Listening with Headphones

Main Types of Errors

- Front-back reversals
- Angle error

Some Experimental Results

- Most front-back errors are front-to-back
- Substantial individual differences
  - Most evident in elevation judgments
- Performance varies with region
  - Best localization: side / front?
  - Worst localization: top rear
- More front-back reversals with headphones than free-field

Headphone Issue: Externalization

"Lateralization" vs "Localization"

Affected by:

- Head movement
  - Need head tracking system and dynamic updating of cues
  - Head-tracked ITD and IID are sufficient for front/back differentiation!!!

Headphone Issue: Externalization

"Lateralization" vs "Localization"

Also affected by:

- Reverberation
  - Stereo reverberation, spatially diffuse
  - Interaural incoherence (more later)
- HRTFs
  - Needed especially if head is static

Headphone Issue: Externalization

"On the Externalization of sound images" Hartmann and Wittenberg (1996)

Externalization depends on:

- Interaural phase below 1K Hz
  - Constant ITD (not frequency-dependent differences)
- IID at all frequencies
- HRTFs at each ear (not interaural differences)

There is a continuous range of percepts from inside the head to outside

Review: Reality vs Ideal

From Begault and Wenzel, 1993

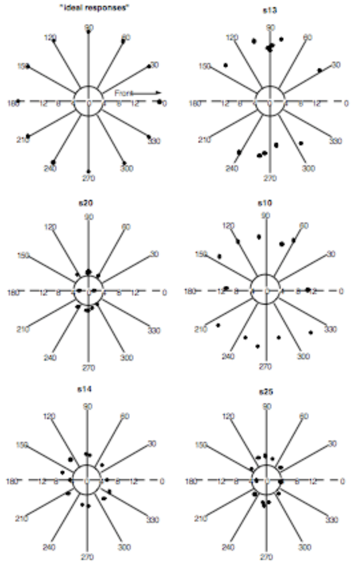
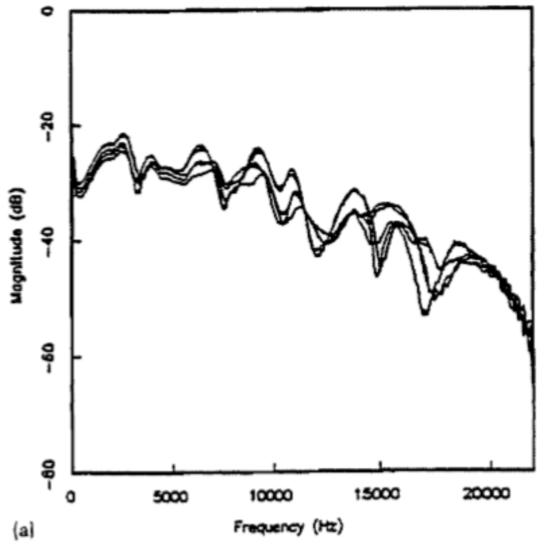
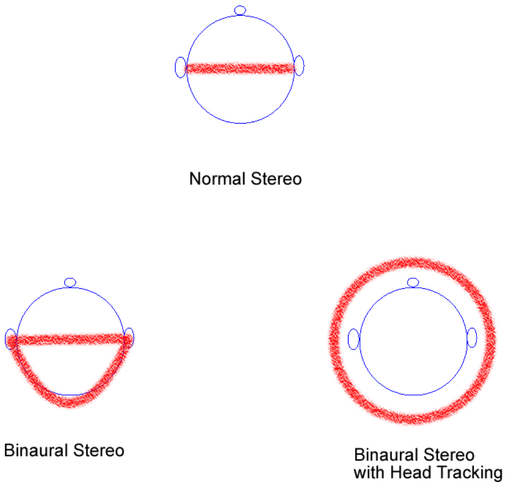


Figure 2.30. Azimuth and distance judgments for speech stimuli (Begault and Wenzel, 1993: see text). Upper left picture shows "ideal responses." *Reproduced with permission from* Human Factors, *Vol. 35, No. 2, 1993. Copyright 1993 by the Human Factors and Ergonomics Society. All rights reserved.*

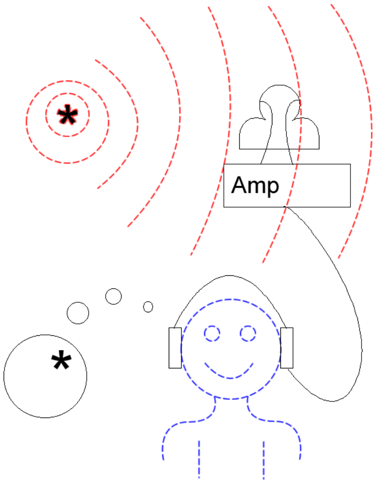
Headphone Issue: Repeatability with Reseating



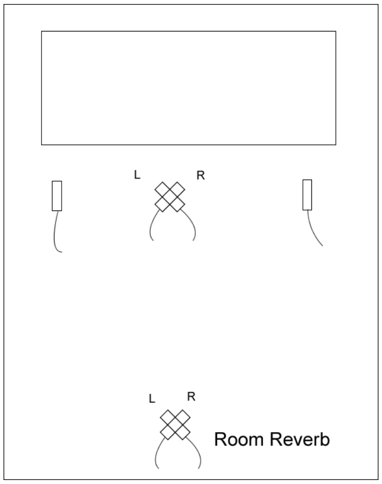
Headphone Reproduction: Comparison of Reproduction Modes



Binaural Recording



Classical Recording

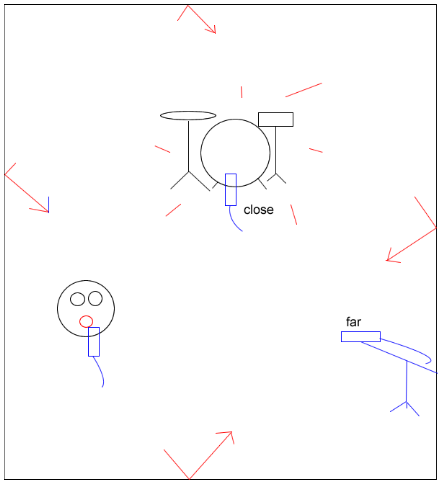


Aesthetic is to attempt to recreate the experience of being in a concert hall.

All sound sources are contained in a signal acoustic space.

The spatial layout attempts to replicate reality.

Rock/Pop Recording

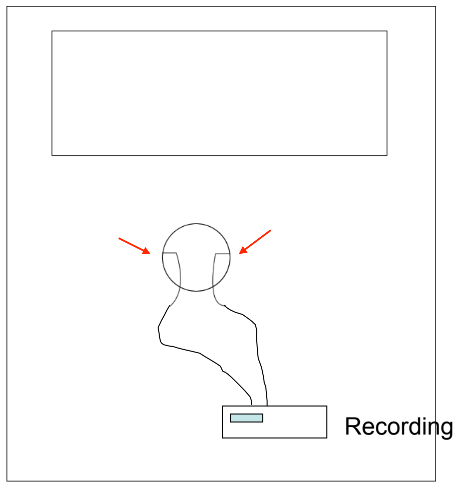


Aesthetic is to create a stereo recording with maximum clarity and impact.

Sound sources are treated in ways that are most appropriate to each: strings are reverberated, but electric bass is not; voice is center and interfering instruments are to the side.

The spatial design does not match any reality.

Dummy Head Recording



Recording Device

Aesthetic is most closely related to classical recording and extends to virtual reality.

All sound sources are contained in a signal acoustic space.

Dummy Heads

Figure 4.26. The Brüel and Kjaer HATS head and torso simulator (model 4128) for acoustical research in areas such as evaluation of headphones, telephones, hearing aids and hearing protectors. The geometry of the head is symmetrical, based on average adult anthropometric data. *Photograph courtesy of Brüel and Kjaer.*

B&K 4128

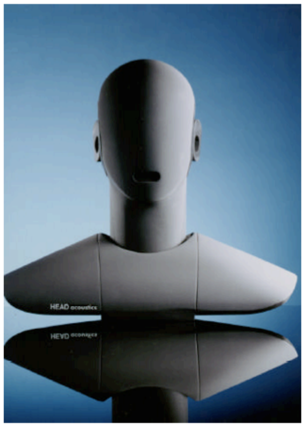


Figure 4.27. The AACHEN Head Model HMS II, manufactured by HEAD acoustics. The outer ears are designed according to a "structural averaging" of HRTFs; note the inclusion of the upper torso. Two versions of the head are available: the HMS II, intended for analysis within an integrated system, and the HRS II, for recording studio applications. *Photograph courtesy of Wade Bray, Sonic Perceptions–HEAD acoustics.*

Aachen Head

Dummy Heads

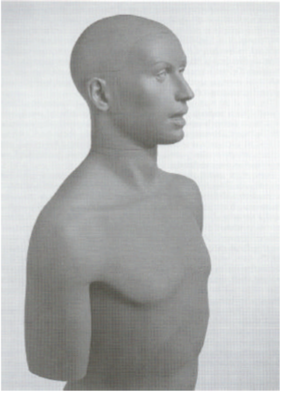


Figure 4.25. The KEMAR mannequin head and torso. KEMAR is used extensively in audiological research, particularly for hearing aids; several types of pinnae are available. Standard laboratory measurement microphones attach to a Zwislocki ear canal simulator inside the head. *Photograph courtesy of Knowles Electronics.*

KEMAR

Dummy Heads

Figure 4.28. The Neumann KU-100 dummy head. This is the third generation of Neumann binaural recording devices (following the KU-80 and KU-81). *Photograph courtesy of Juergen Wahl, Neumann USA.*

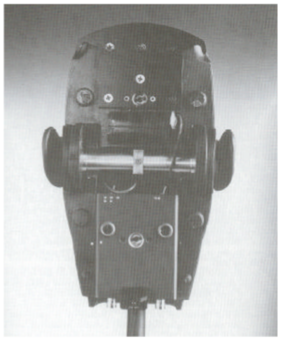
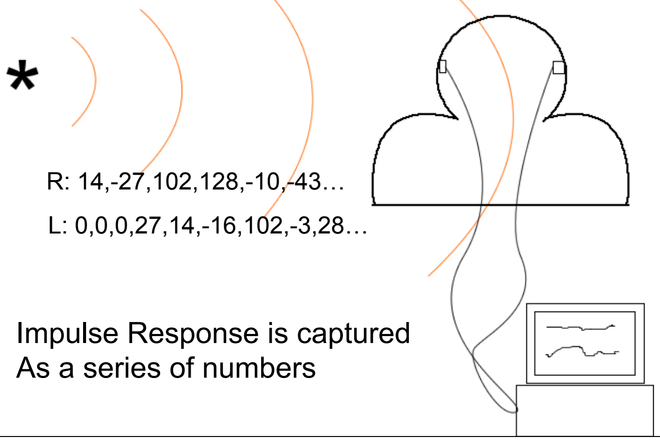


Figure 4.29. Inside of the Neumann KU-100 showing microphone capsules and power supply. *Photograph courtesy of Juergen Wahl, Neumann USA.*

Neumann KU-100

Simulating Binaural Recording Through Digital Signal Processing (DSP)

HRTFs can be captured and stored



R: 14,-27,102,128,-10,-43…

L: 0,0,0,27,14,-16,102,-3,28…

Impulse Response is captured
As a series of numbers

Simulating Binaural Recording Through Digital Signal Processing (DSP)

System Overview

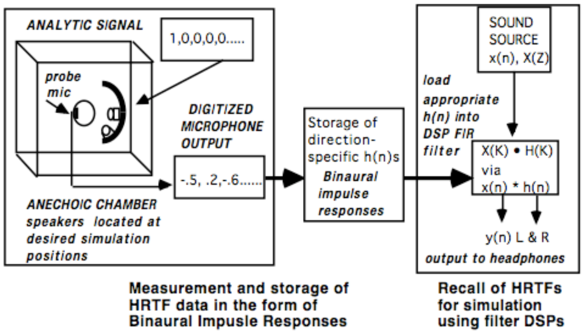


Figure 4.17. Overall plan of the HRTF measurement-storage-simulation technique. HRTFs are obtained using acoustic signals from loudspeakers that are recorded with probe microphones in the vicinity of the ear canals. The data are stored for input to DSP filters during the simulation stage.

Simulating Binaural Recording Through Digital Signal Processing (DSP)

The HRTFs are represented as HRIRs. These are simply a series of numbers, h(nT).

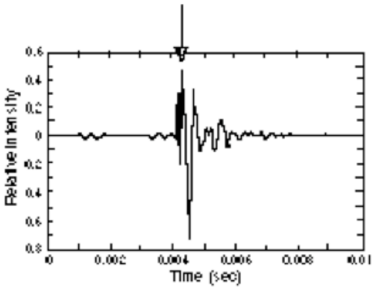


Figure 4.18. A time domain **impulse response** measured in the ear canal in an anechoic chamber. Its FFT would be equivalent to the HRTFs shown in Figures 2.13–2.15. The arrow indicates the peak of the response.

Simulating Binaural Recording Through Digital Signal Processing (DSP)

The HRIR, h(nT), can be used directly to recreate the HRTFs as FIR (finite impulse response) filters. Of course, you need filters for both ears.

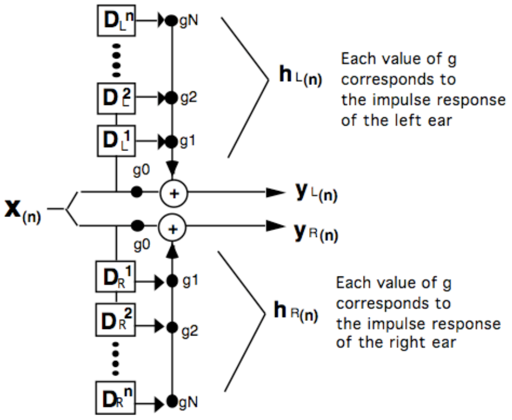


Figure 4.19. Convolution using two separate FIR filters for binaural output. Two impulse responses, one each for left and right ear HRTFs, are applied to the single input, resulting in a two-channel output.

Simulating Binaural Recording Through Digital Signal Processing (DSP)

The amount of computation for an FIR filter is proportional to the number of data points, ie. filter coefficients. Reducing the number of coefficients alters the HRTF.

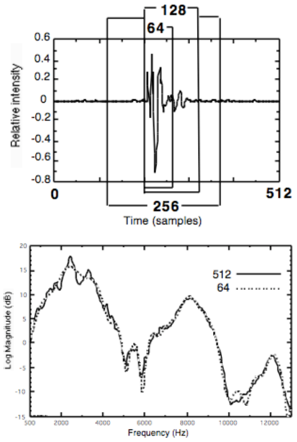


Figure 4.31. Time domain windowing (top) of an original impulse of 512 points narrowed to 256, 128, and 64 points; (bottom) a comparison of the magnitude response of the original impulse and a

Simulating Binaural Recording Through Digital Signal Processing (DSP)

The HRIR can be processed to create a 'minimum phase' version that has the fewest possible coefficients, but the overall HRIR delay has to be recreated by the system.

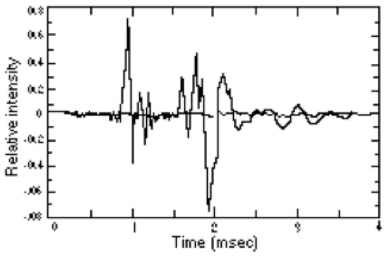
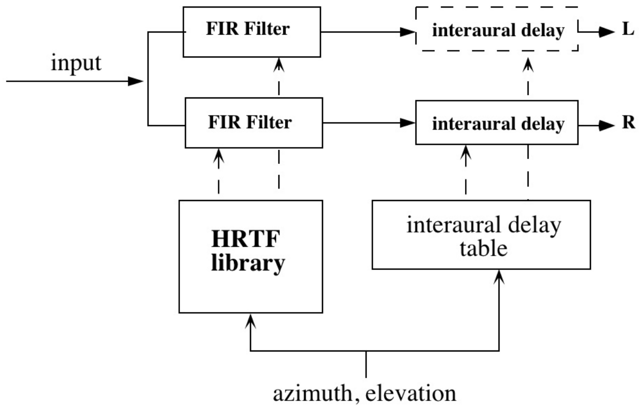


Figure 4.32. Minimum-phase version (dotted line) of a measured impulse response measurement. The process requires "reinserting" an approximation of the delay.

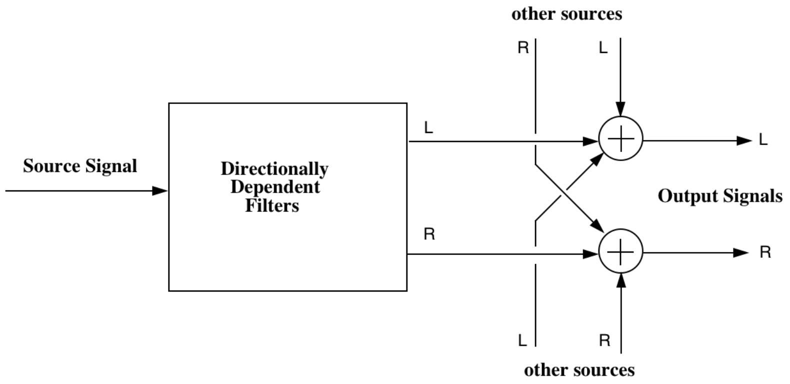
Simulating Binaural Recording Through Digital Signal Processing (DSP)

Here the system selects a pair of HRTFs and an interaural delay on the basis of source azimuth and elevation. The delay is implemented in series with the FIR filter.



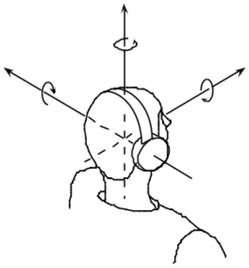
Simulating Binaural Recording Through Digital Signal Processing (DSP)

Each sound source needs its own pair of HRTF filters with delay. The output for each sound source is added together to create a composite output.



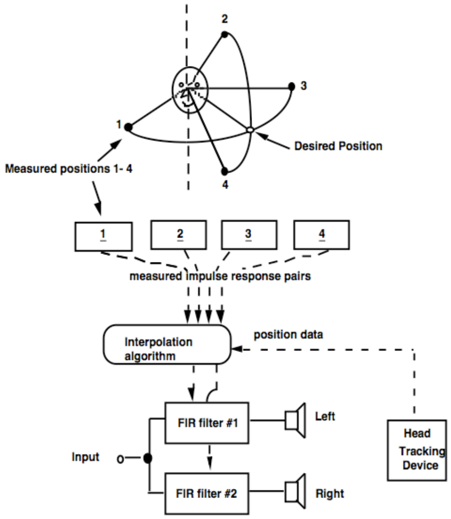
Simulating Binaural Recording Through Digital Signal Processing (DSP)

Since sound sources change position in space, the HRTF filters need to be updated. This is especially true if one includes head-tracking!



Simulating Binaural Recording Through Digital Signal Processing (DSP)

Instantaneous locations do not always match the location of the HRTFs in the library. Therefore, for locations that fall inbetween the recorded ones, new filters need to be interpolated.



Simulating Binaural Recording Through Digital Signal Processing (DSP)

FIR interpolation can produce unexpected results. Best results are obtained with minimum-phase HRIRs.

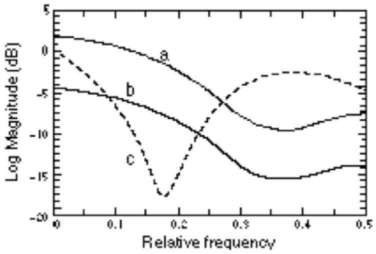


Figure 4.39. Linear interpolation of two simple FIR filters, a and b. Ideally, one would obtain an intermediate curve from interpolation that would be inbetween a and b; instead, the curve labeled c results.