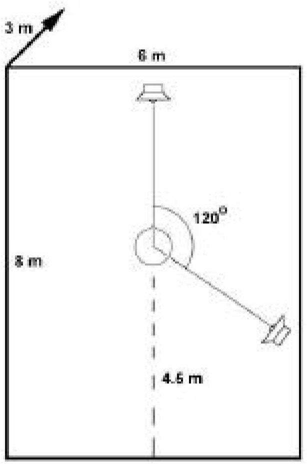
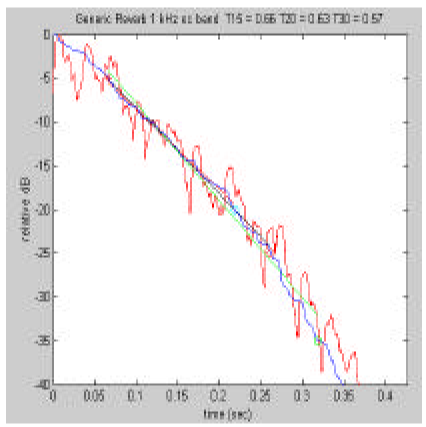
Challenges and solutions for realistic room simulation





ASA June, 2002

Durand R. Begault

Human Factors Research and Technology Division

NASA Ames Research Center

Moffett Field, California

Background

- Perception of the acoustical environmental context
- Basic technique of auralization (virtual room synthesis)

Challenges for improved simulation

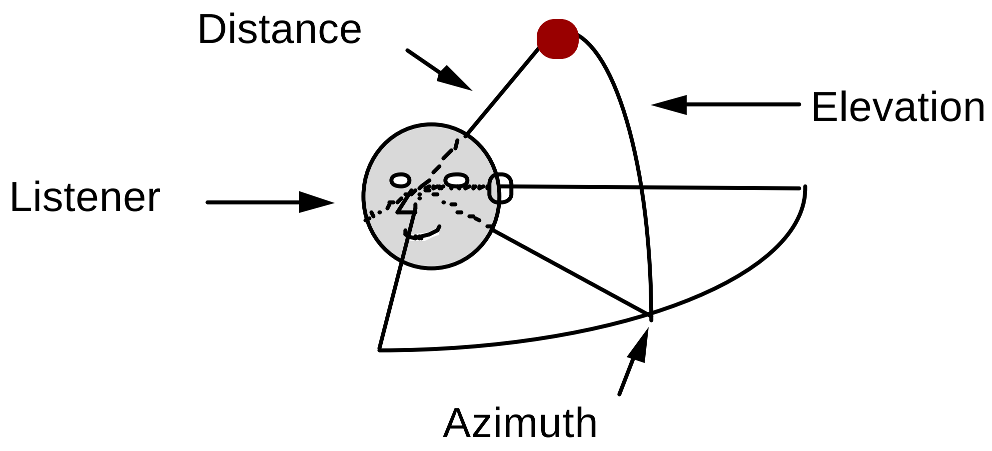
- Determining acceptable data reduction in measurement and synthesis (*echo thresholds)*

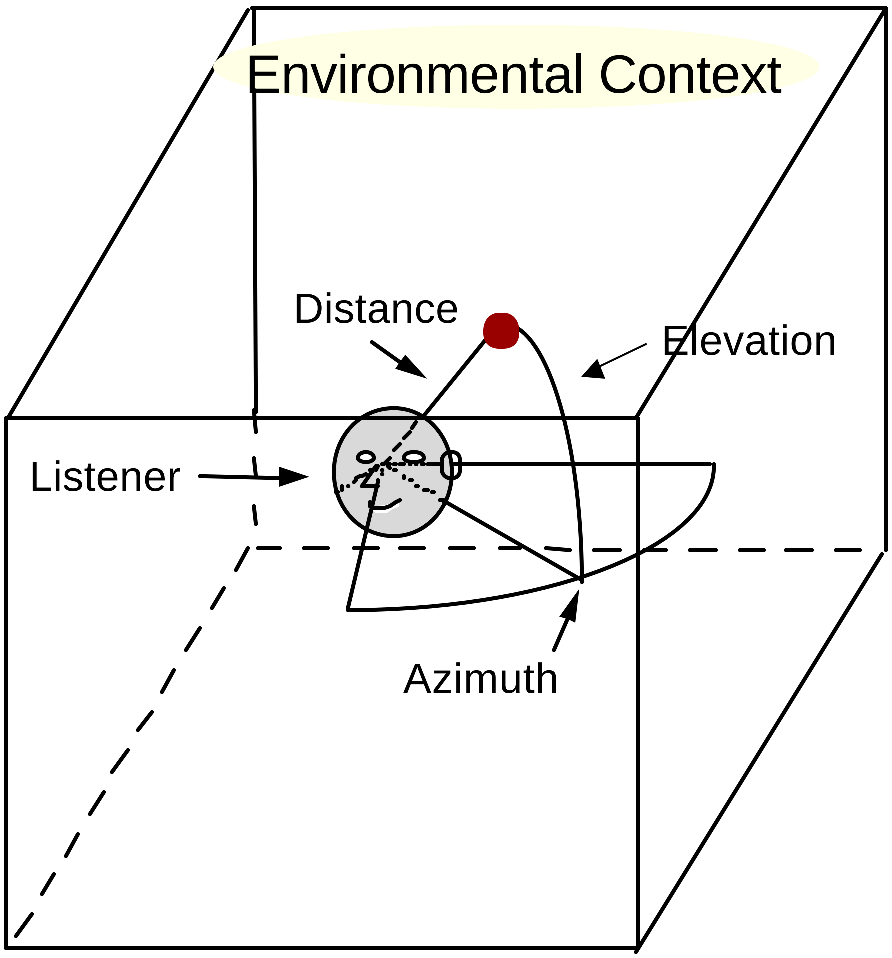
- Simulation of low frequency energy, coupled spaces

- Accounting for localization error due to reproduction method

- Accurate simulation of source directivity

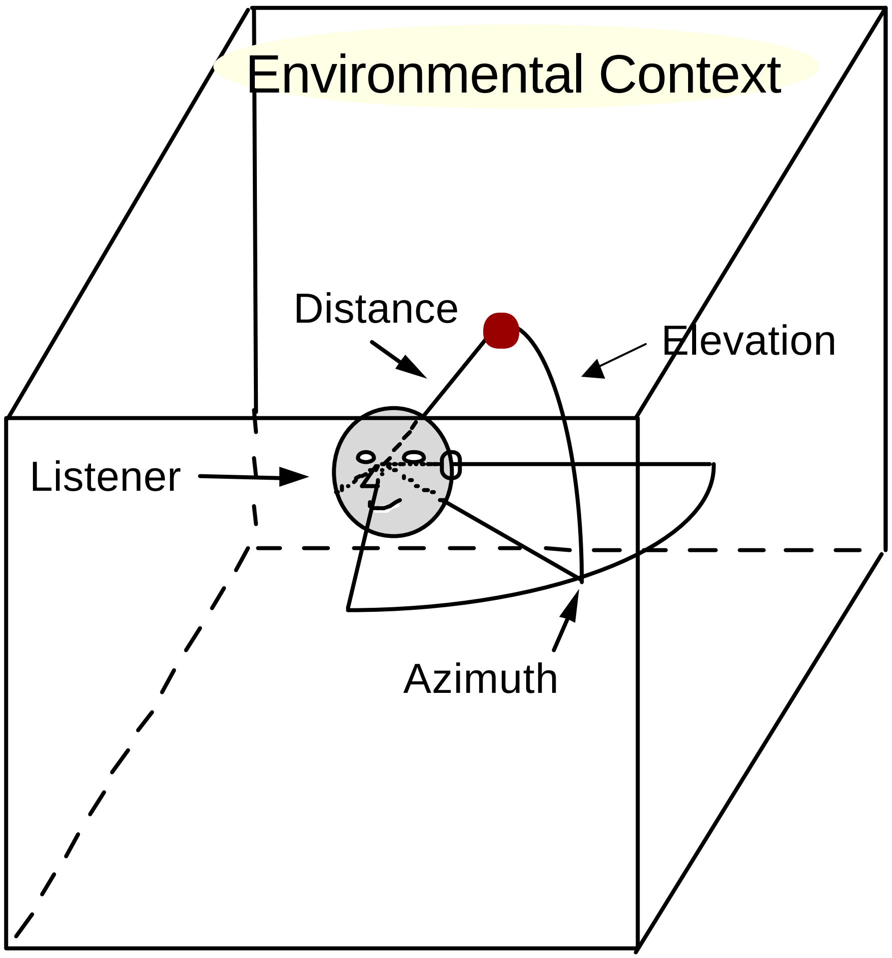
Spatial hearing fundamentally involves perception of the location of a sound source ● at a point in space (azimuth, elevation, distance).





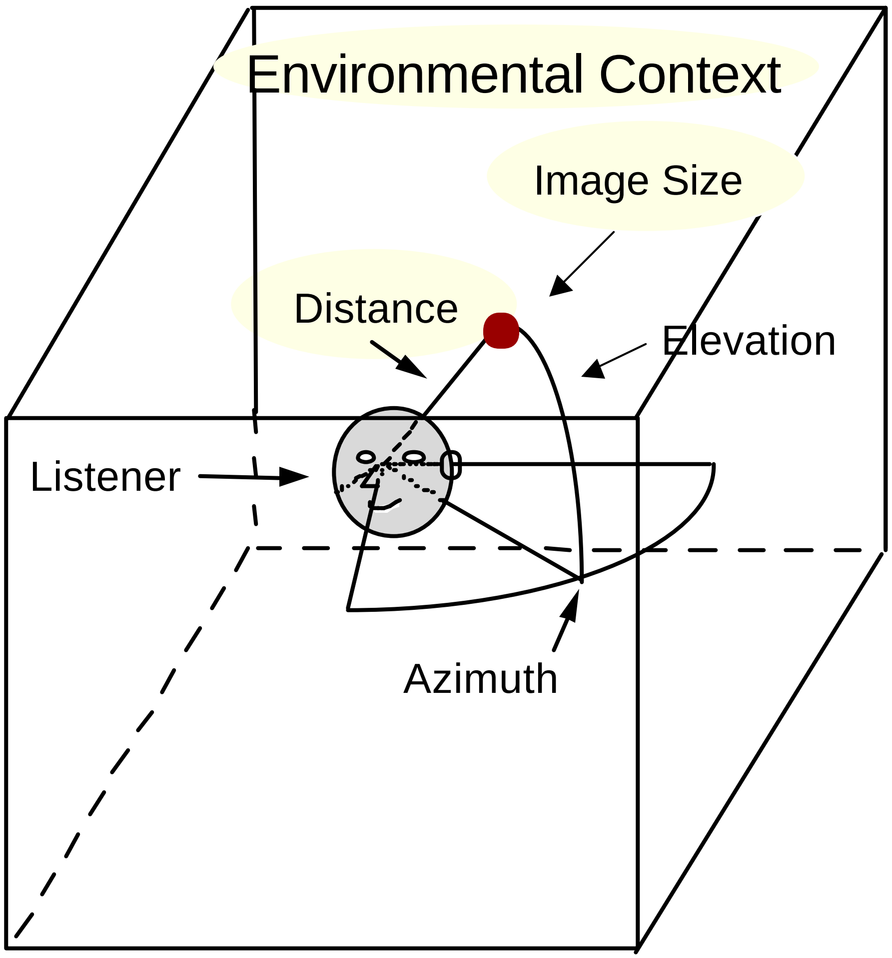
Spatial (and non-spatial) hearing simultaneously reveals information about the environmental context (EC) of a sound source.

- Reverberance, echoes, surface characteristics reveal dimension, extent of space via cognitive associations



Spatial (and non-spatial) hearing simultaneously reveals information about the environmental context (EC) of a sound source.

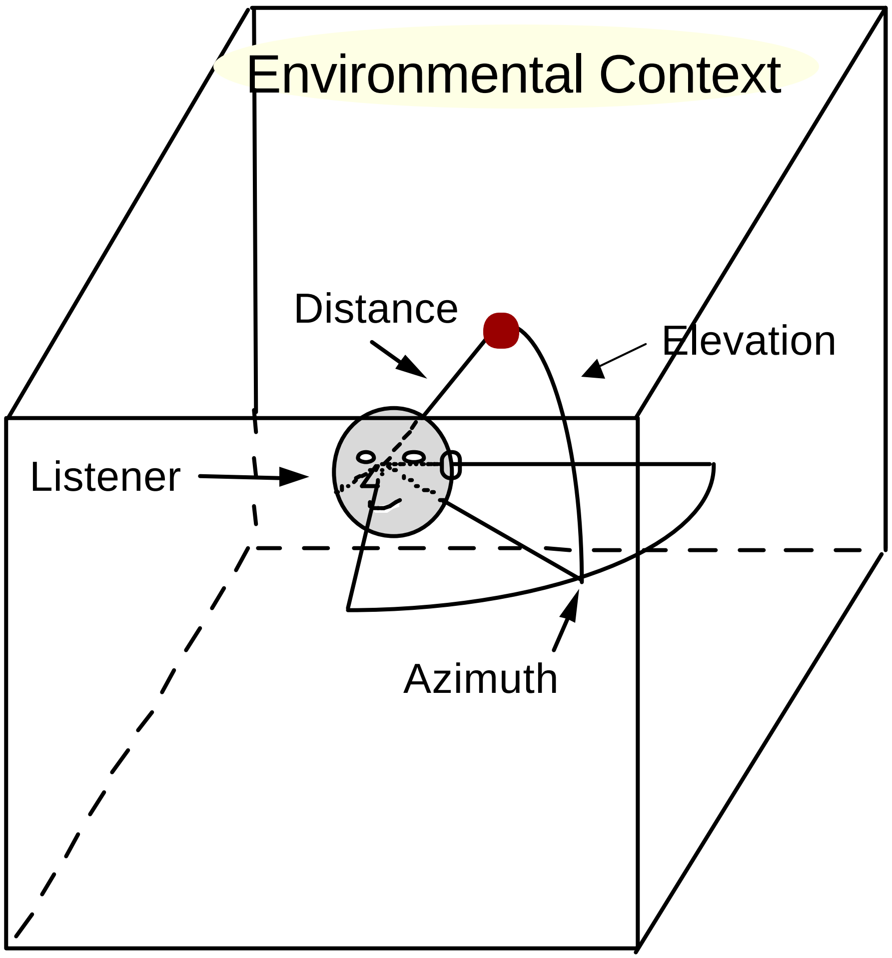
- Reverberance, echoes, surface characteristics reveal dimension, extent of space via cognitive associations
- Interactive cues: EC's effect on speech effort, intelligibility



Spatial (and non-spatial) hearing simultaneously reveals information about the environmental context (EC) of a sound source.

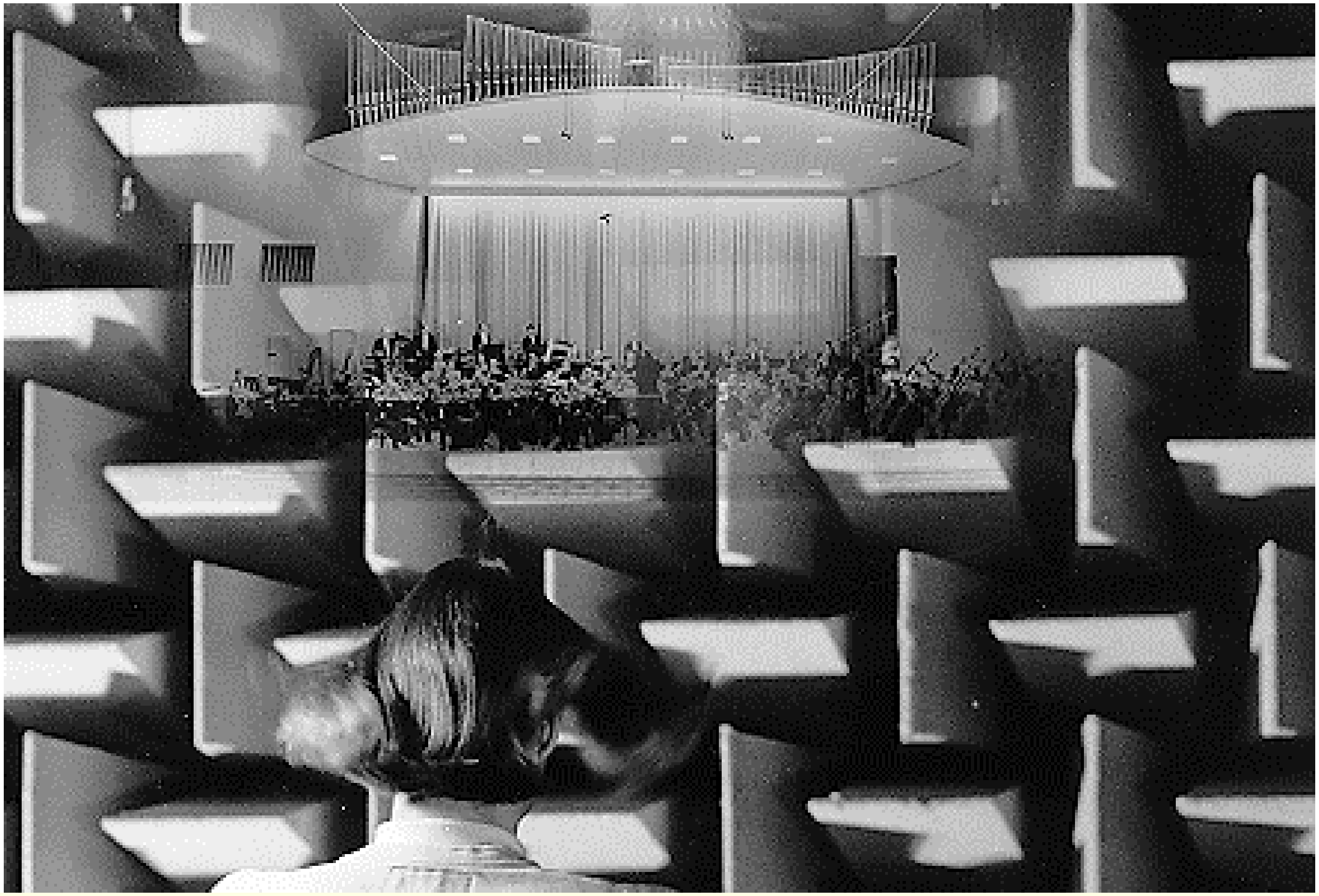
- Reverberance, echoes, surface characteristics reveal dimension, extent of space via cognitive associations
- Interactive cues: EC's effect on speech effort, intelligibility
- Perceived sound source width and distance affected by EC

("concert hall" musical acoustic factors - intimacy, envelopment)



Spatial (and non-spatial) hearing simultaneously reveals information about the environmental context (EC) of a sound source.

- Reverberance, echoes, surface characteristics reveal dimension, extent of space via cognitive associations
- Interactive cues: EC's effect on speech effort, intelligibility
- Perceived sound source width and distance affected by EC
- Simultaneous sources: ambient sound & background noises (inside & outside) contribute to EC perception



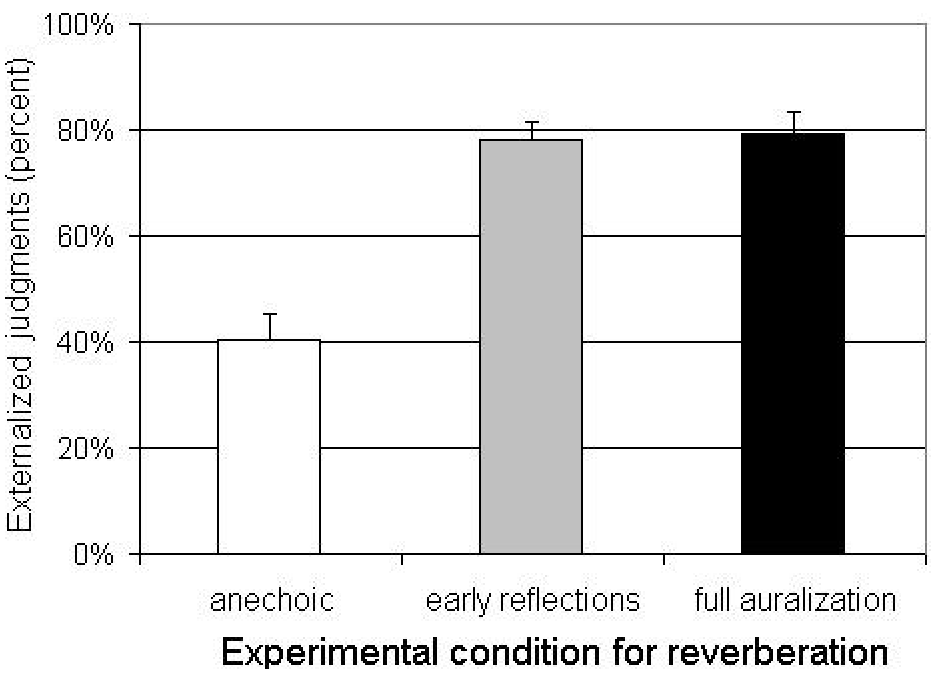
Auralization: virtual simulation of sound sources within an environmental context, using loudspeakers or headphones.

Useful applications for auralization technology include:

• Psychoacoustic investigations of spatial hearing in realistic acoustical environments (e.g., echo thresholds)

• Acoustical engineering (listening to a virtual room as a part of the design phase)

• Improved virtual environment simulation and 3-D sound reproduction →

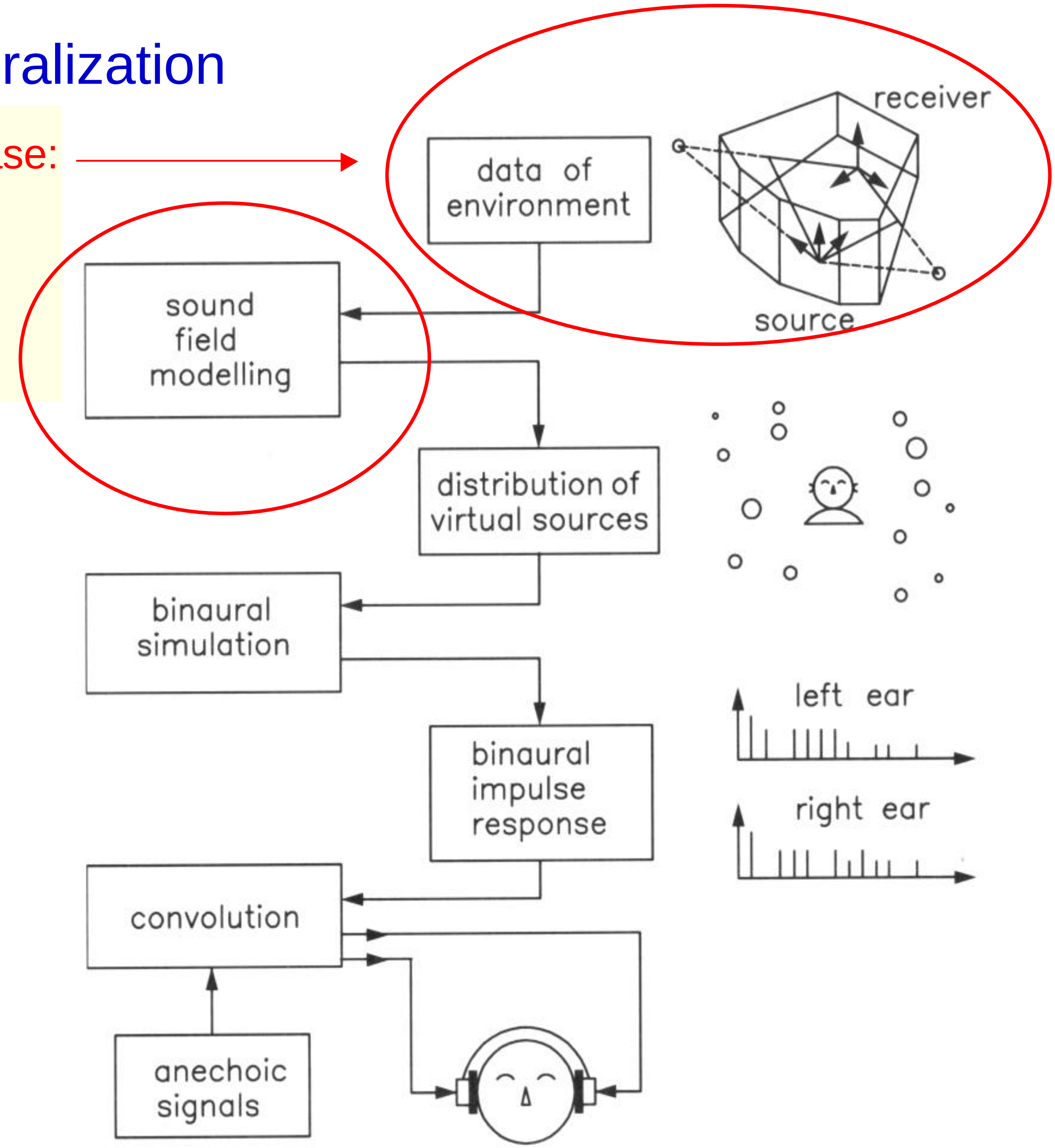


• Increased realism for entertainment (music reproduction)

• Evaluation of sound quality for industrial applications

Basic technique for auralization

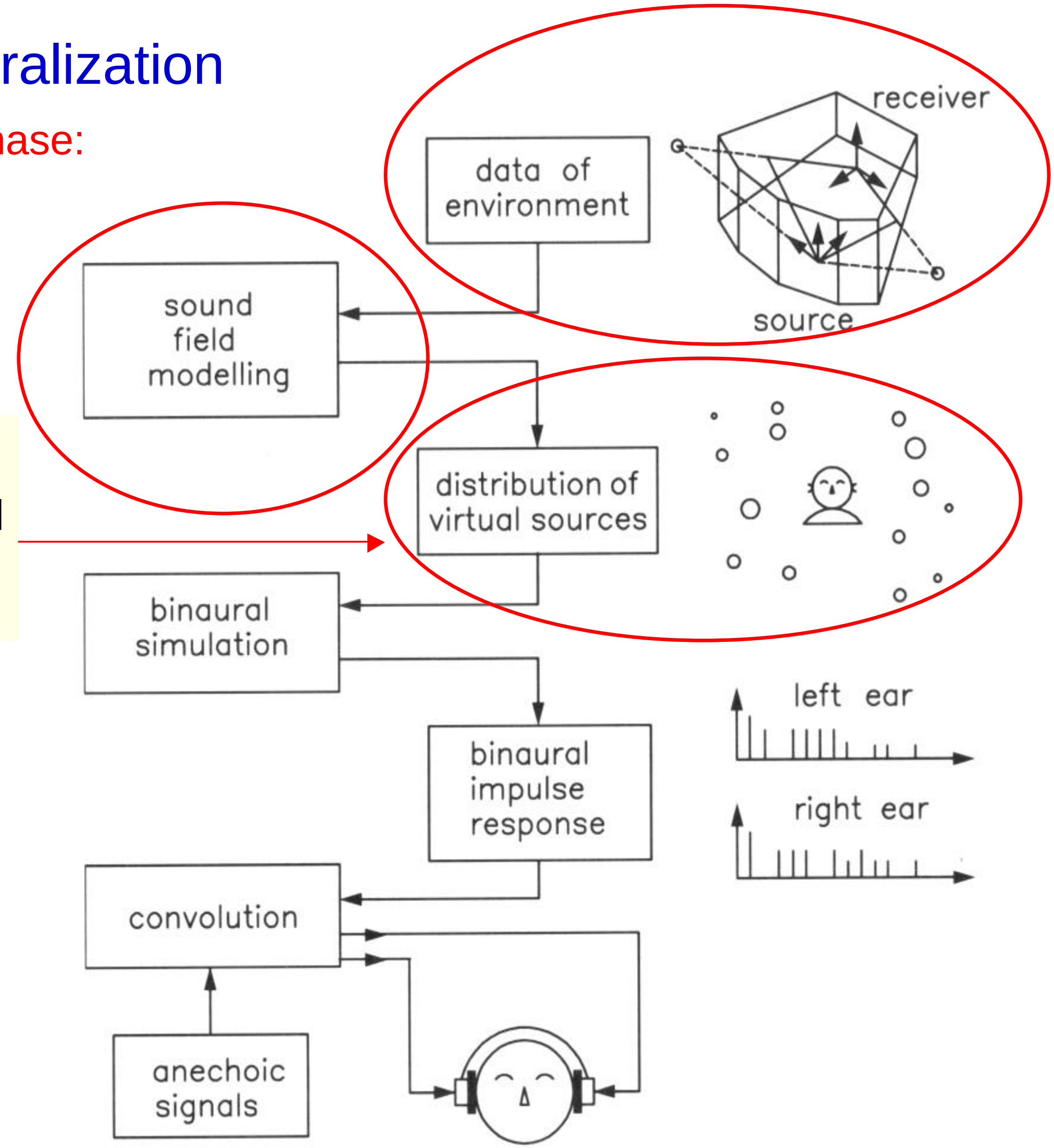
- Modeling (pre-synthesis) phase:
  either measure or predict
  room impulse response
  (data reduction is usually
  necessary)



Basic technique for auralization

- Modeling (pre-synthesis) phase:
  either measure or predict
  room impulse response
  (data reduction is usually necessary)

- Translate impulse response to listener-referenced virtual source locations

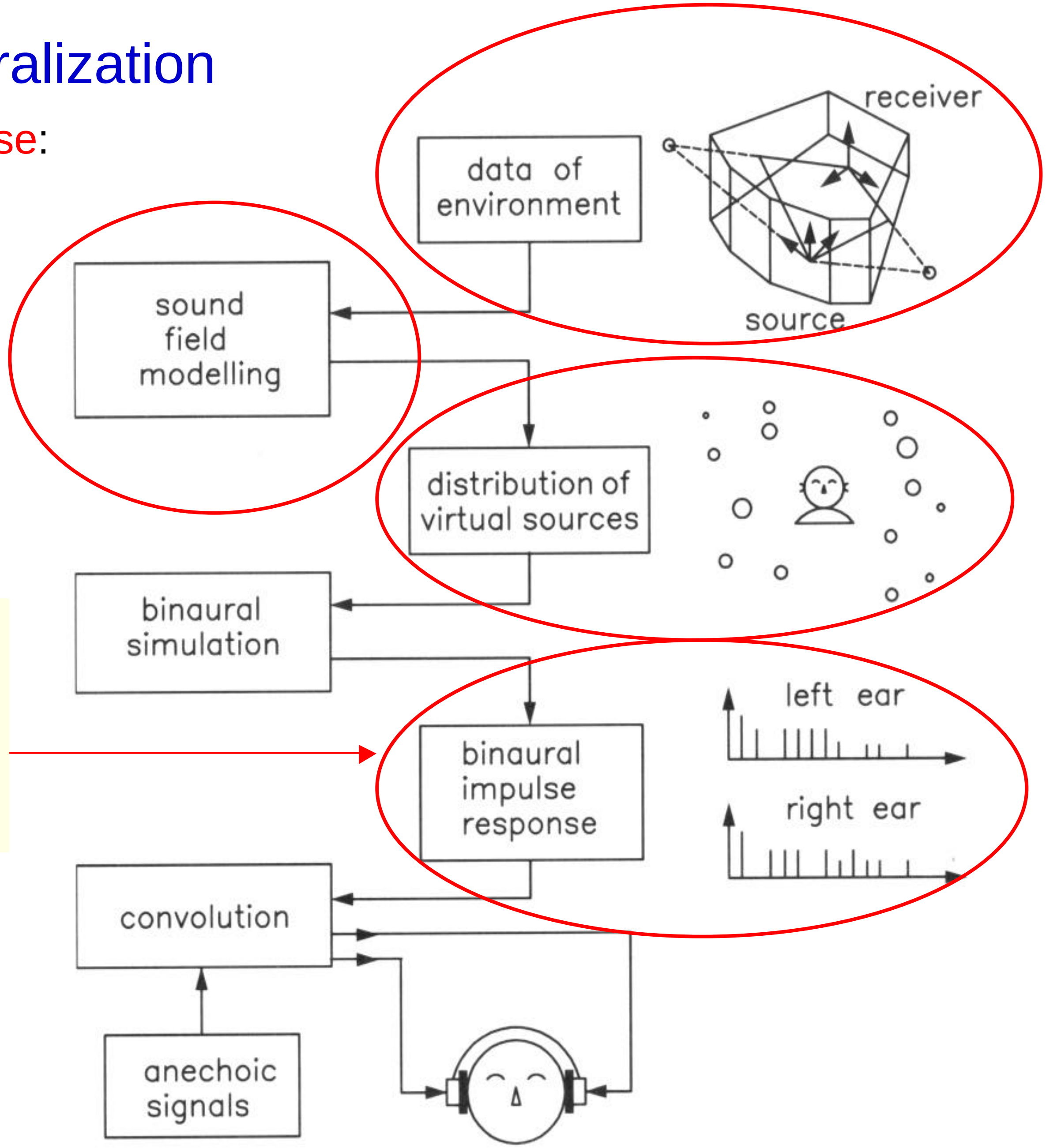


Basic technique for auralization

- Modeling/pre-synthesis phase:
  either measure or predict room impulse response
  (data reduction is usually necessary)

- Translate impulse response to listener-referenced virtual source locations

- Apply virtual acoustic techniques (HRTF filtering) to derive the binaural impulse response



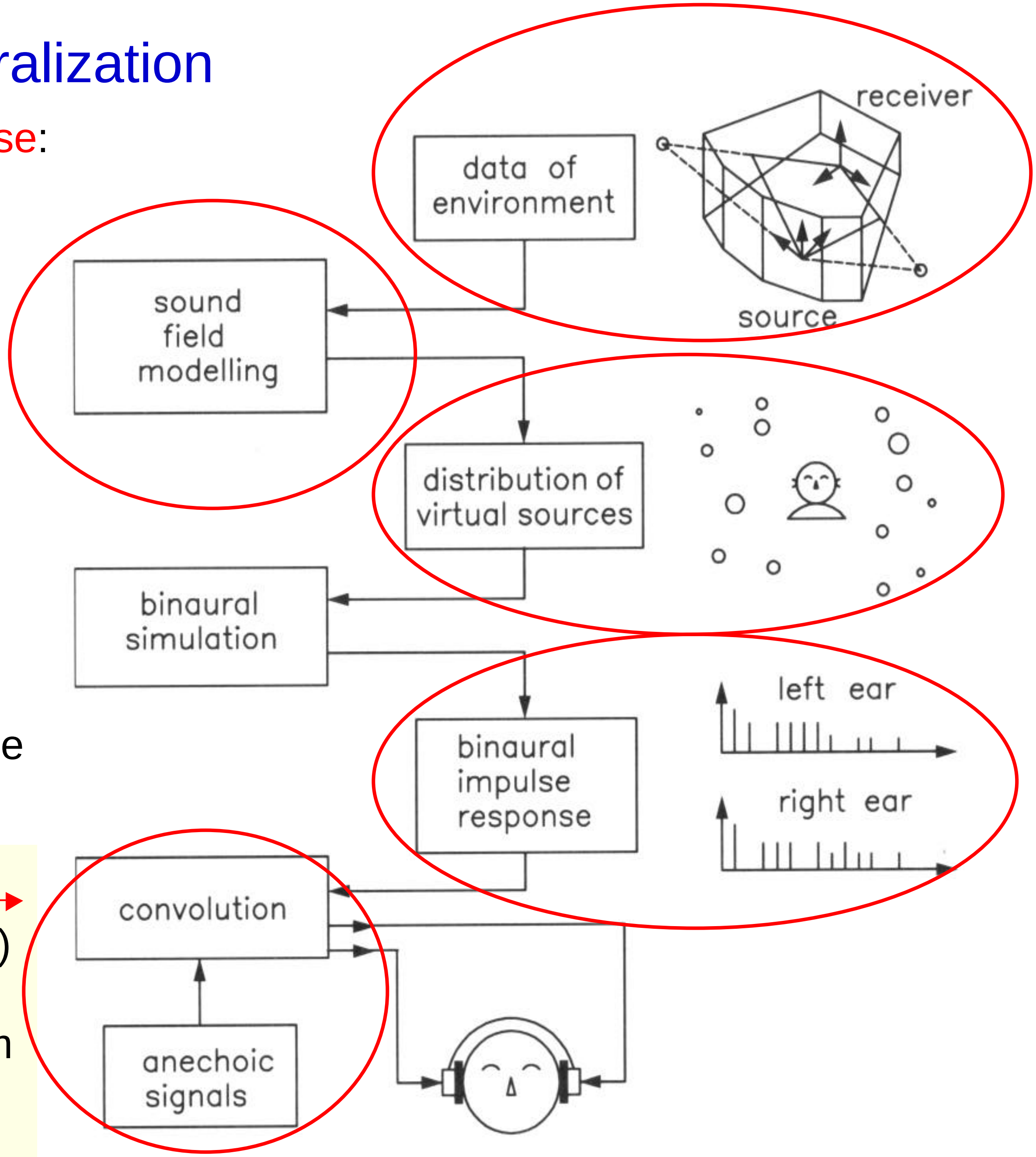
Basic technique for auralization

- Modeling/pre-synthesis phase: either measure or predict room impulse response *(data reduction is usually necessary)*

- Translate impulse response to listener-referenced virtual source locations

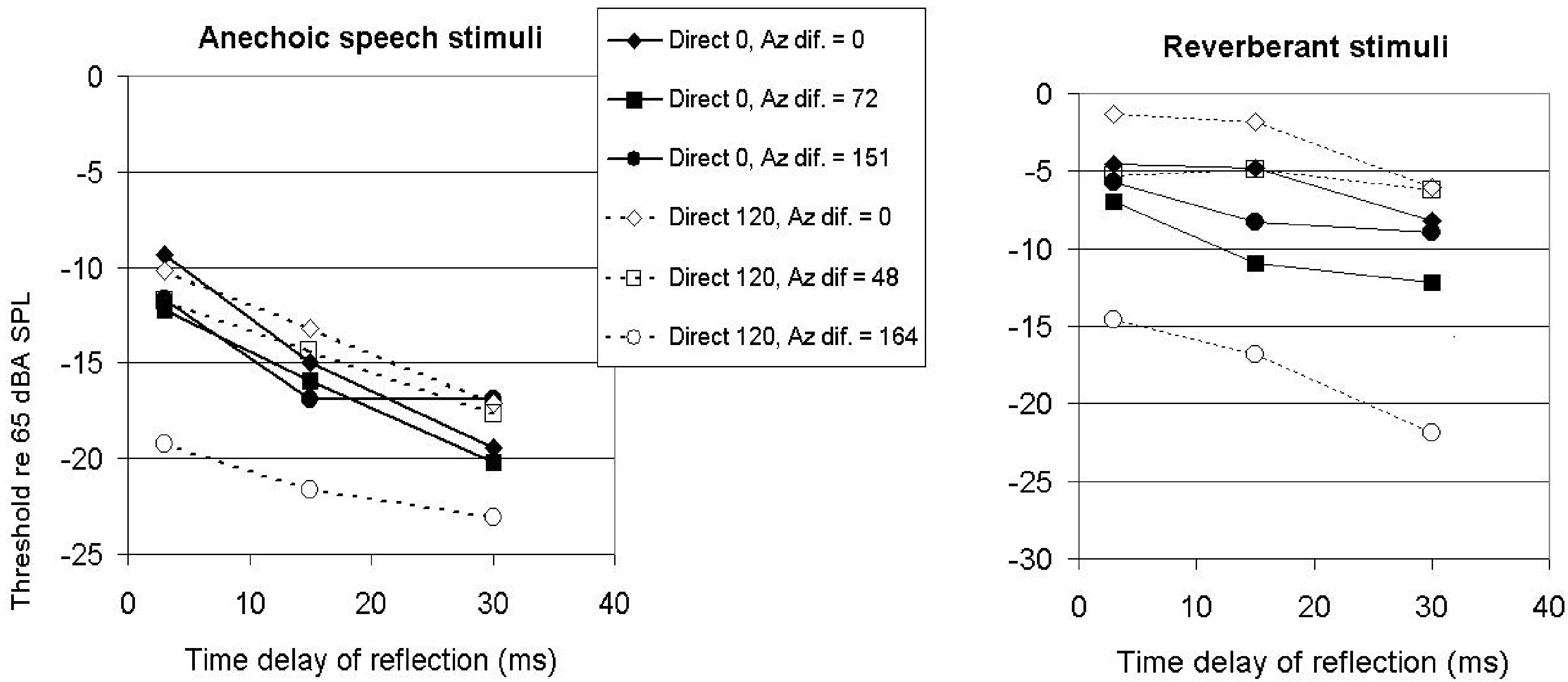
- Apply virtual acoustic techniques (HRTF filtering) to derive the binaural impulse response

- Synthesis phase: Convolve “neutral” (anechoic) signals with the binaural impulse response of the room *(data reduction is usually necessary for rendering)*

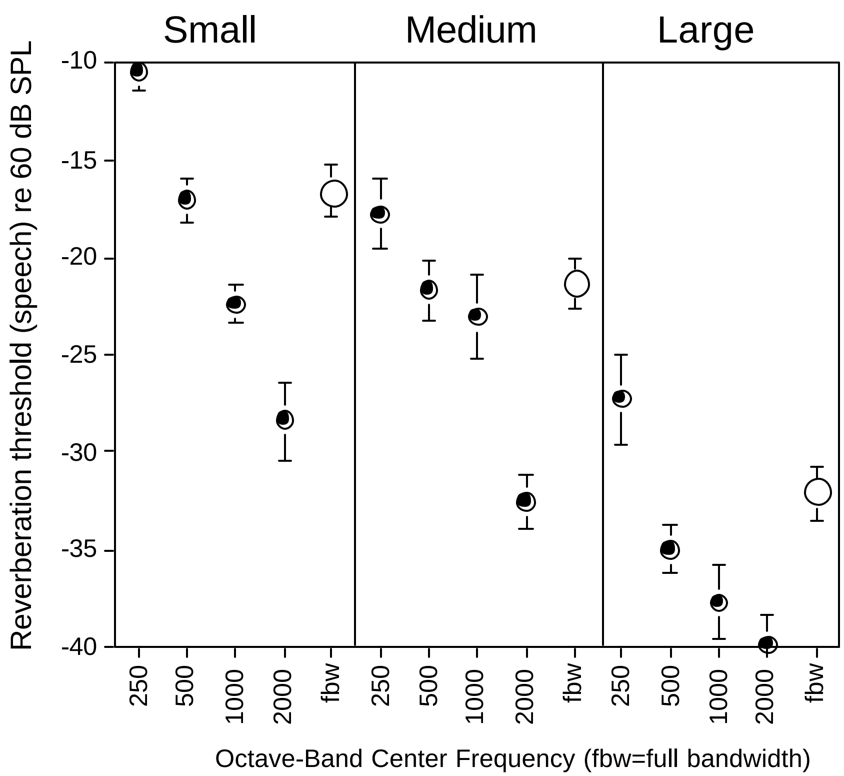


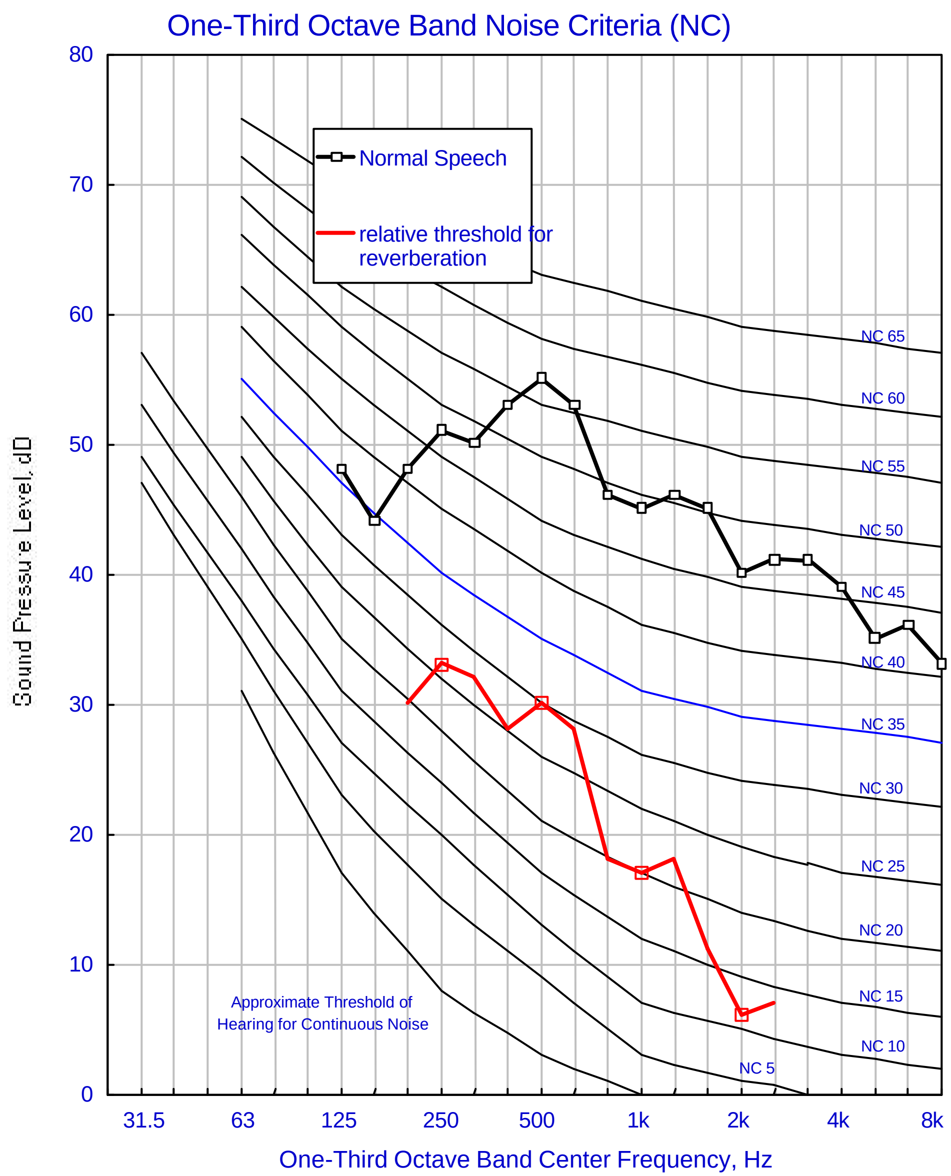
• Recent study indicates that echo thresholds in real and virtual environments are similar

• Engineering rule of thumb: early reflections re direct sound should be inaudible < -21 dB @ 3 ms and < -30 dB @ 15-30 ms (add 5 dB for speech stimuli).



Although thresholds for reverberation are relatively low, background noise (e.g., NC 35) can mask a significant portion of reverberant energy (particularly the reverberant decay).

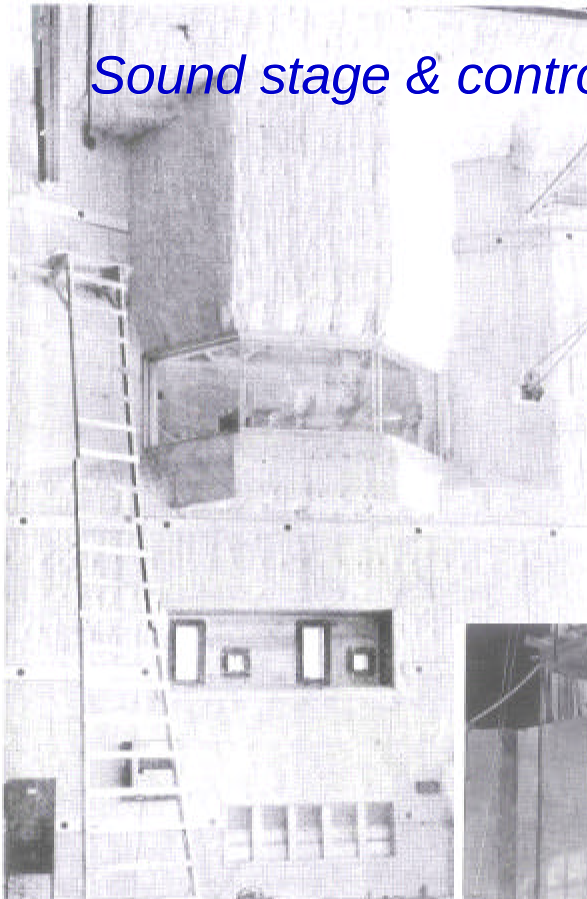




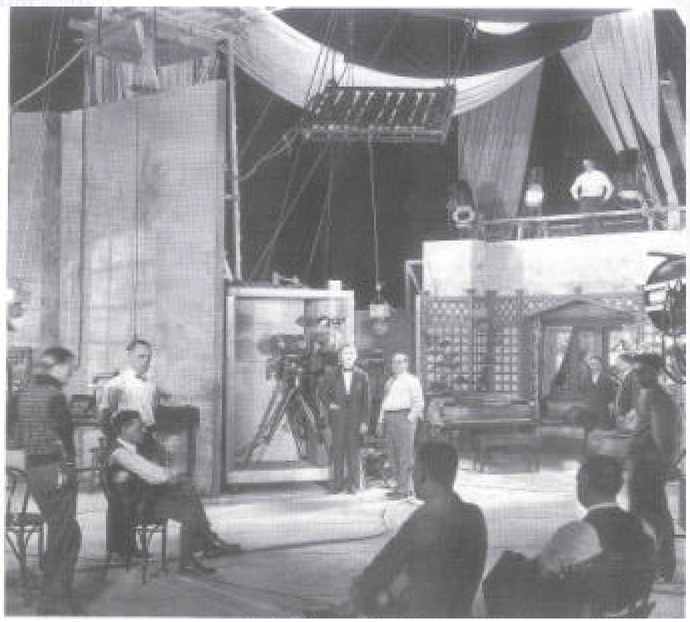
- Creating virtual spatial sound environments: playback

Historically, the use of acoustical absorption to ‘neutralize’ acoustical characteristics of rooms began ca. 1930s

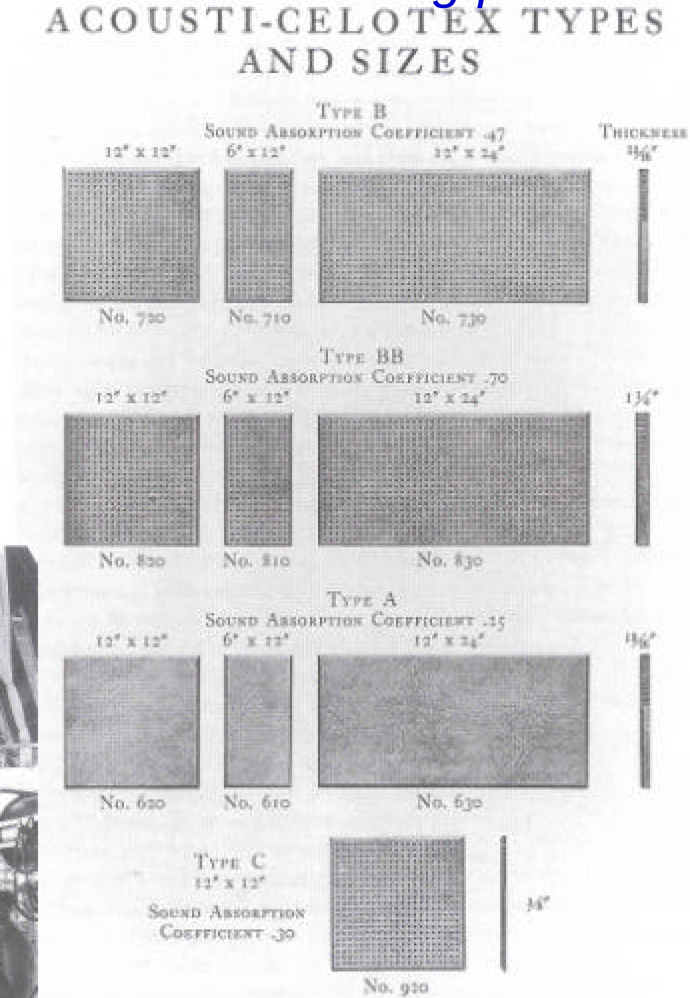
Sound stage & control room

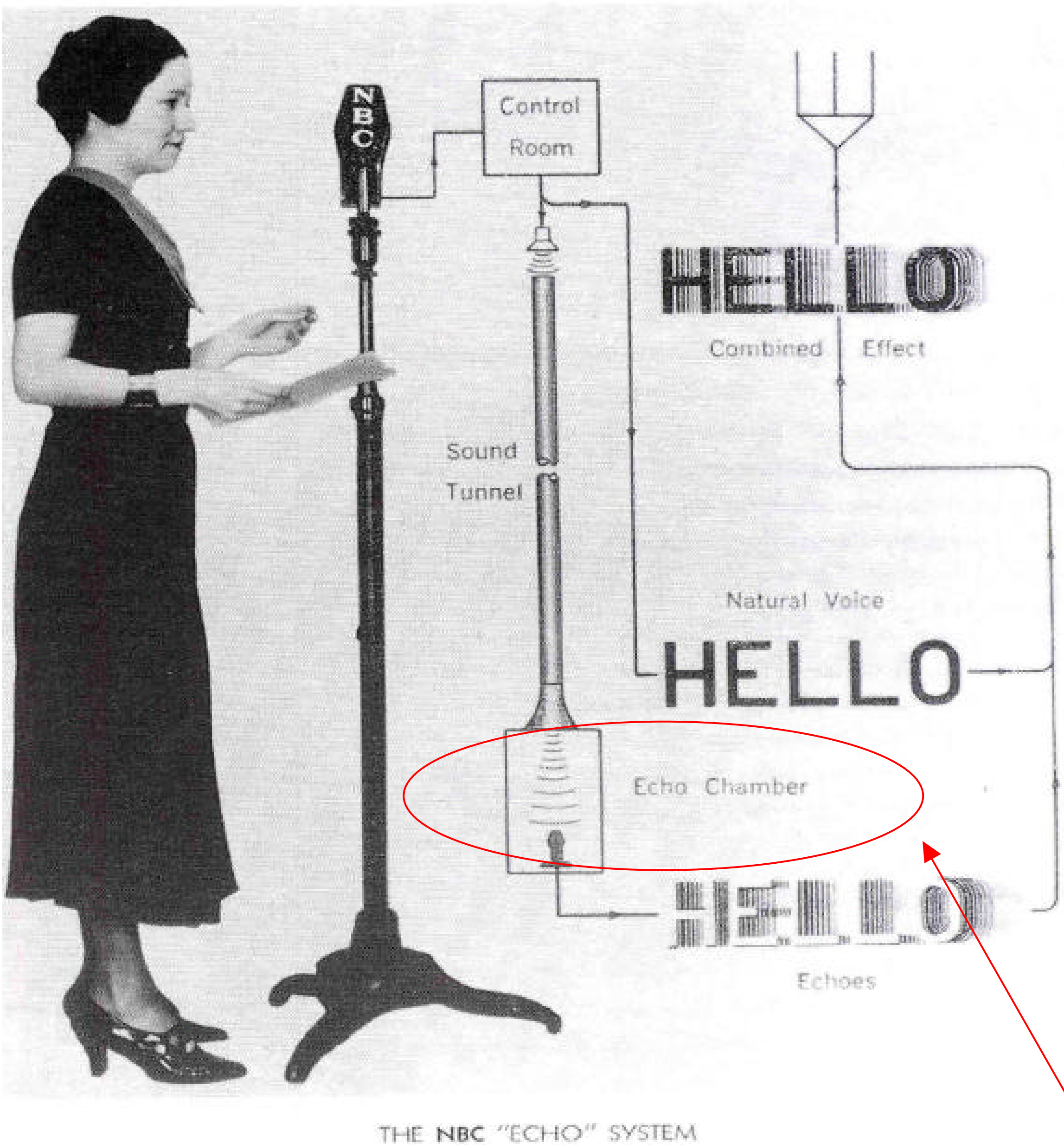


Recording for film



Acoustical wall/ceiling panels



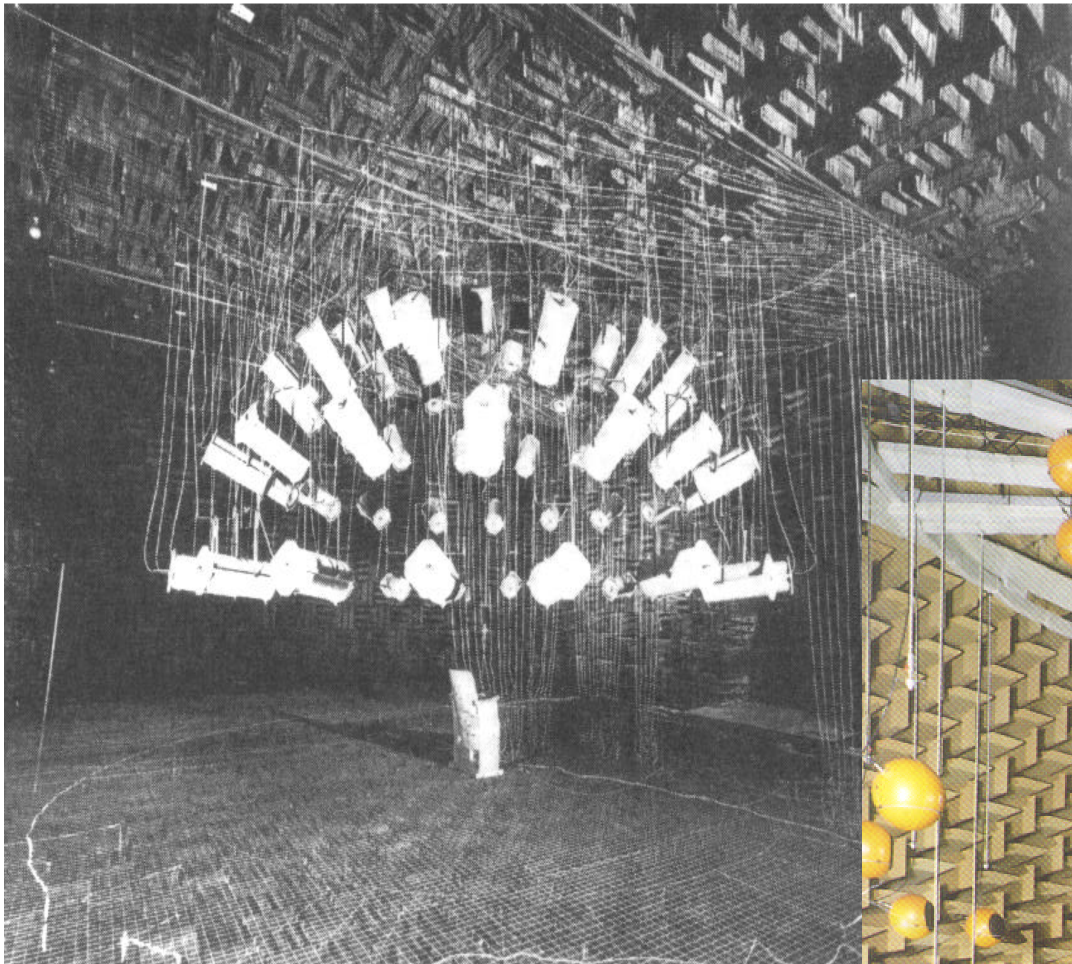


The realism of environmental context simulations has increased in tandem with the "neutralization" of rooms via absorption

- Echo chambers
- Reverberation plates
- Spring reverberators
- ↓
-
- Real-time convolution for head-tracked auralization

Reverberation using echo chamber, NBC, 1930s

Auditorium simulators using actual sound sources

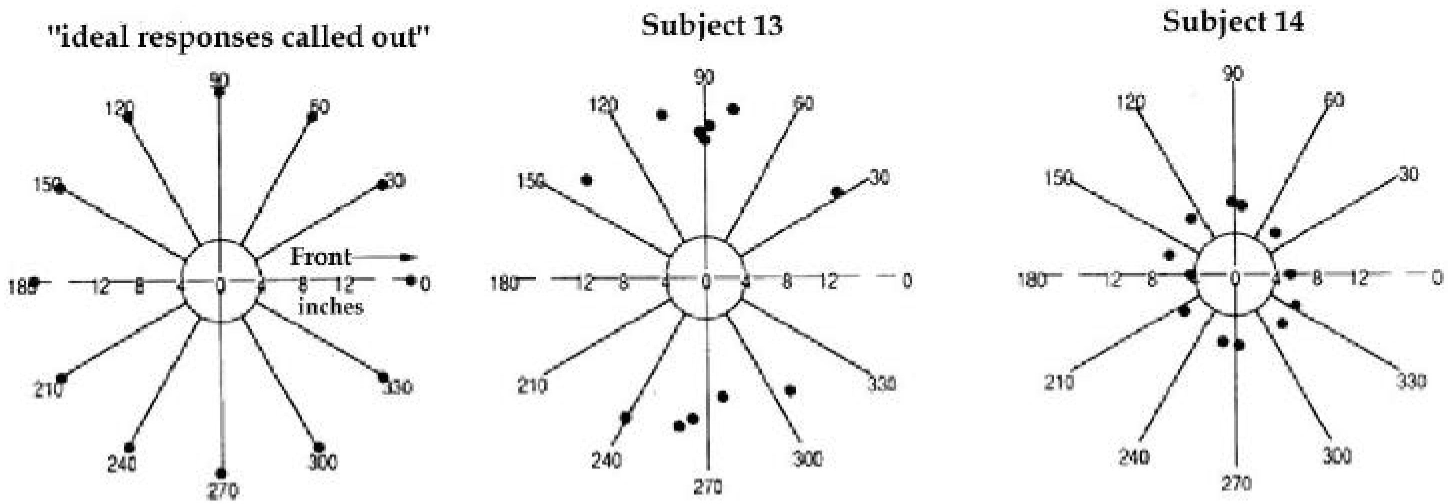


University of Goettigen 1965.

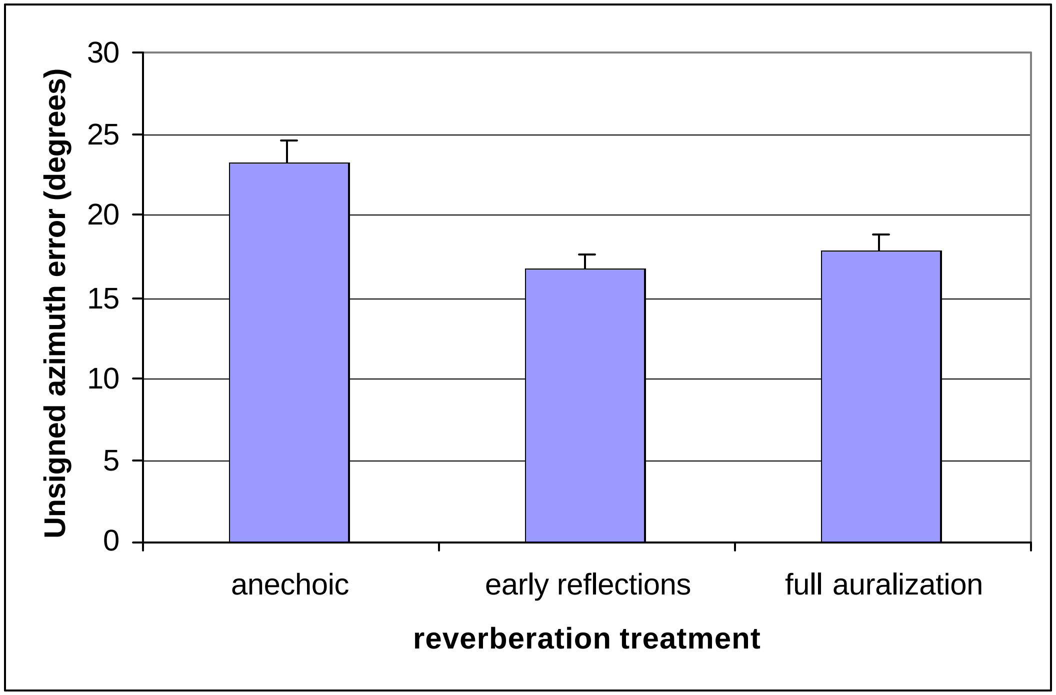
Technical University of Denmark, Lyngby, 1992.

Localization error for headphone stimuli (azimuth)

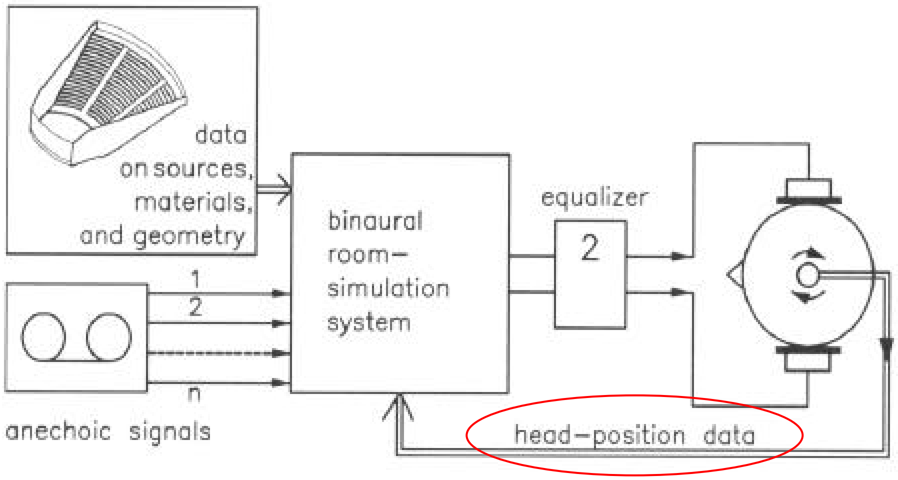
Anechoic Speech: Individual differences

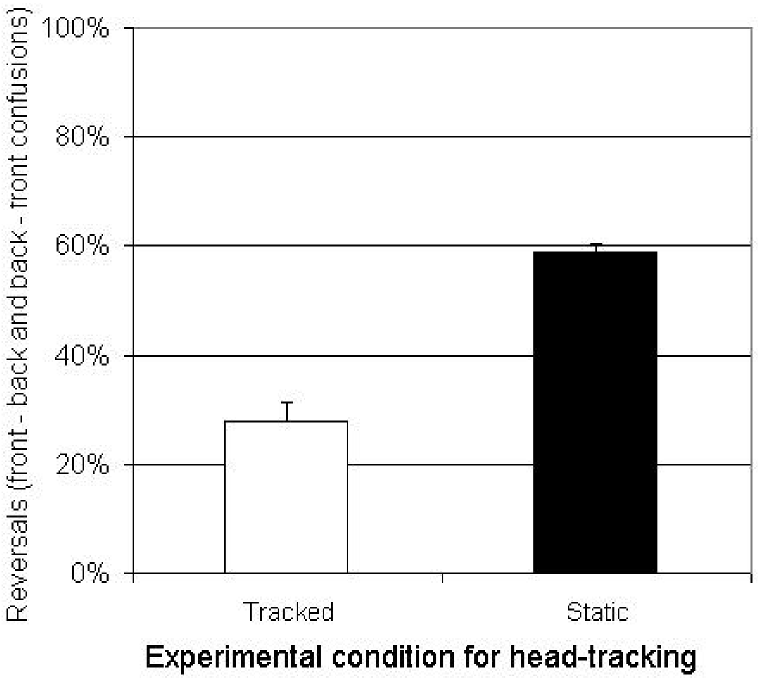


Mean values for different reverberation conditions

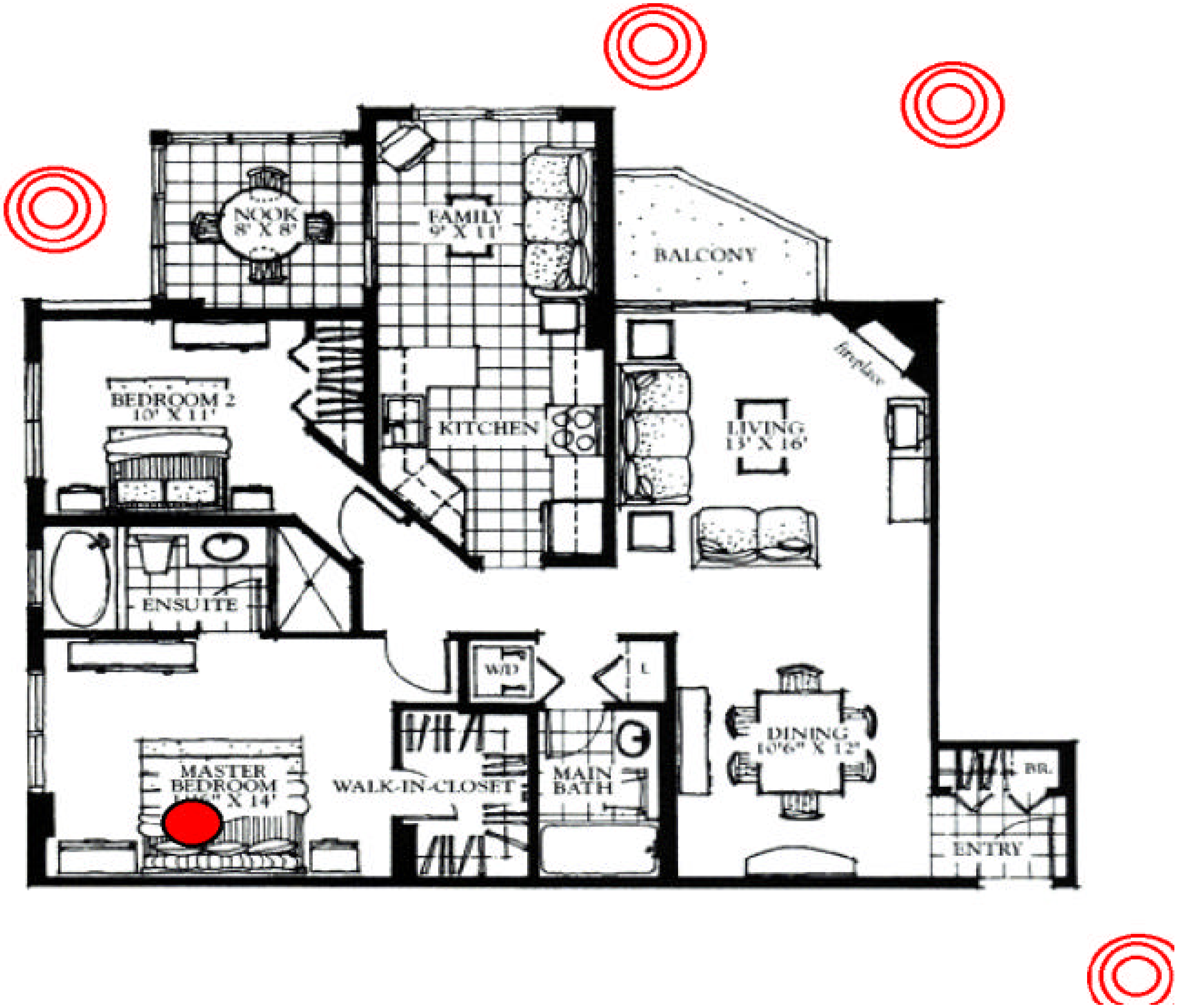


Head-tracked systems increase realism of 3D audio simulations in general (minimization of reversals)





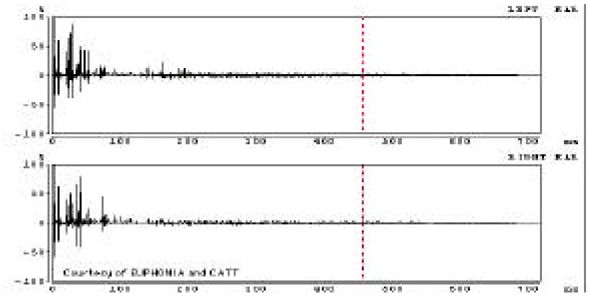
Auditory localization can be influenced or biased by cognitive references and visual capture

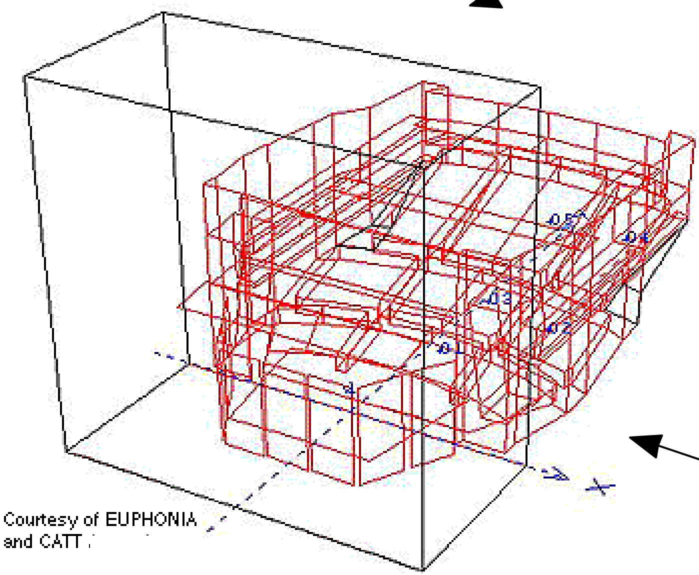


- Creating virtual spatial sound environments: measurement

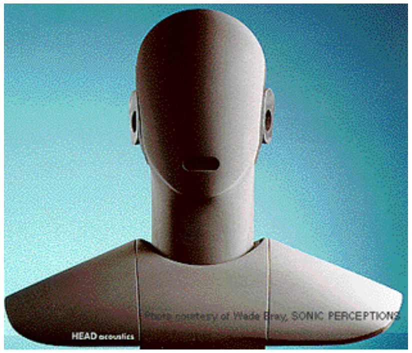
Spatial room impulse responses can be obtained from a existing room by…

1. Calculating the response from a room model (ray tracing, image modeling)



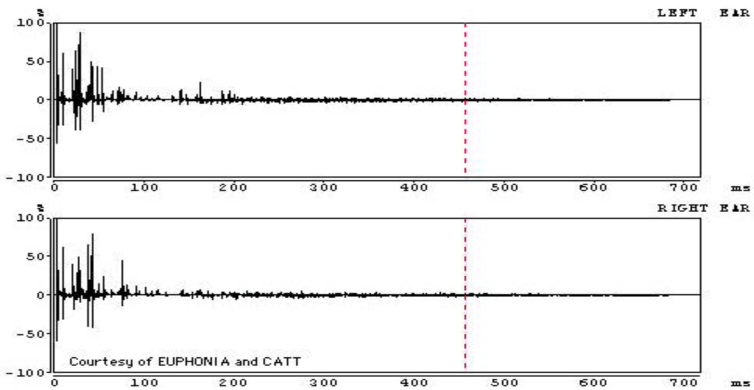


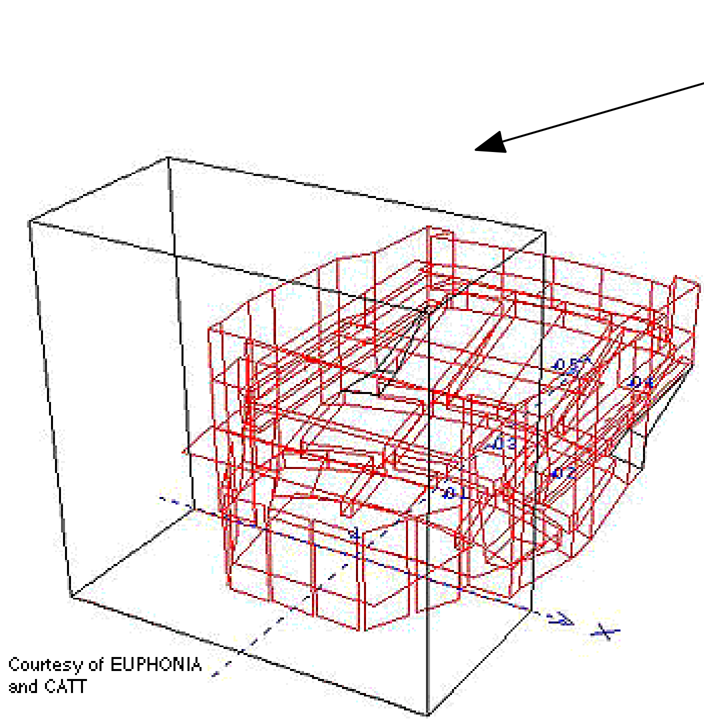
2. Recording the binaural impulse response



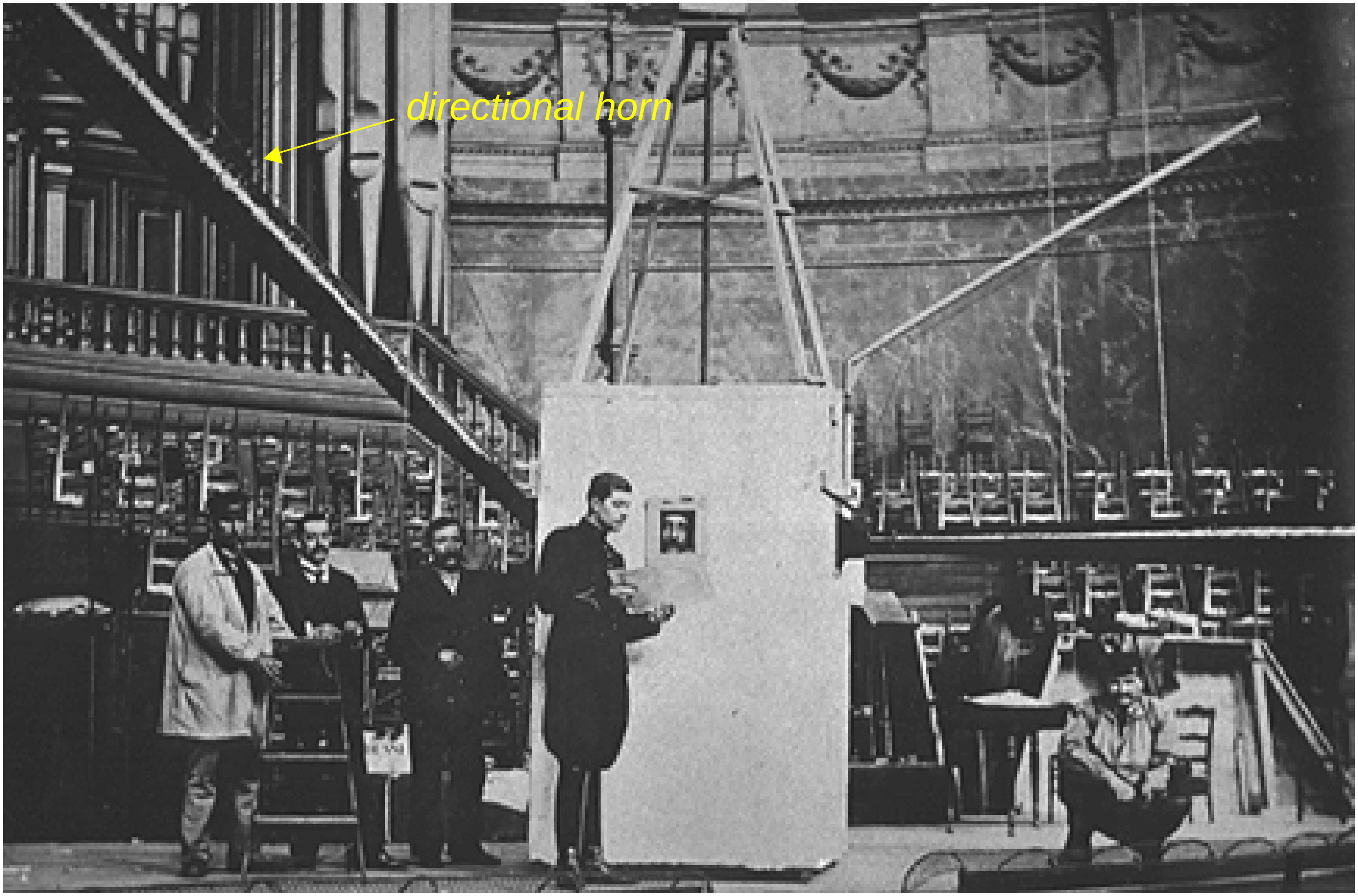
3. Recording a directional impulse response for post-processing of directional information

Calculate the impulse response
from a room model
(using ray tracing, image modeling)



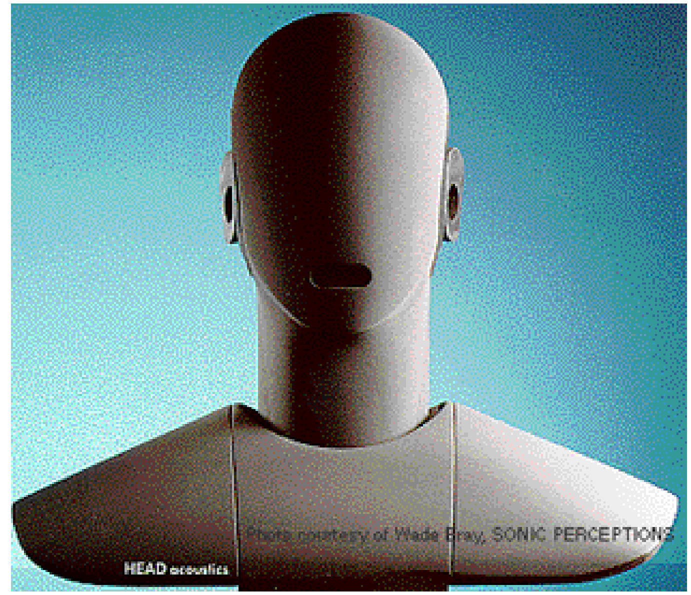


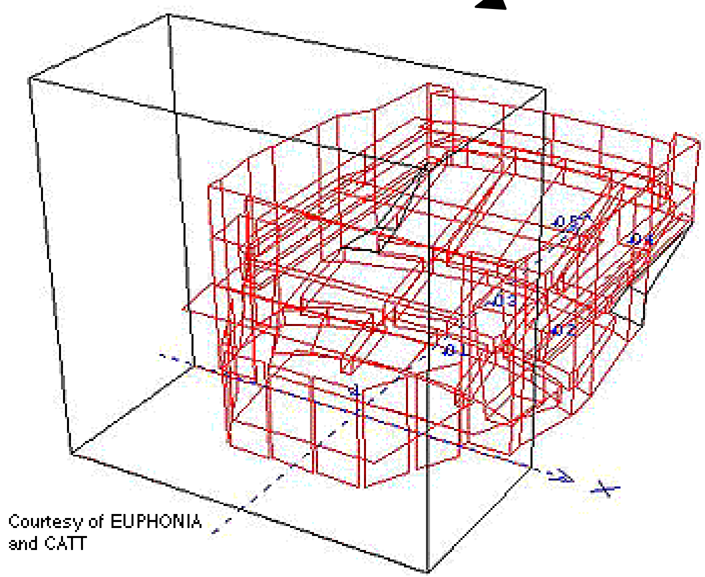
Challenges:
-adequate yet reasonable
sampling of sources & receivers
-directivity data for sound sources
limited; sound source movement
usually unaccounted for
-modeling low frequency behavior



Ray tracing, while imperfect, can be an efficient means for finding the location of disturbing echoes

Measure the binaural
room impulse response
via a dummy head recording





Challenges:
-localization error due to non-individualized HRTF
-deriving spatial information from a 2-channel source
-multiple measurements for each receiver position required

Directional mic room impulse response.

Essert (1996) used a B-format output from a *SoundField* MKV microphone to obtain one omnidirectional (W) and three dipole IRs (oriented left-right X, back-front Y, and down-up Z, respectively)

The omnidirectional response reveals the arrival time of significant early reflections;

Cross-correlations between the monopole and dipole responses indicate reflection direction of arrival.

Individualized HRTFs can be applied during the synthesis phase to significant early reflections

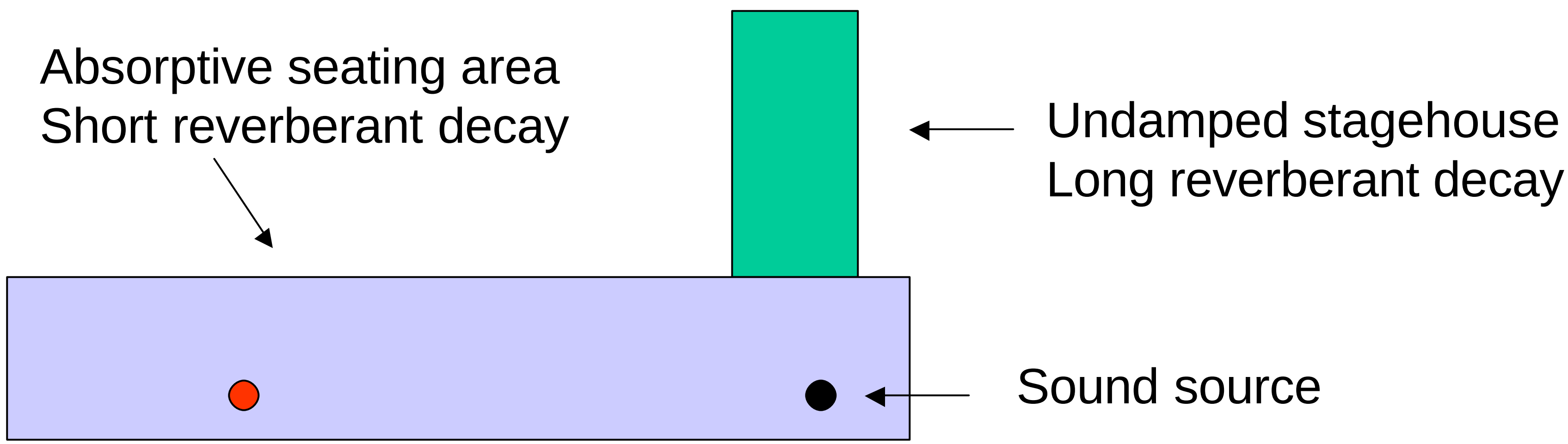
Intensity measurements have also been investigated in the literature.

- Simulation of coupled spaces, low frequency energy

Reverberation can cause spatial movement of reverberant sound within coupled spaces

-examples: theater; organ music in a gothic cathedral

-difficult to model



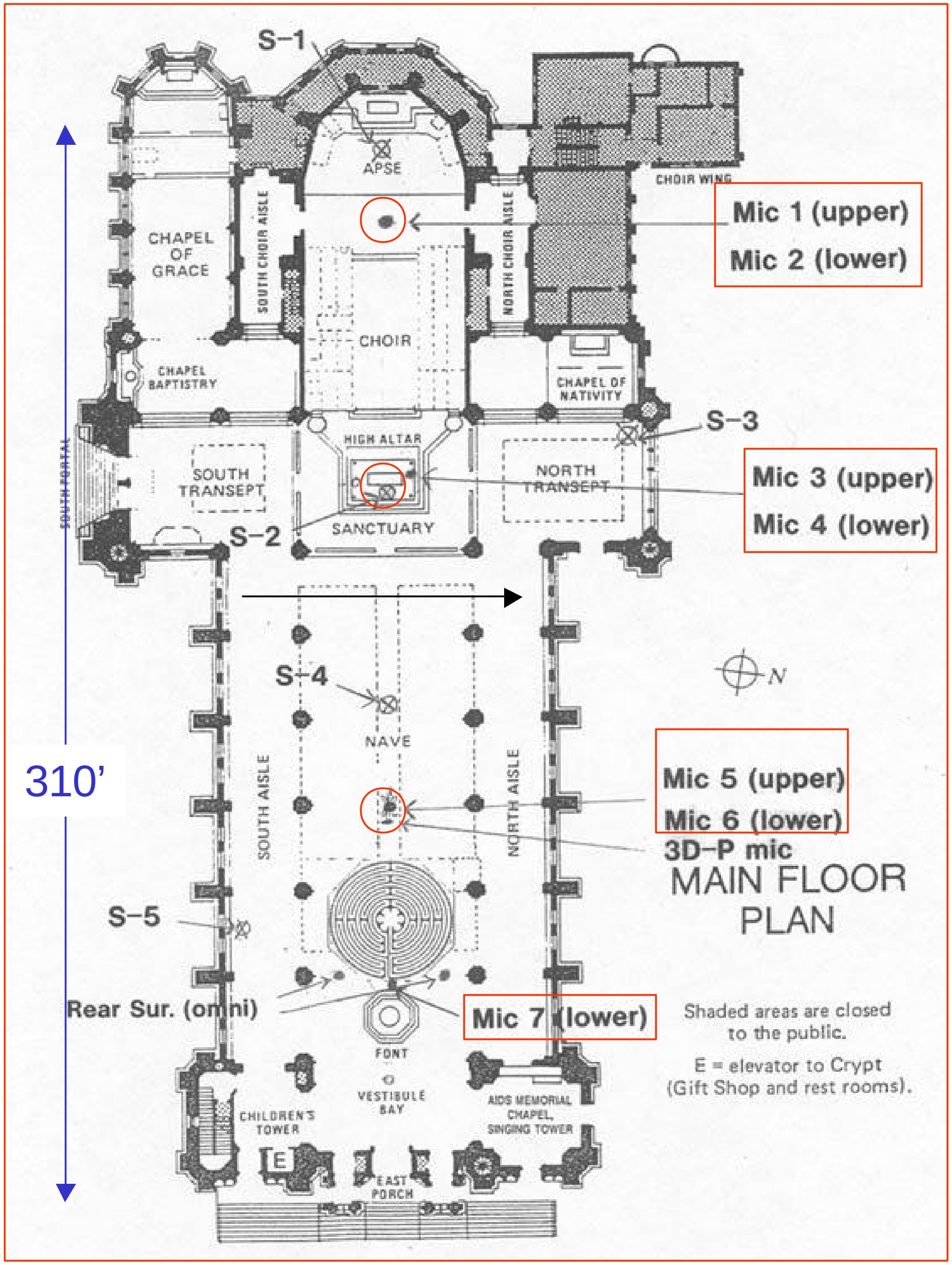
Measurement of “moving reverberation”:

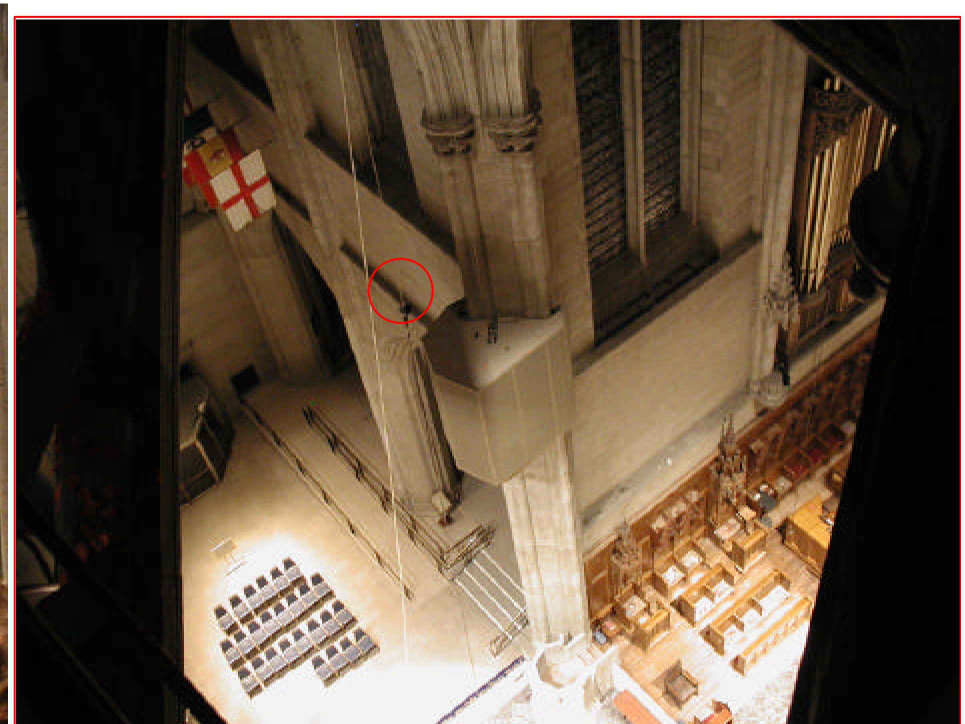
Grace Cathedral, San Francisco

-starter pistol sound source at 5 locations

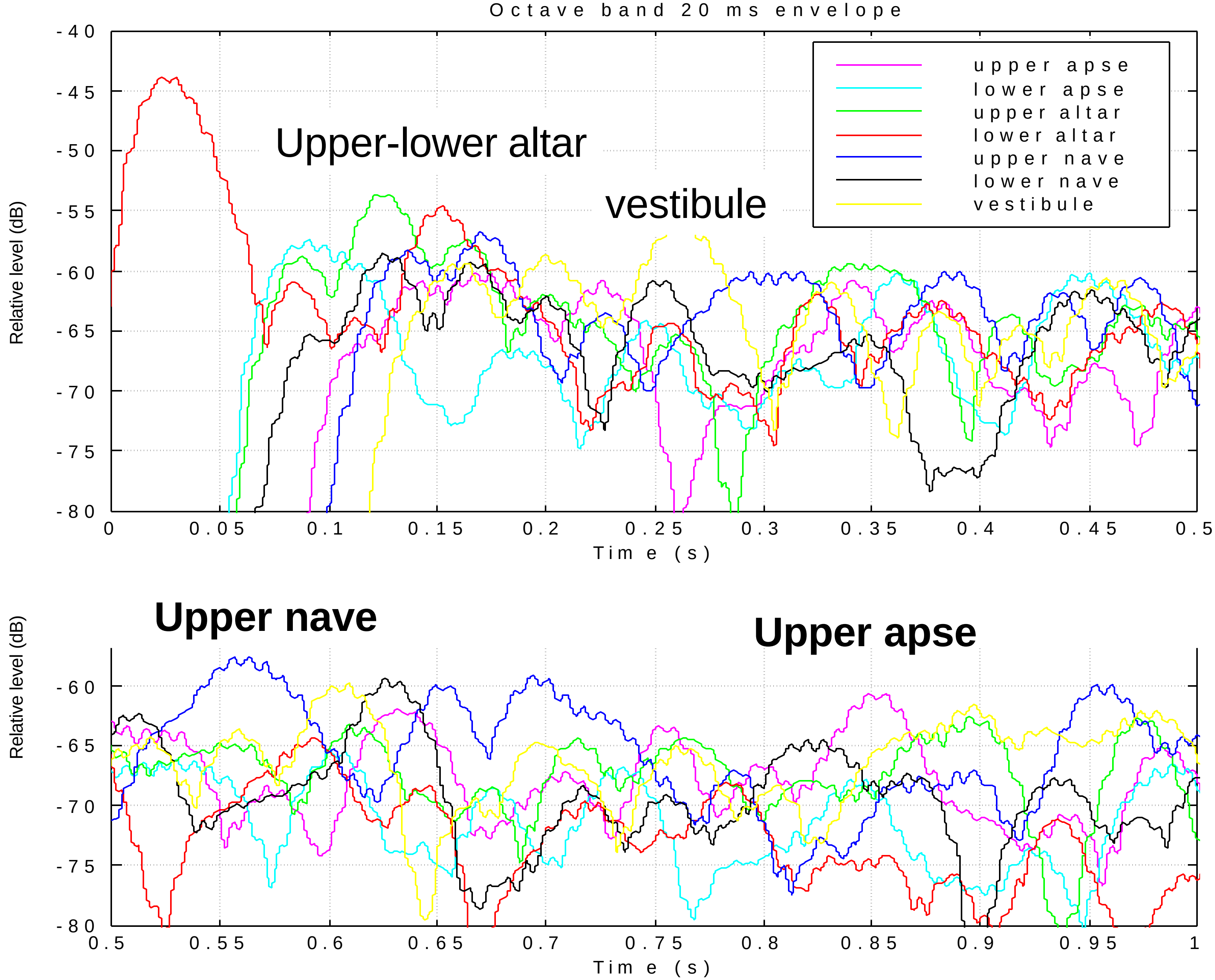
-7 measurement microphones simultaneously recorded with synch pulse

-Superior to binaural measurement since location/time can be tracked



Sound source at altar: 20 ms envelope



Overlap of frequency sensitivities for haptic and acoustic stimuli

"Haptic" includes both kinesthetic (muscle) and touch (cutaneous) sensations.

- Touch: 0 – 1 kHz
- Kinesthesia: 0 – 30 Hz
- Audition: .02- 20 kHz: overlap from .02 –1 kHz

Very low-frequency vibration can also be sensed via vestibular, visceral, and visual sensory systems

Transmission paths from a "felt-heard" source to a receiver

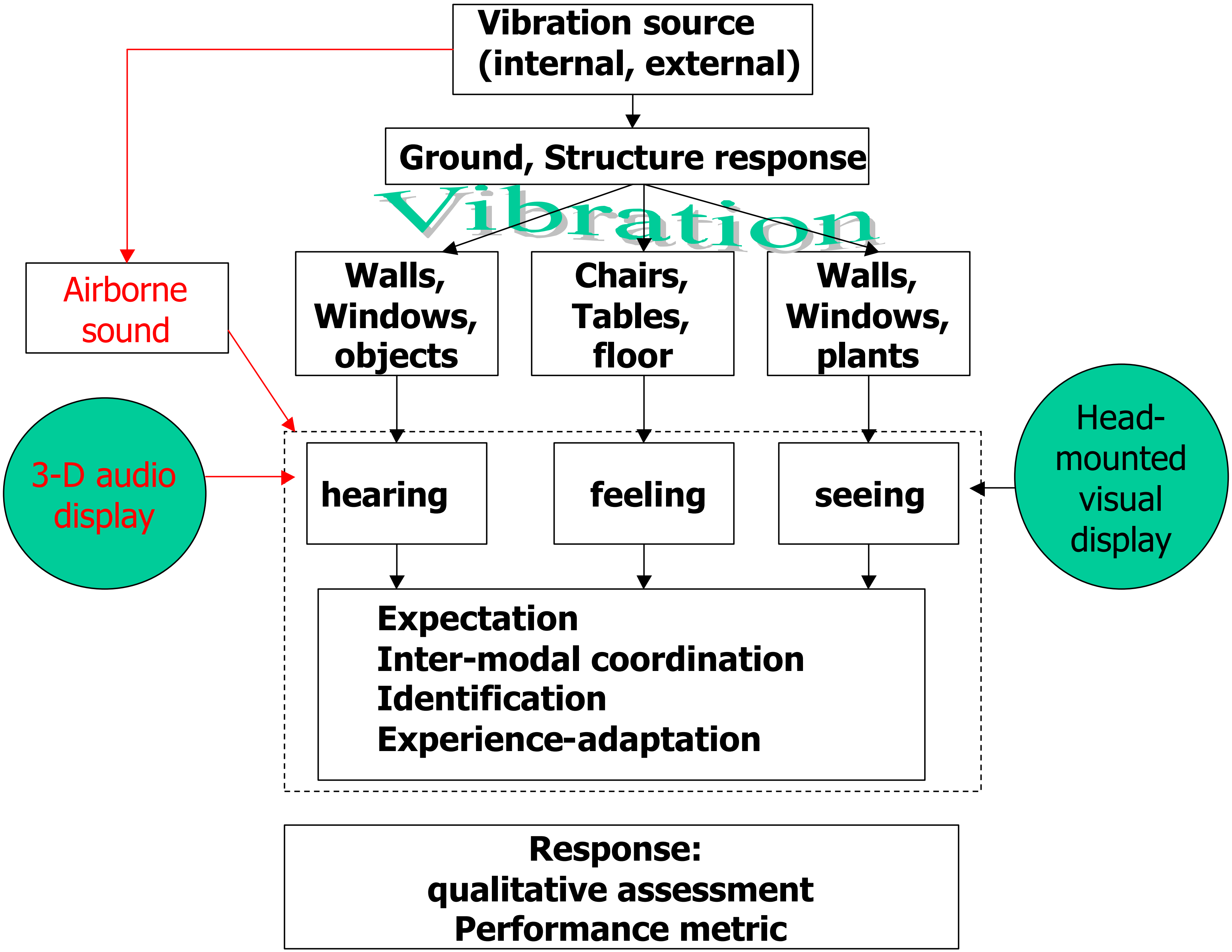
-Airborne pressure waves (sonic & infrasonic)

- -transfer of acoustic wave to mechanical vibration possible
- -simultaneous auditory-haptic stimuli possible
  - (touch; bone conduction; visceral)

-Ground / structure-borne vibration

- -"Full-body" vibration
  - from 8–80 Hz (displacement threshold 1.5 x 10^{-6} m)
- -Haptic sensation via kinesthetic, touch sensors
- -Transfer of vibration into acoustic excitation

-Perception of haptic and acoustical stimuli influenced by presence of coordinated visual stimuli



Vibration source
(internal, external)
Ground, Structure response
Vibration
Airborne
sound
Walls,
Windows,
objects
Chairs,
Tables,
floor
Walls,
Windows,
plants
3-D audio
display
hearing
feeling
seeing
Head-
mounted
visual
display
Expectation
Inter-modal coordination
Identification
Experience-adaptation
Response:
qualitative assessment
Performance metric

Why simulate- and manipulate- both haptic and acoustical phenomena simultaneously?

-Improved "realism" ,"immersion" in virtual enviroments
(experiments; entertainment; training)

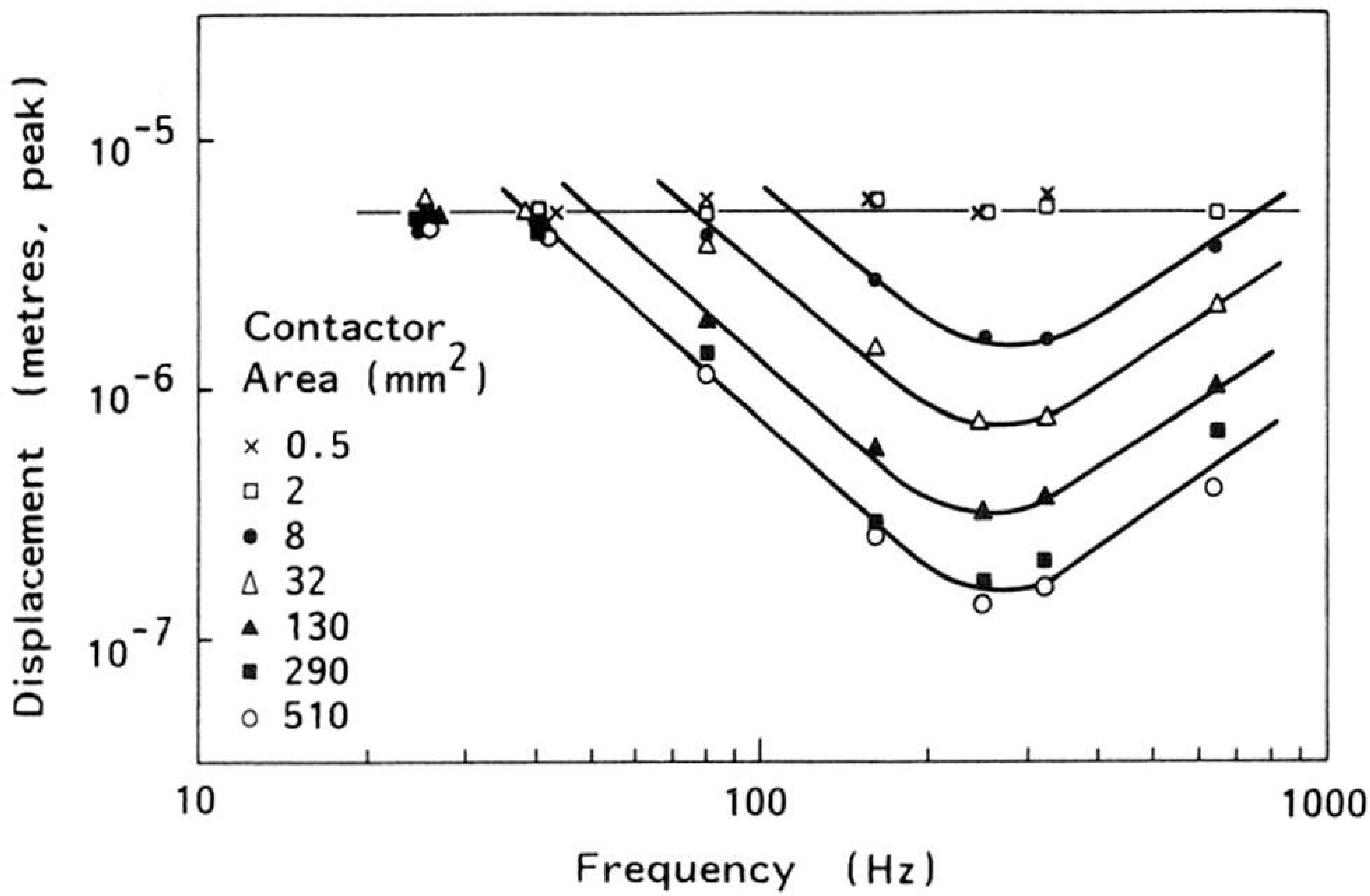
Simulations of musical performance (as player or listener) in particular seem incomplete without tactile cues

-Product quality evaluation (e.g., automotive industry)

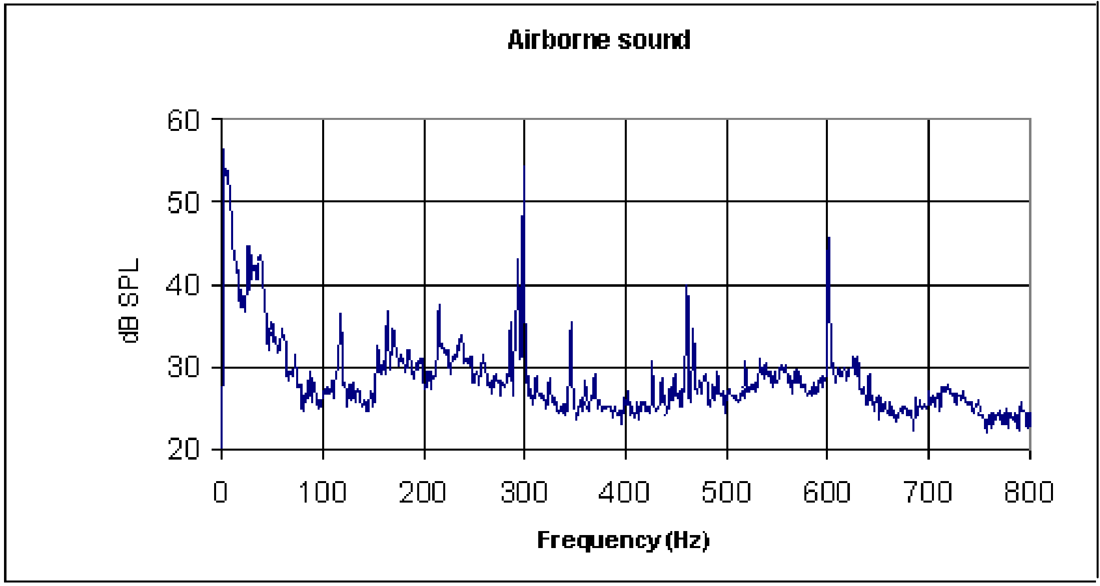
-Performance enhancement for tasks in VE
(e.g., transportation simulators)

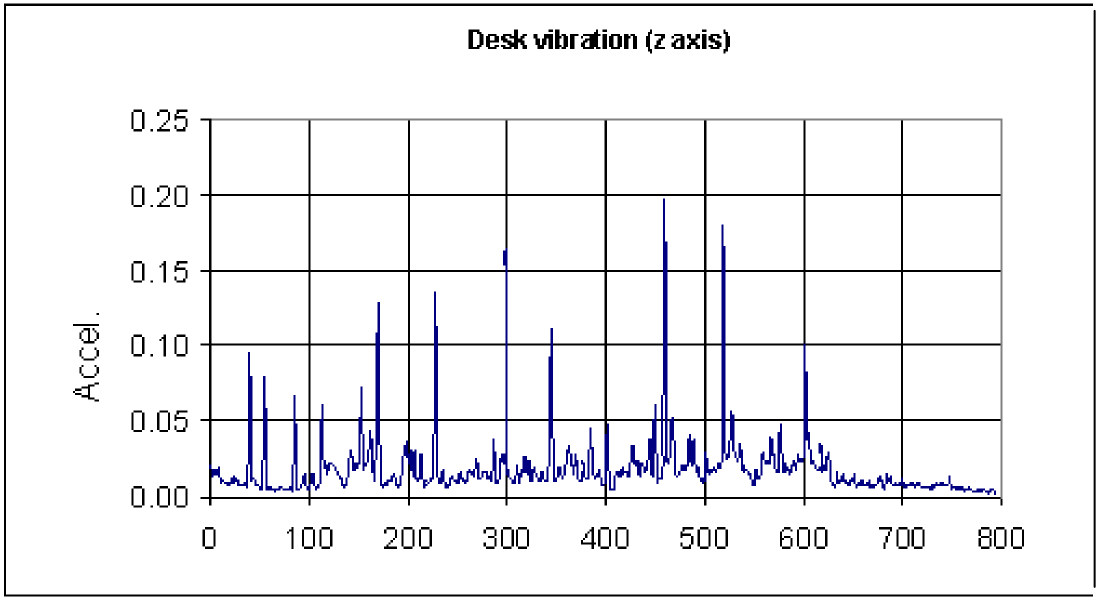
Thresholds for vibrotactile perception, ball of thumb

-peak at 250 Hz corresponds to acceleration of 0.24 m/s^2



Workplace sound and vibration at a desk





Summary

• Challenges for improved simulation include determining acceptable data reduction in measurement and synthesis (*echo thresholds)*

• Data reduction based on echo thresholds and reverberant masking can relax demands on auralization synthesis computation *(important for real-time rendering that includes head tracking).*

•Simulation of low frequency effects, modal prediction, vibration, and source directivity require further research efforts