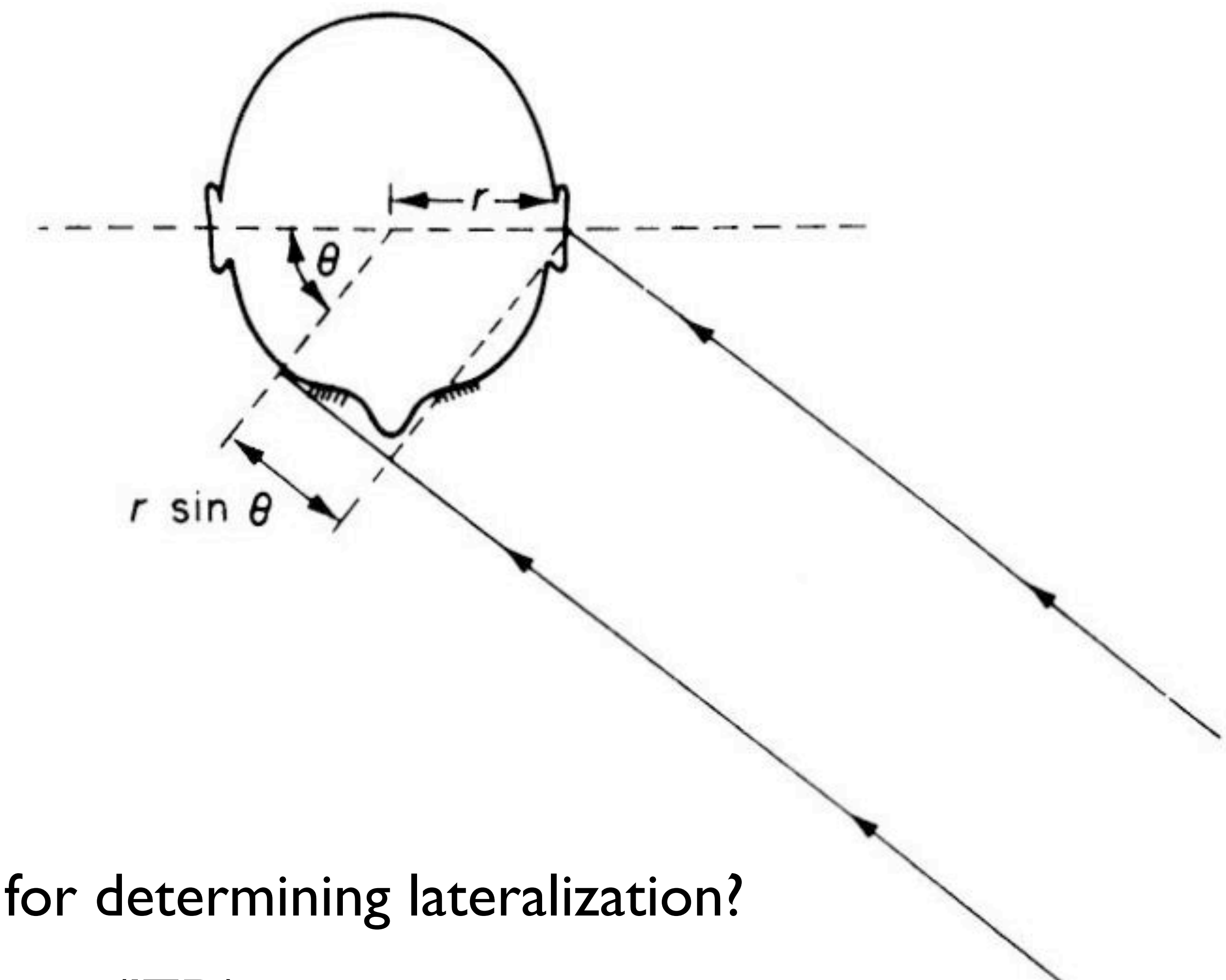
1. Problem statement: what do you want to do?

2. Simplification or idealization

3. Mathematical description of problem

4. Deriving algorithms or formulas

Simple(r) case: lateralization



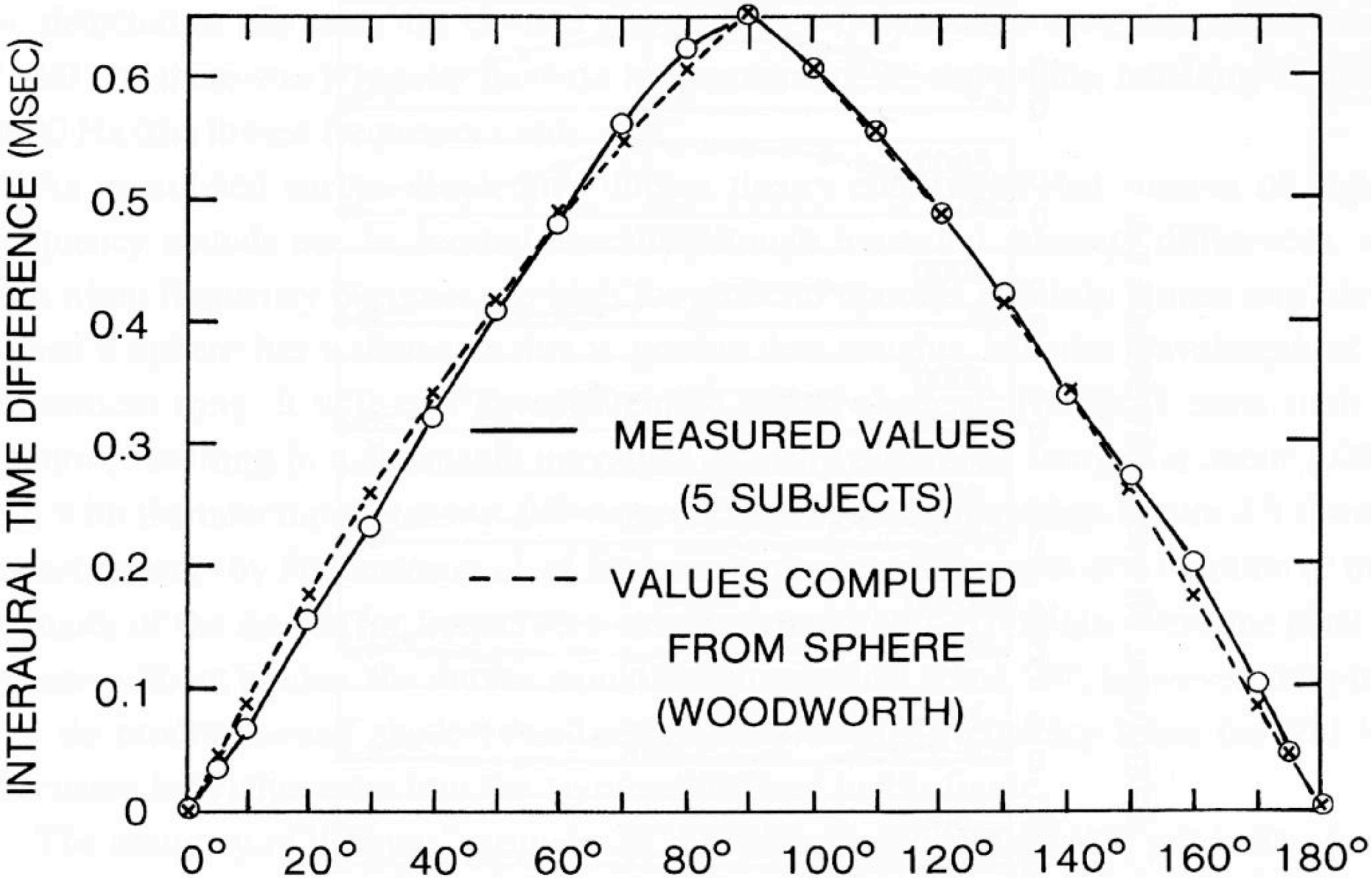
from Moore, 1997

What are the acoustic cues for determining lateralization?

- interaural time differences (ITD)
- interaural intensity differences (IID, or ILDs)

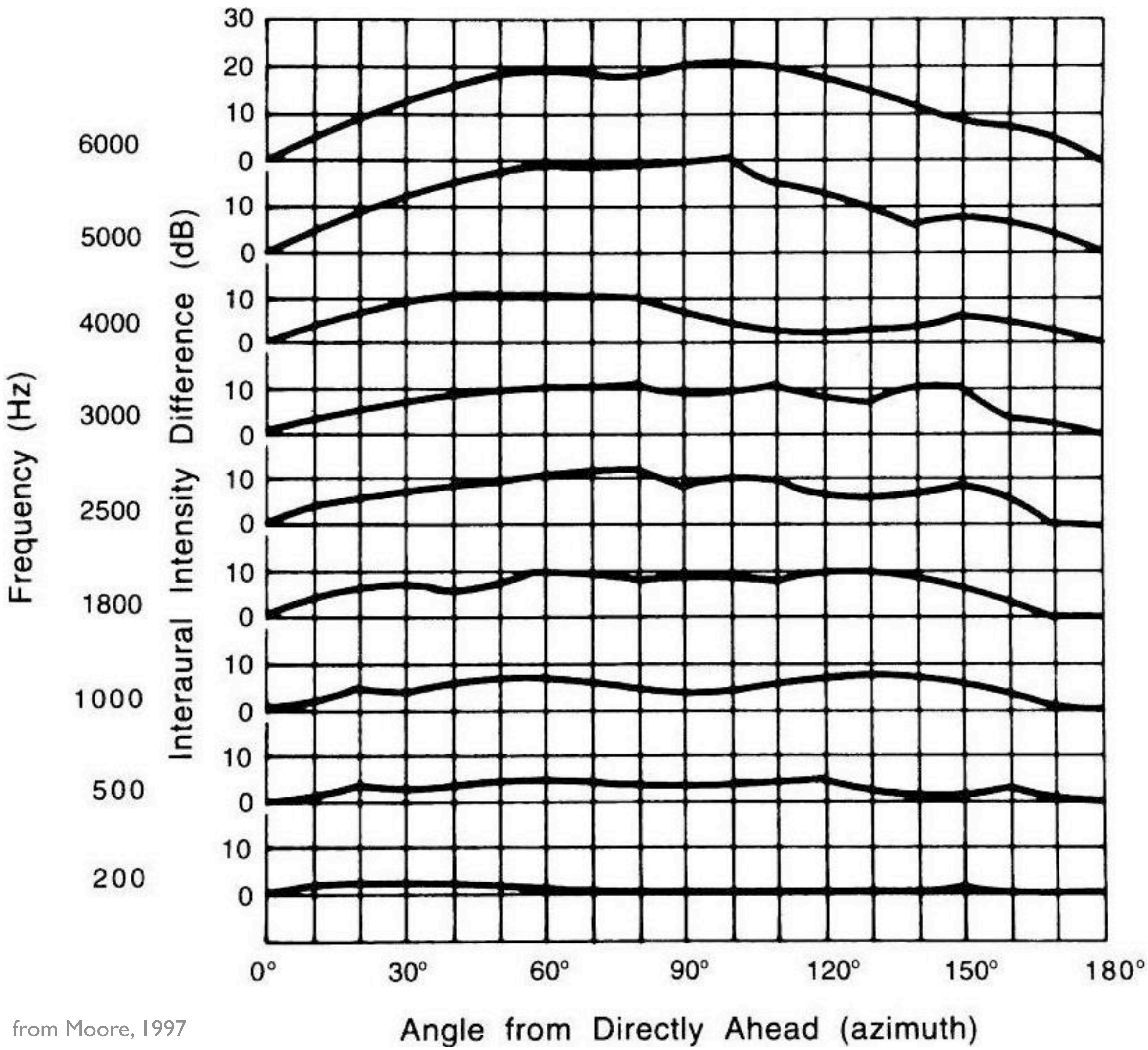
Are these reasonable assumptions?
How could we validate them?

ITD: measured vs predicted



from Warren, 1999

Measured IIDs



from Moore, 1997

What can you observe?

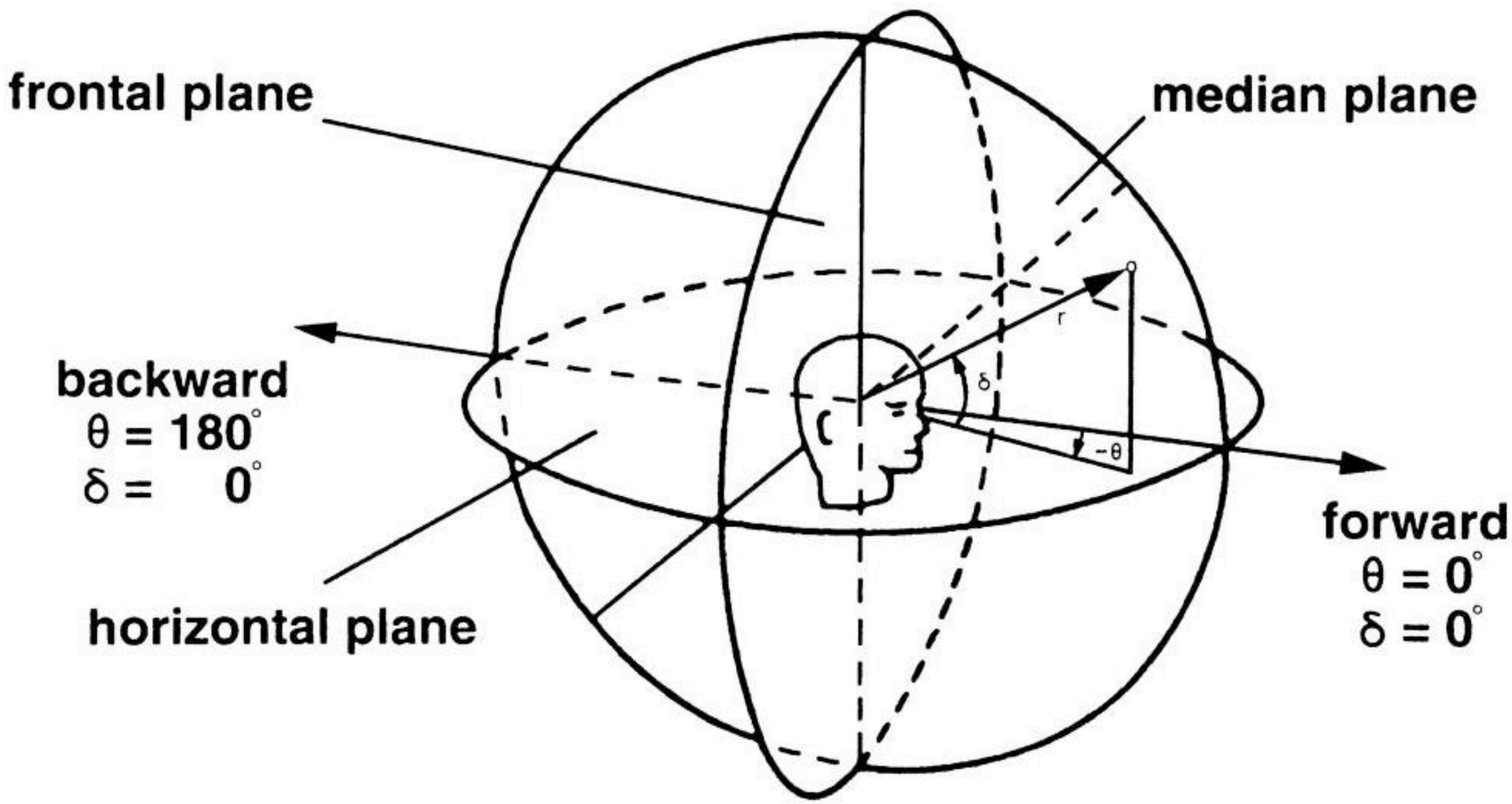
The duplex theory of sound lateralization

Lord Rayleigh, 1907

- Use ITDs for low frequencies.
- Use IIDs for high frequencies.

How could you test this theory?

Localization in 3D space



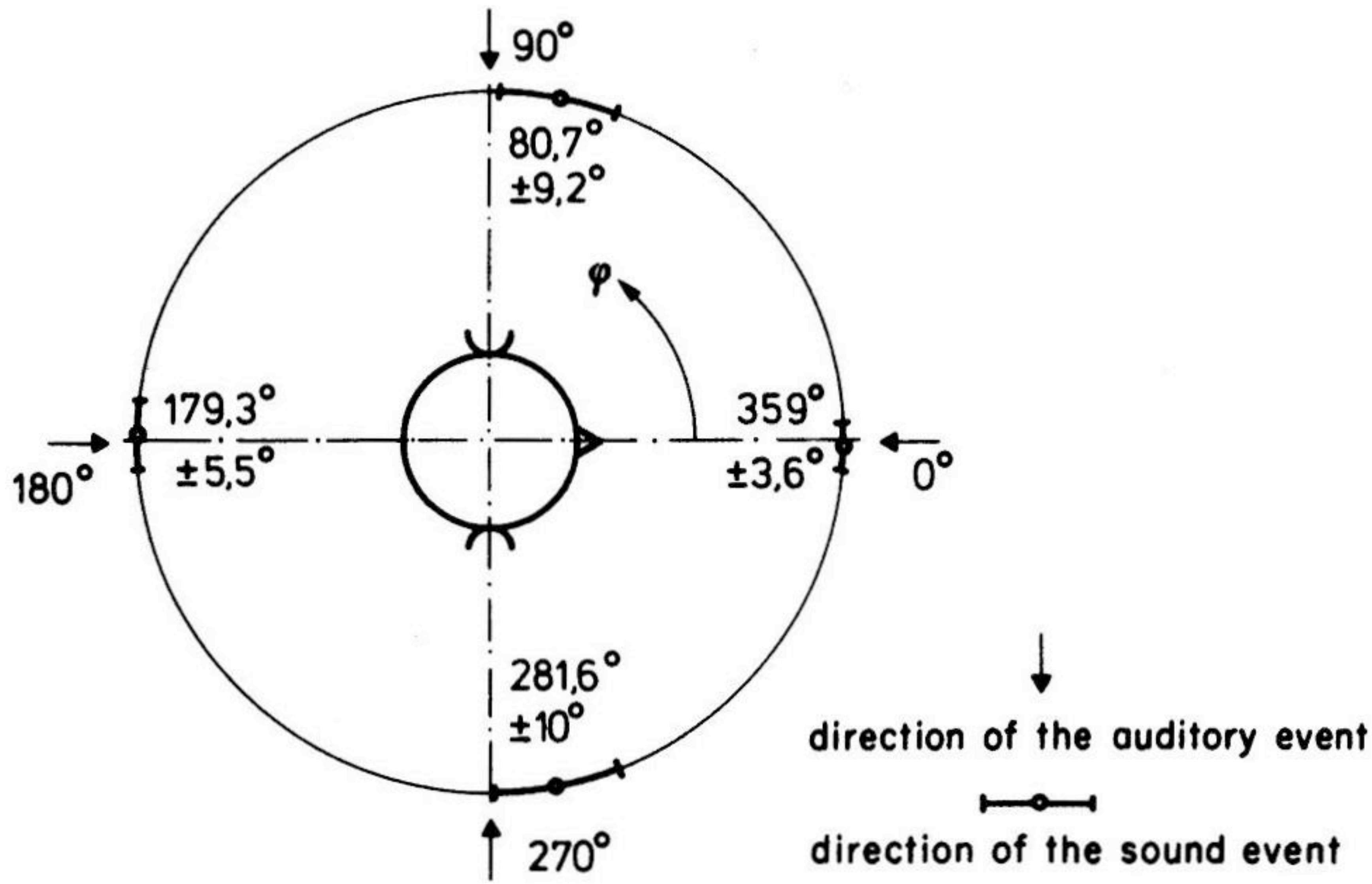
from Moore, 1997

Localization performance

How can we measure human sound localization performance?

- perceived direction vs actual direction
- detection of a sound source shift

Direction acuity: horizontal plane



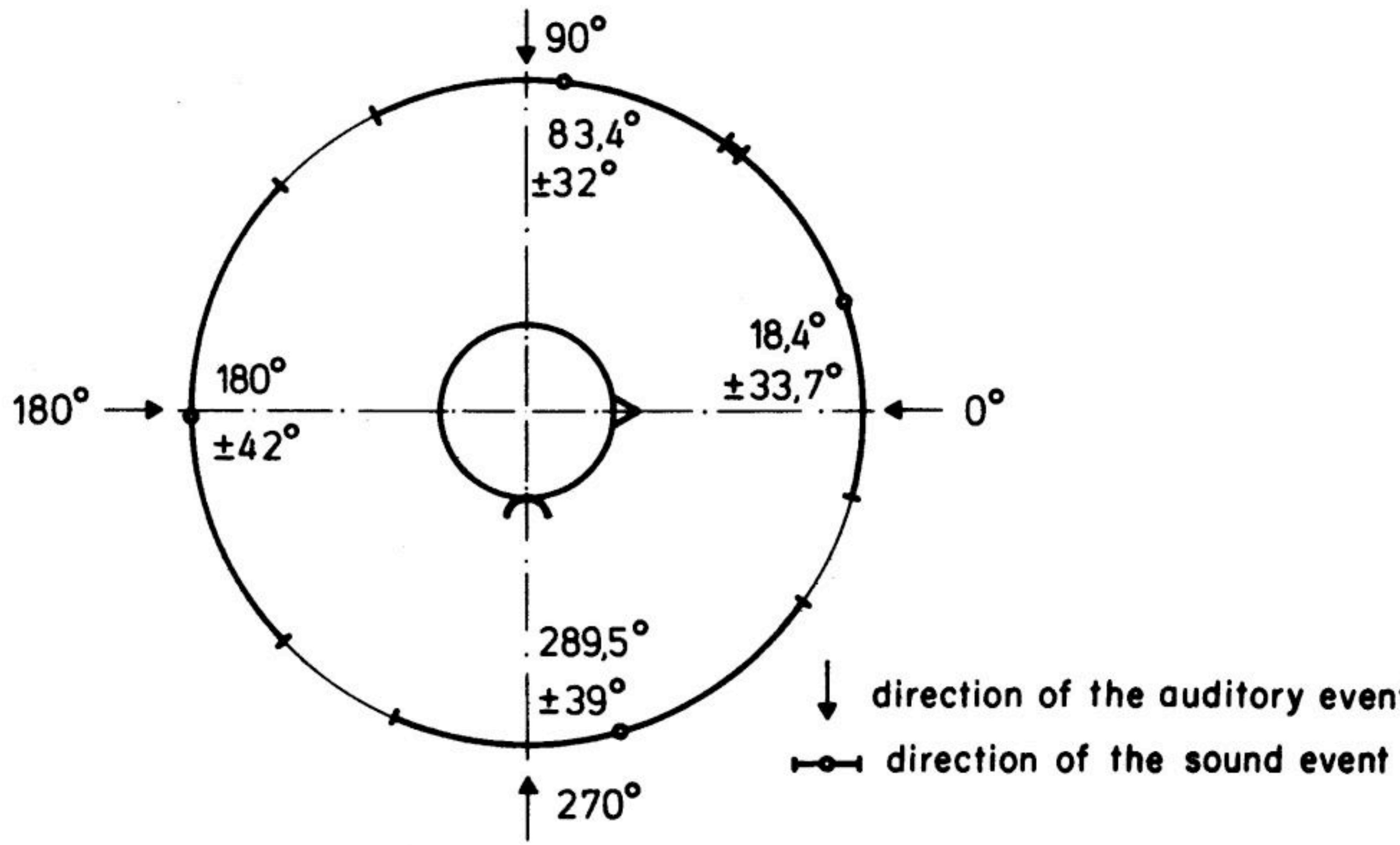
from Blauert, 1997

Sound shift acuity: minimum audible angle

Signal Type	Acuity
clicks	0.75° - 2°
sinusoids	1° - 4°
tone bursts	0.8° - 3.3°
speech	0.9° - 1.5°
noise	3.2°

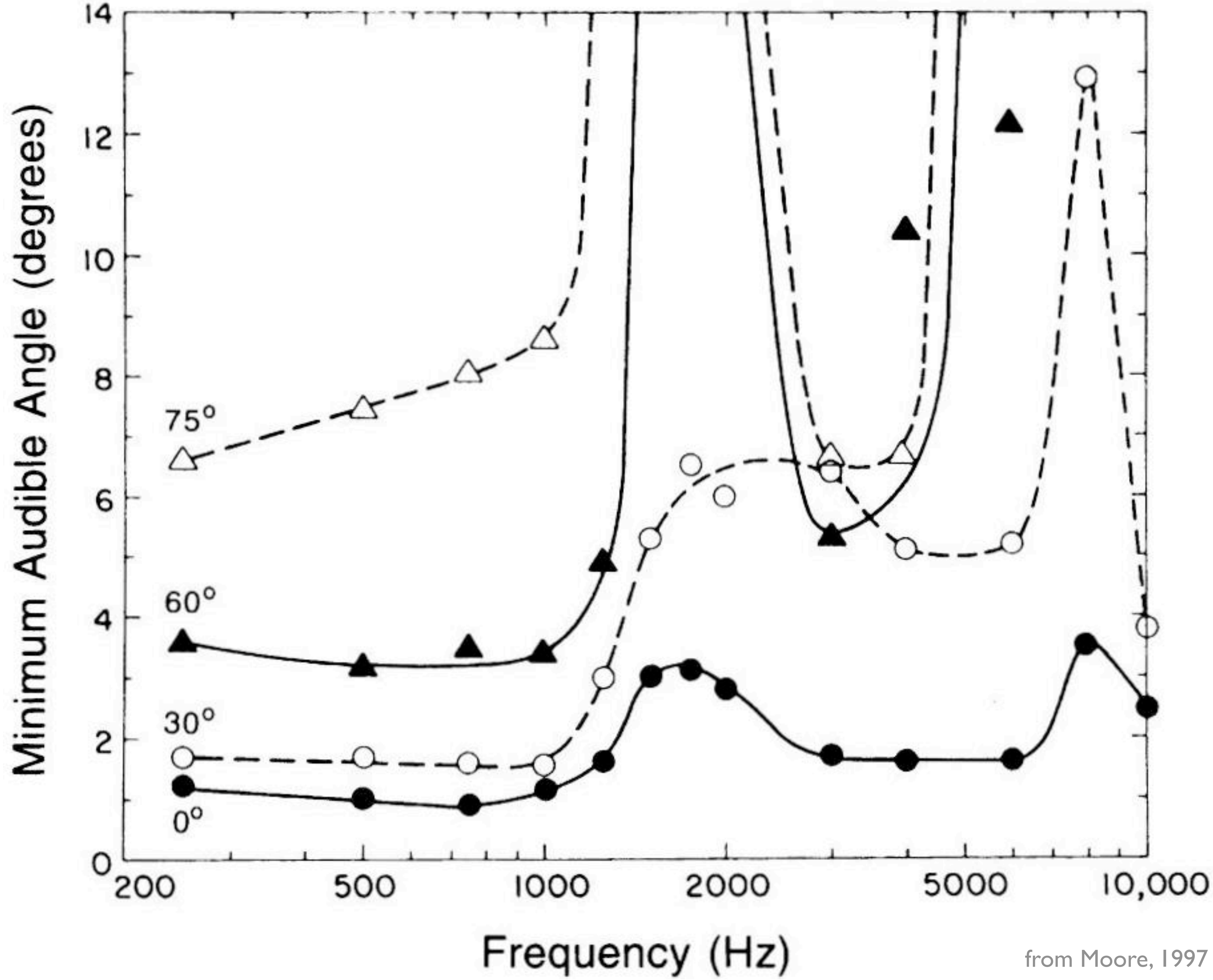
Measuring minimal perceivable displacement from forward direction, under various "ideal" conditions.

Direction acuity: medial plane



from Blauert, 1997

MAA vs frequency for different azimuths



from Moore, 1997

Experimental Tests of the Duplex Theory

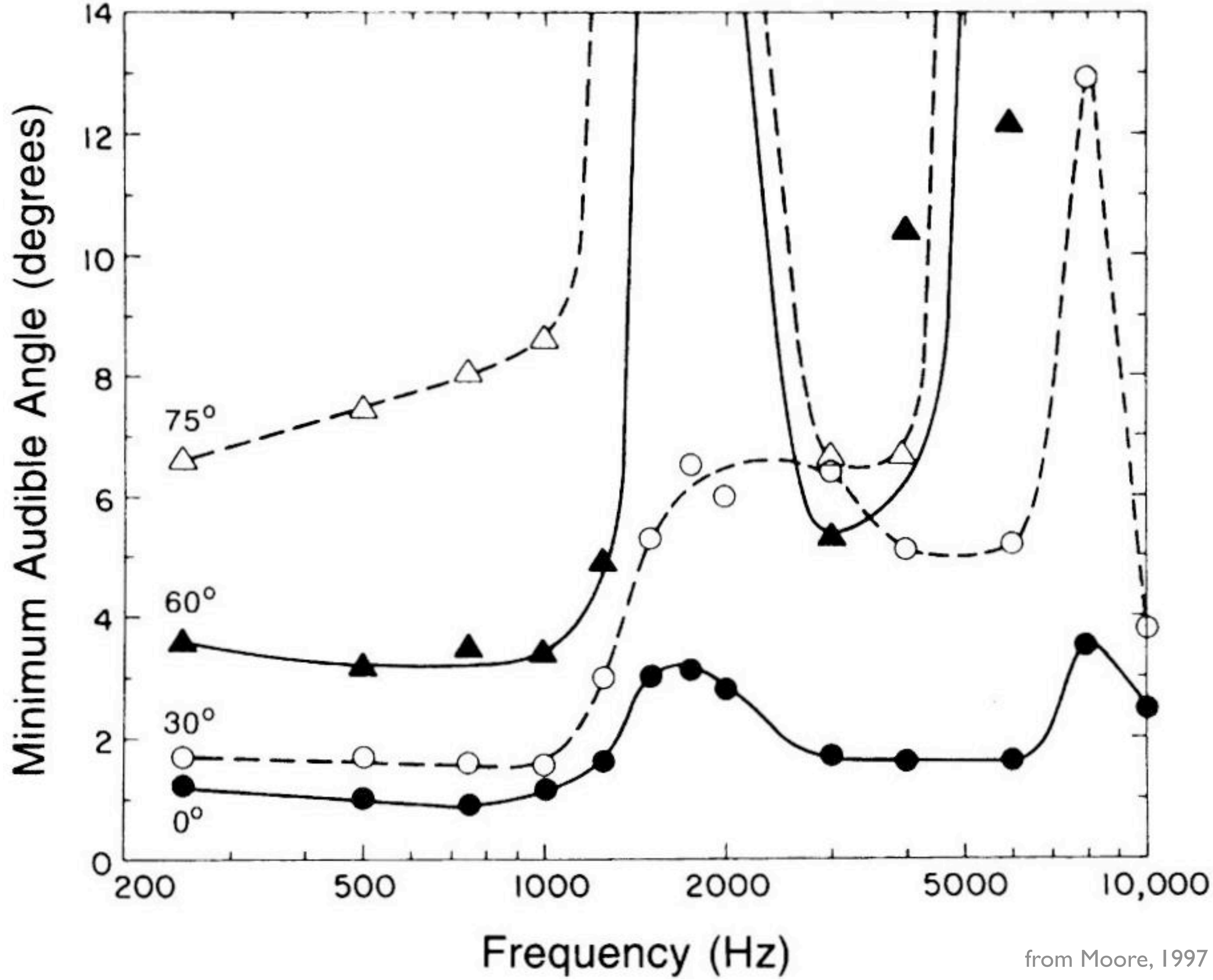
ITD:

- most sensitive at 0°
- smallest detectable change is ~10 µs or 1°
- $f > 900$ Hz, ITD sensitivity drops dramatically

IID:

- most sensitive at 0°
- smallest detectable change is ~1 dB
- Real world IIDs are small $f < 1800$ Hz, but still perceptible.

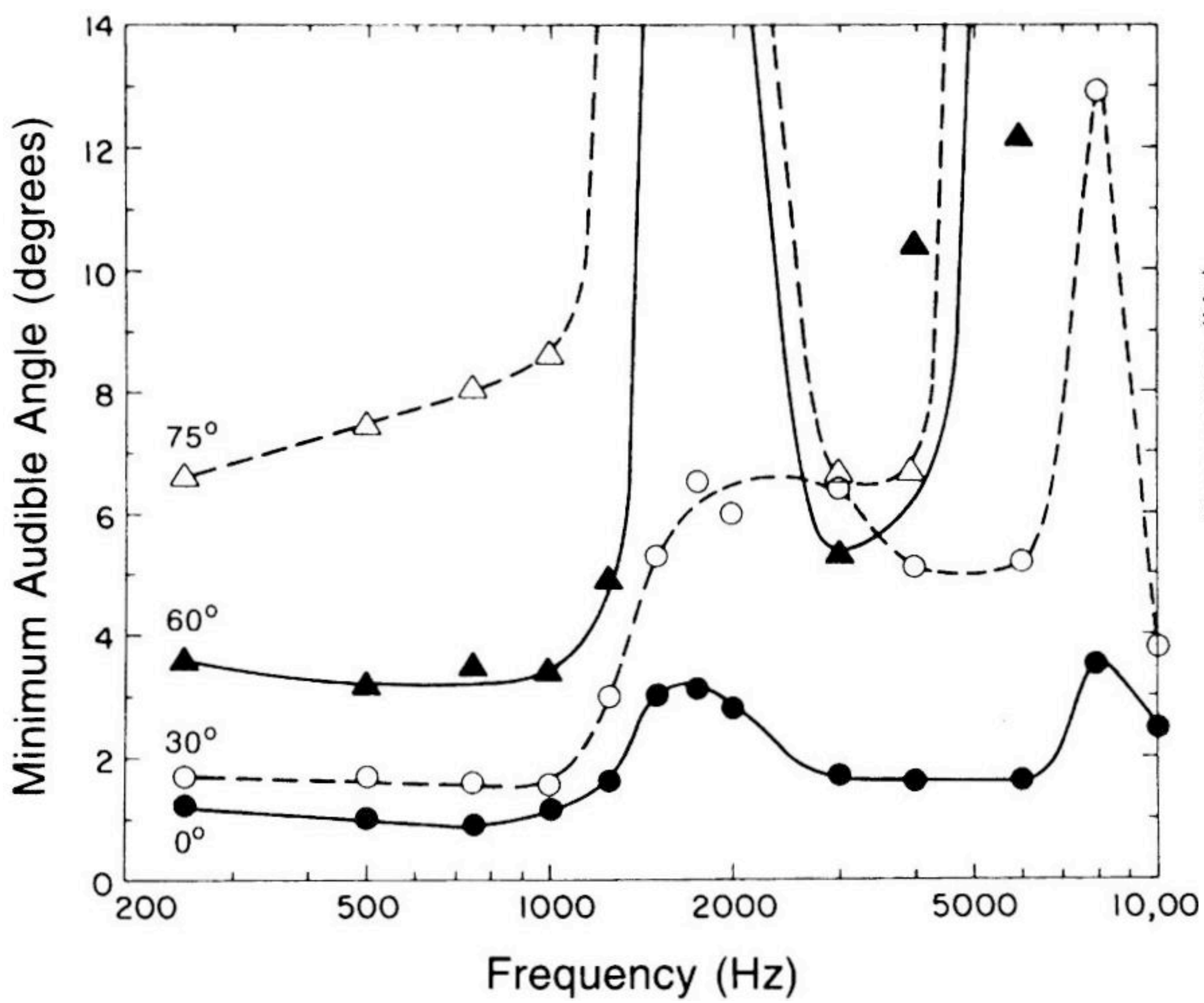
MAA vs frequency for different azimuths

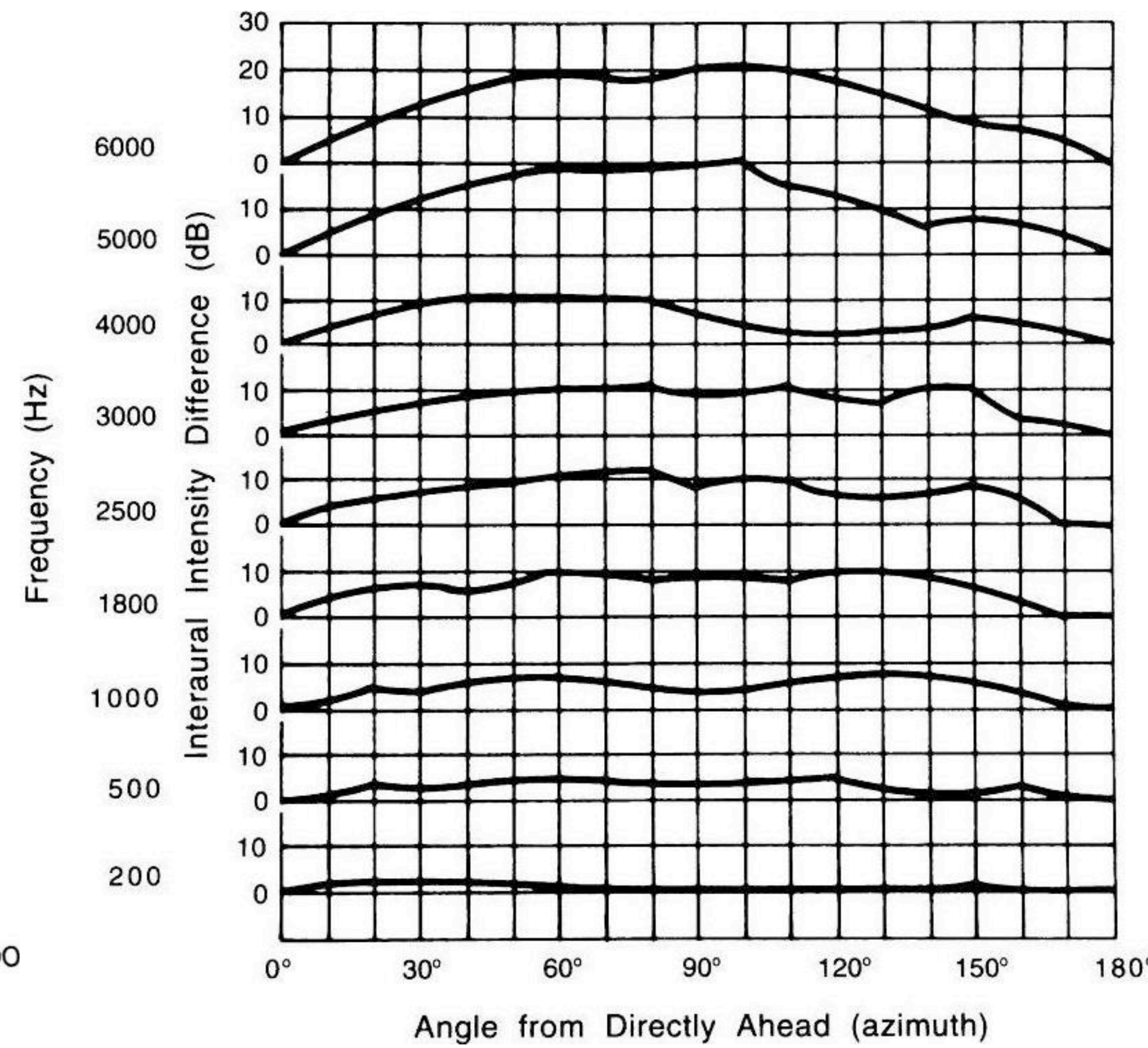


from Moore, 1997

How does this compared to measured ITD and IID?

Can we explaining the MAA data?





How do we compute ITD?

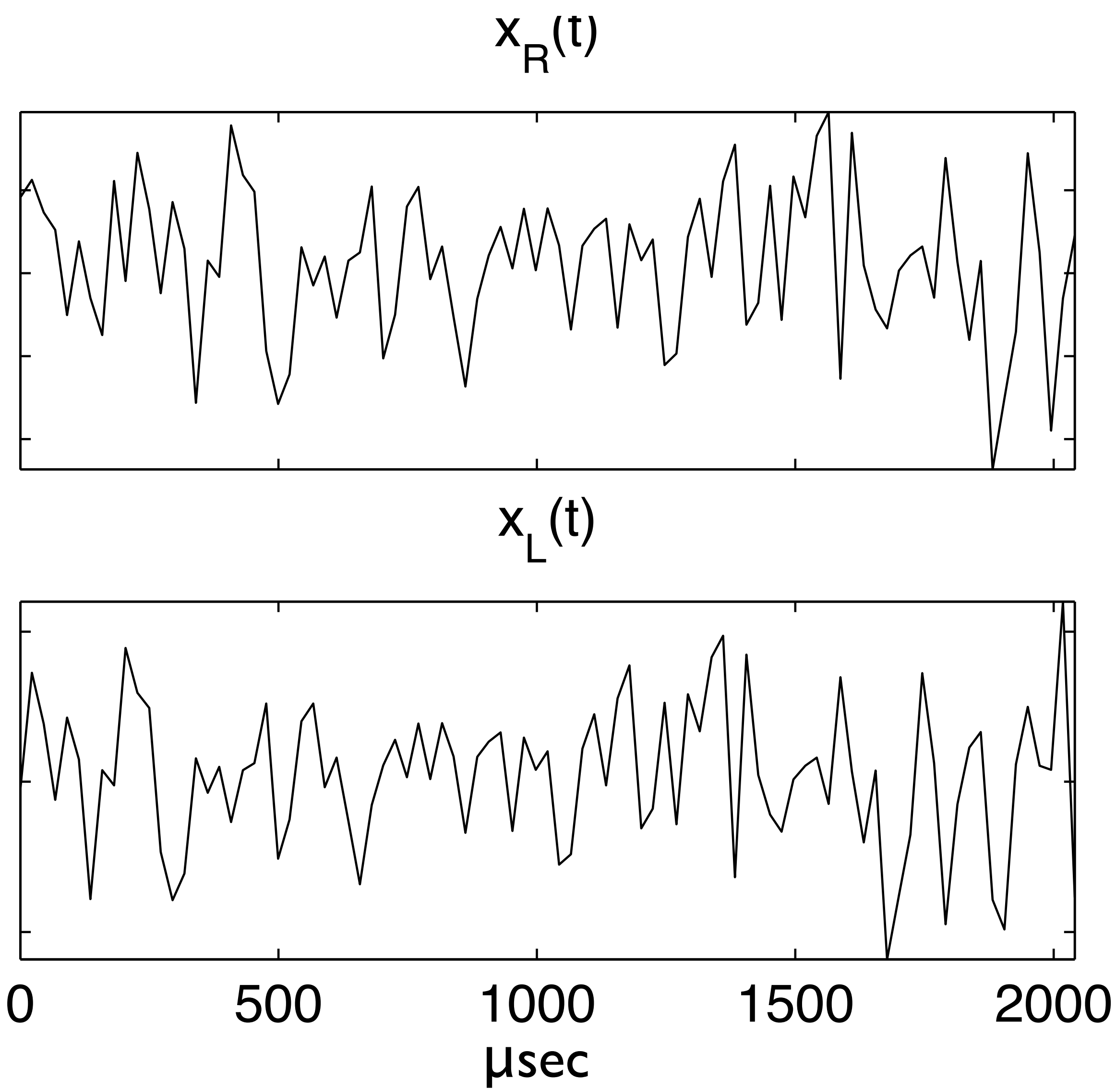
- Cross correlation:

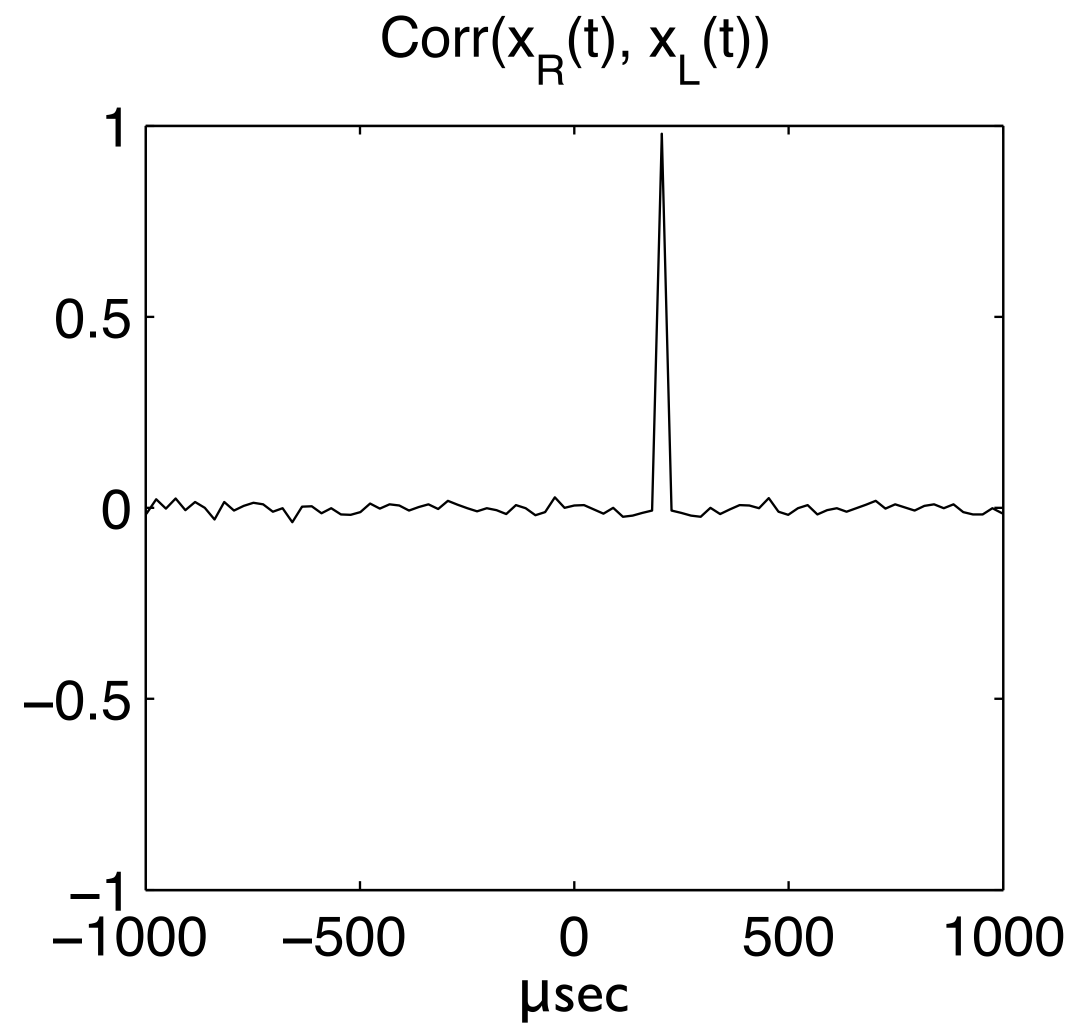
$$\begin{aligned}\mathrm{Corr}(x_R, x_L)(t) &= \frac{1}{Z}\int_{-\infty}^{\infty} x_R(\tau)x_L(t+\tau)d\tau \\ &= \frac{1}{Z}\sum_{k=1}^{N-t} x_R(k)x_L(k+t)\end{aligned}$$

- This assumes signals have zero mean.
- Correlation is normalized so Corr(x_R, x_L) = 1 when $x_R = x_L$

$$Z = \int_{-\infty}^{\infty} x_R(\tau)x_L(\tau)d\tau$$

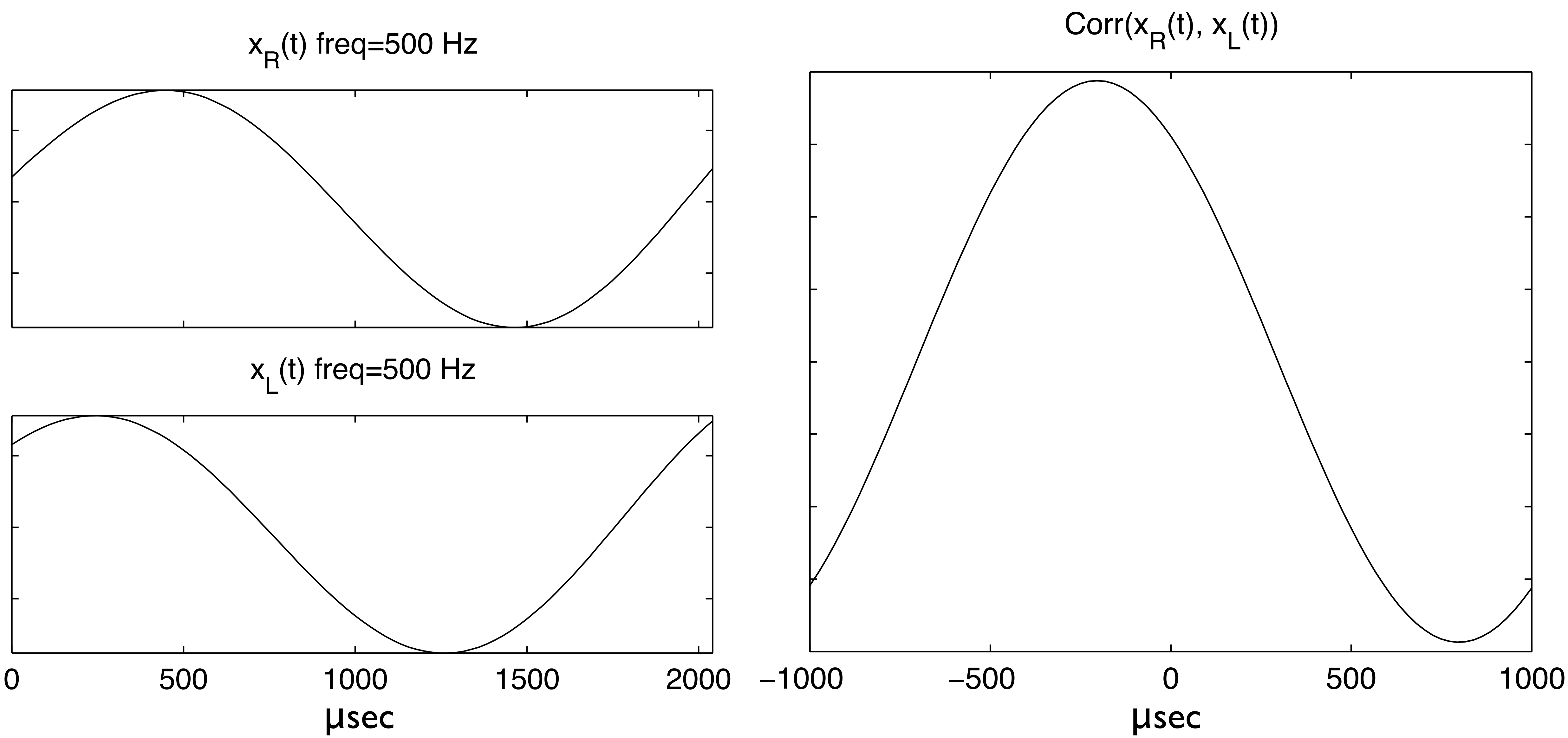
Example: white noise



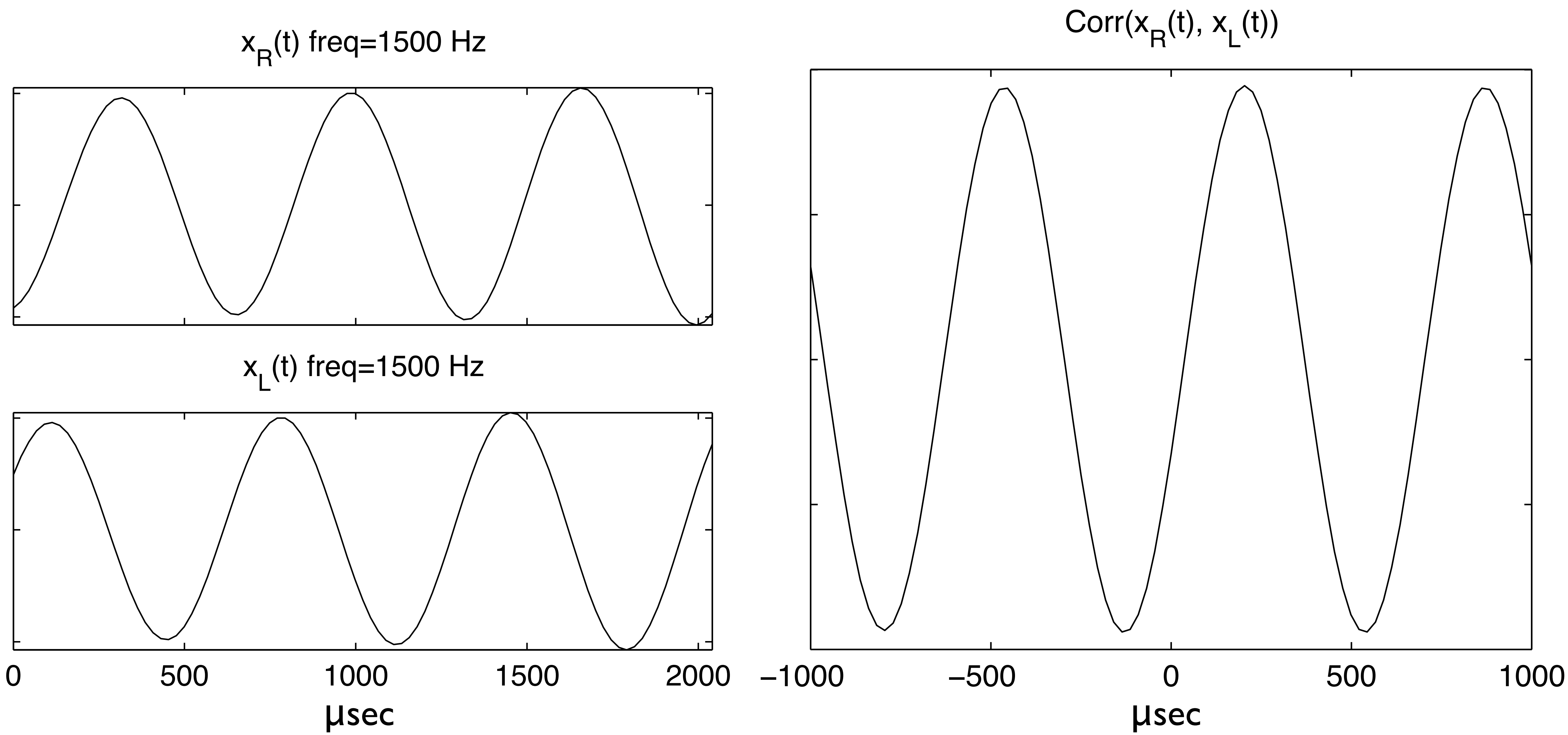


Will this work for all signals?

Example: low frequency sinusoids



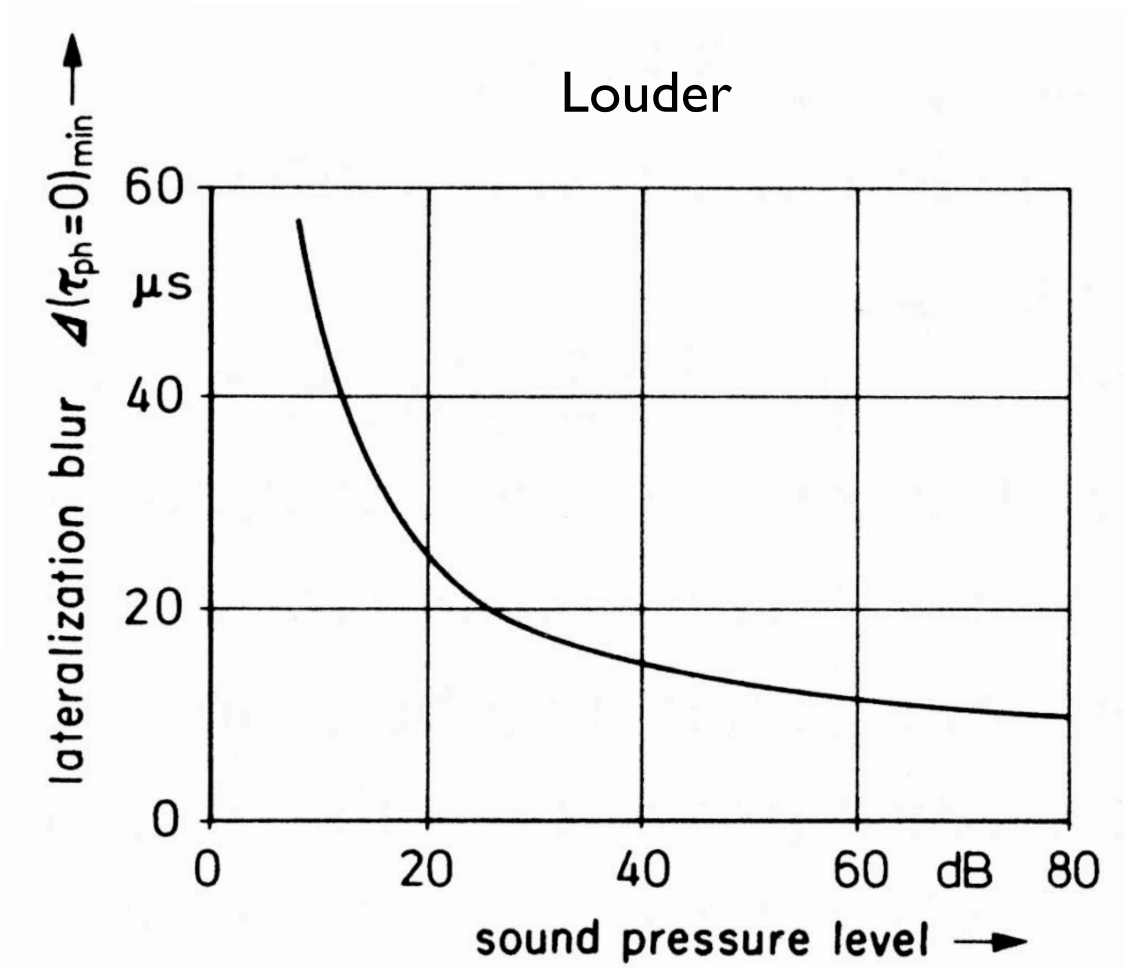
Example: high frequency sinusoids

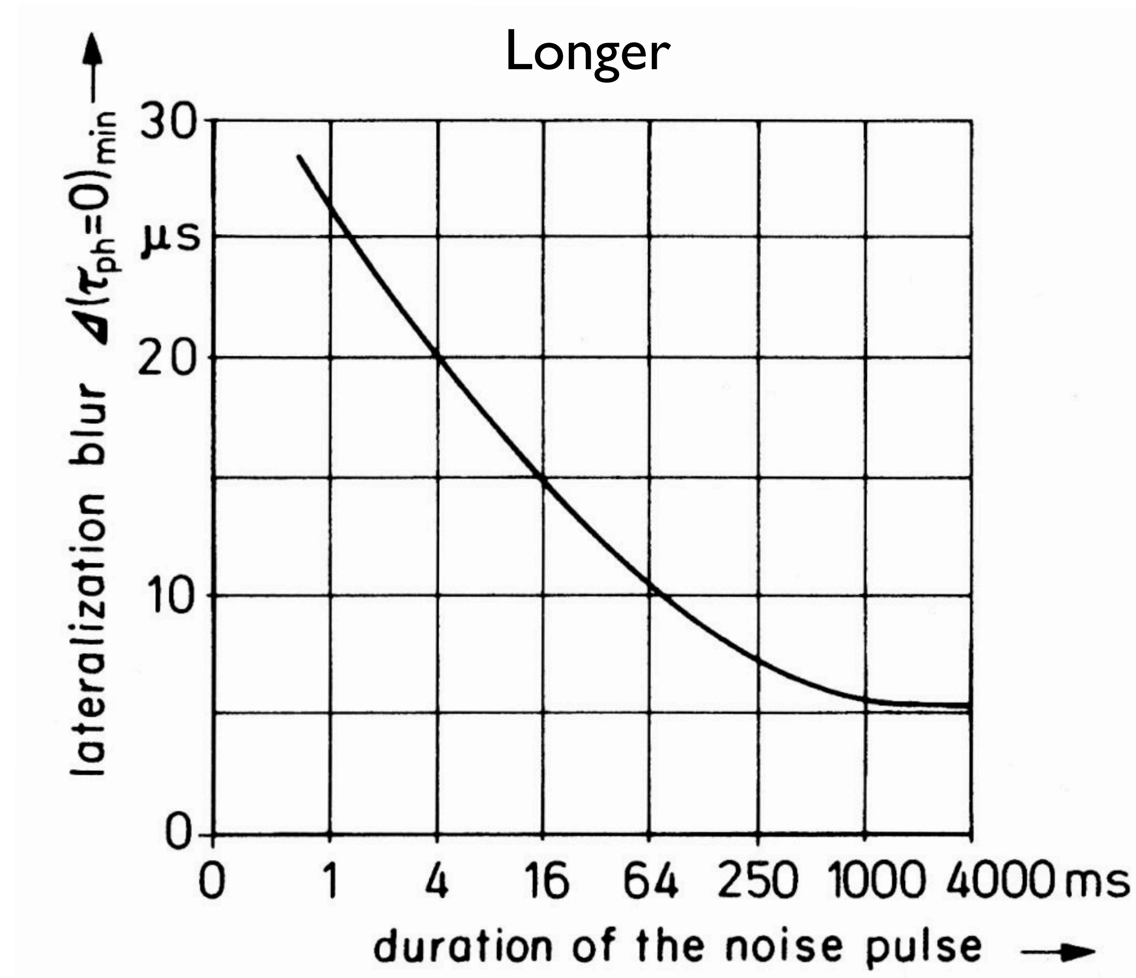


This is called *phase ambiguity.*

What do people hear?

More signal improves localization accuracy





from Blauert, 1997

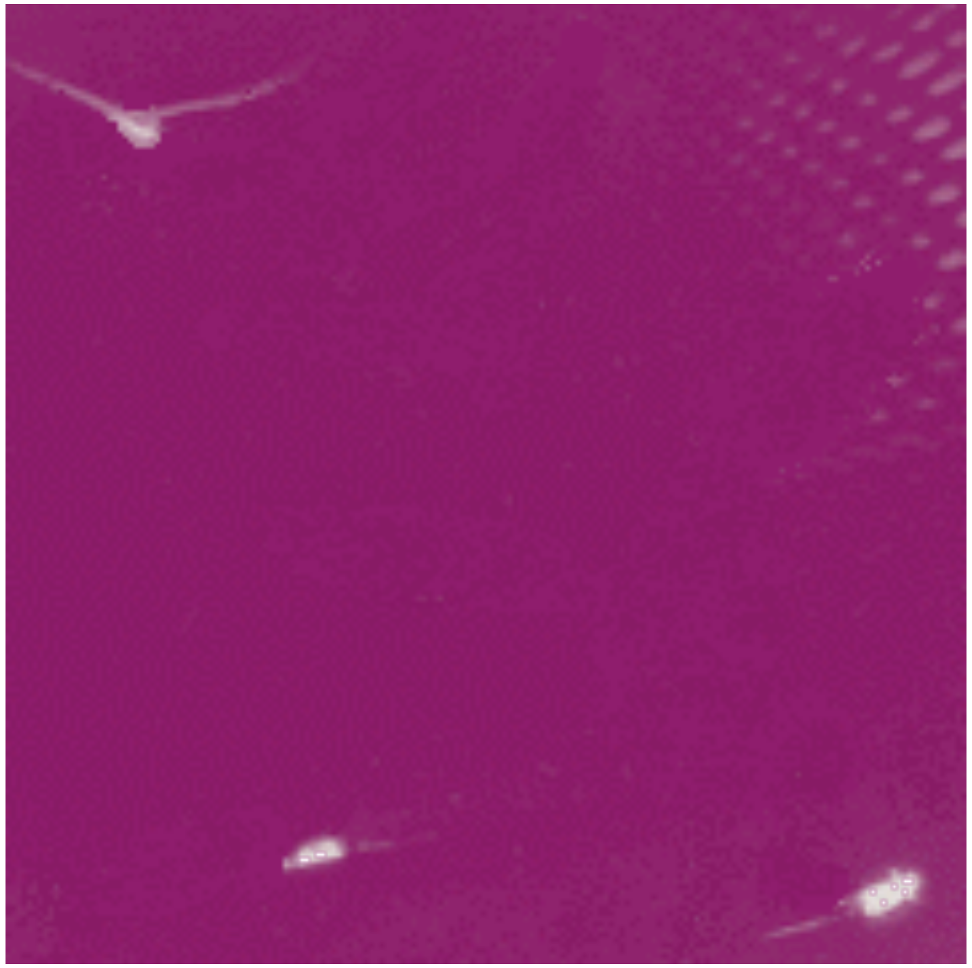
Limitations of the Duplex Theory

- limited to lateralization
- doesn’t do front-back discrimination
- doesn’t explain why are sounds are outside your head
- neglects acoustic environment
- can’t handle multiple sound sources

The Barn Owl

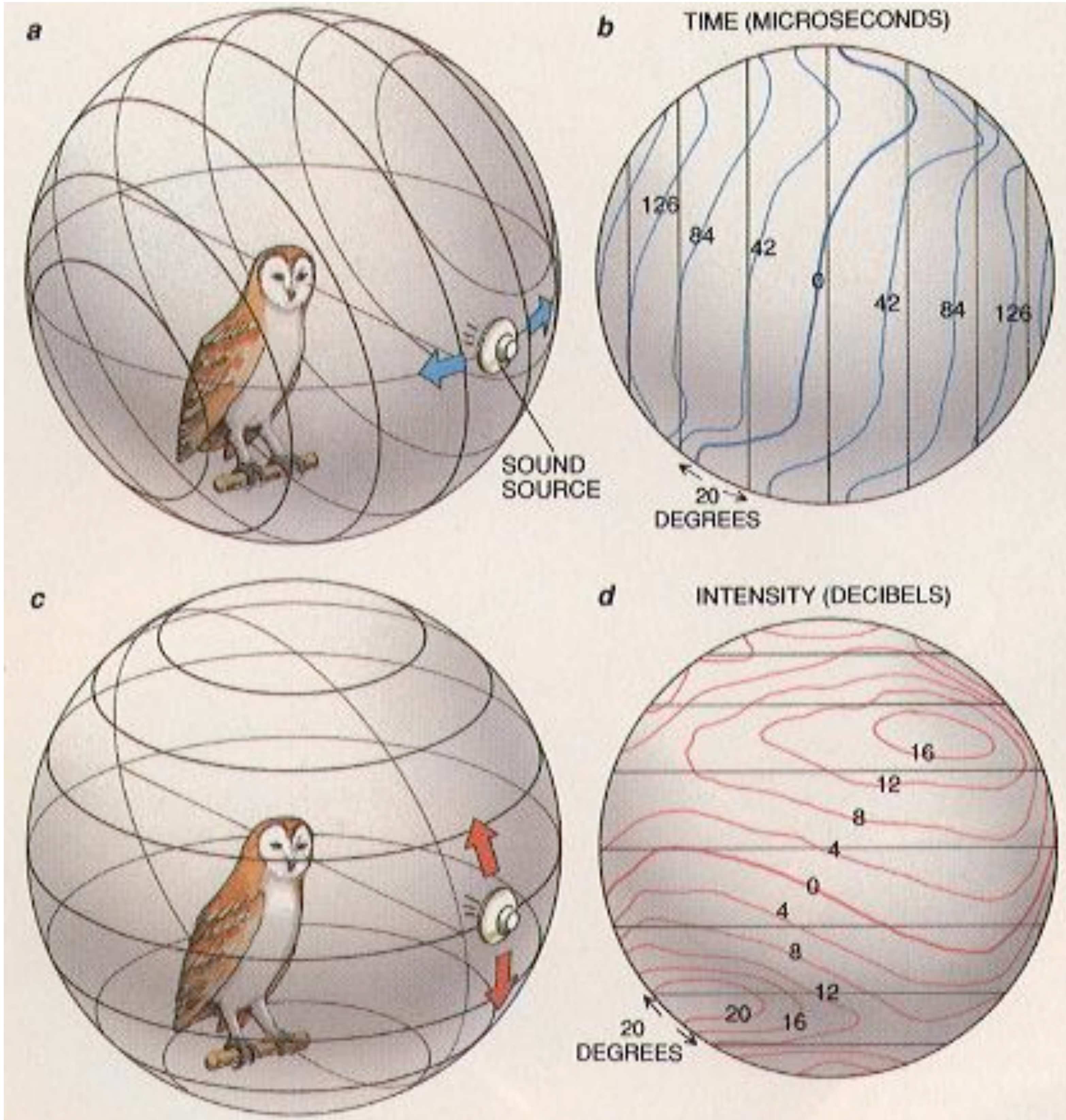
Filmed with a stroboscopic infrared camera.

The owl grabs the mouse in total darkness.



images by M. Konishi

3D Localization in the barn owl



from Konishi, 1993

On the scientific method

Georg von Békésy (1899-1972): Won the 1961 Nobel Prize in Physiology for his discovery of how sound is transduced in the cochlea.

"Of great importance in any field of research is the selection of problems to be investigated and a determination of the particular variables to be given attention."

"When a field is in its early stage of development the selection of good problems is a more hazardous matter than later on, when some of the principles have begun to be developed."

Taxonomy of Scientific Problems (von Békésy, 1960)

- Theoretical vs mosaic approach
- The classical problem: much effort but no solution
- The premature problem: poorly formulated or not susceptible to attack
- The strategic problem: seeks data to distinguish between two basic assumptions or principles
- The stimulating problem: may lead to re-examination of accepted principles and may open up new areas of exploration
- The statistical question: a survey of possibilities
- The unimportant problem: easy to formulate easy to solve
- The embarrassing question
- The pseudo problem: differences in terminology or methods of approach

Issues

- What are the scientific questions?
  - What do you want to understand?
  - What do you want to explain?

- What are the computational problems?
  - What is our scope?
  - Do we need to solve the same problem in a similar way?

- What behavioral questions would you ask?

Next time: sound localization in 3D space