What Audio Engineers Should Know About Human Sound Perception

Part 2. Binaural Effects and Spatial Hearing

AES 112th Convention, Munich
AES 113th Convention, Los Angeles

Durand R. Begault

Human Factors Research & Technology Division
NASA Ames Research Center
Moffett Field, California
Overview

• ILD, ITD differences and lateralization
• HRTF spectral changes for 3D imagery
• Binaural versus monaural influence of echoes
• Effects of reverberation on perception of the environmental context
• Cues to auditory distance
• Cognitive and multisensory cues
Communication chain for acoustic events

Sound source(s), interaction with room acoustics

**SOURCE** ➔ **MEDIUM** ➔ **RECEIVER**

*Frequency, Amplitude, Spectrum, Location*
Communication chain for acoustic events

Sound source(s), interaction with room acoustics

Recording & playback: acoustical-electrical-acoustical transformation

SOURCE

MEDIUM

RECEIVER

Frequency
Amplitude
Spectrum
Location
Communication chain for acoustic events

Sound source(s), interaction with room acoustics

Recording & playback: acoustical-electrical-acoustical transformation

Hearing: perception, cognition, multisensory interaction

SOURCE

MEDIUM

RECEIVER

Frequency
Amplitude
Spectrum
Location

Pitch
Loudness
Timbre
Localization
Communication chain for acoustic events

- **Source(s)**, interaction with room acoustics
- Recording & playback: acoustical-electrical-acoustical transformation
- Hearing: perception, cognition, multisensory interaction

Mismatch between prescribed & perceived spatial events
Model of the binaural hearing system

Binaural hearing (localization; signal separation & detection):
forming spatial auditory events from acoustical (bottom-up) and psychological (top-down) inputs

Model of the binaural hearing system

Binaural hearing
(localization; signal separation & detection)

Filtering of acoustic signal
by pinnae, ear canal
Model of the binaural hearing system

**Binaural hearing**
(localization; signal separation & detection)

**Filtering by inner ear; frequency-specific neuron firings**

**Filtering of acoustic signal by pinnae, ear canal**
Model of the binaural hearing system

Binaural hearing (localization; signal separation & detection)

Physiological evaluation of interaural timing and level differences

Filtering by inner ear; frequency-specific neuron firings

Filtering of acoustic signal by pinnae, ear canal
Model of the binaural hearing system

Acoustic signal-driven

- Filtering of acoustic signal by pinnae, ear canal.
- Filtering by inner ear; frequency-specific neuron firings
- Filtering of acoustic signal by pinnae, ear canal.

Physiological evaluation of interaural timing and level differences

Binaural hearing (localization; signal separation & detection)

Multi-sensory information; cognition

Psychologically-driven

- Non-auditory sensory information
- Formation of the auditory event
- Binaural—activity pattern

Multi-sensory information; cognition
Two important functions of the binaural hearing system for recording engineers:

• Localization
  (lateral and 3-dimensional)

• Binaural masking:
  Echo supression, room perception
Lateral localization of auditory images

“Duplex” theory of localization

- ILD (interaural level difference)
- ITD (interaural time difference)
Lateral spatial image shift

- **ILD** (interaural level difference) caused by head shadow of wavelengths $> 1.5$ kHz

---

**Graphs:**

- **6,000 Hz**
- **200 Hz**
Perceptual decoding of spatial cues in a cross-coincident microphone recording is based on ILDs.
Lateral image shift

- ITD (interaural time difference)
Lateralization demo. A simple time or level difference can make headphone images move from side to side inside the head.

1. ILD DEMO:
   - 2 dB
   - 4 dB
   - 6 dB
   - 8 dB
   - 12 dB

2. ITD DEMO:
   - 0.00 ms
   - 0.25 ms
   - 0.50 ms
   - 0.75 ms
   - 1.00 ms
   - 1.50 ms

Adapted from Toole & Sayers, 1965 and Blauert, 1983: click stimuli
Adapted from Blauert, 1983: broadband noise
Elevation and front-back discrimination: HRTF, pinnae cues
The cone of confusion causes reversals for virtual sources with identical or near-identical ITD or ILD.
Head-related transfer function cues (HRTFs) provide cues for front-back discrimination and elevation.

3. audio example: HRTF “clock positions”
Variation in HRTF magnitude with elevation at one azimuth

4. Audio example:

120 degree azimuth: at

+36,

0,

-36 degrees elevation

Graphic by William L. Martens,
University of Aizu
Perceptual errors with headphone 3-D sound include inside-the-head localization (solution: reverberation cues) and reversals (solution: head tracking)
Localization error for headphone stimuli (azimuth)

Anechoic Speech: Individual differences

Mean values for different reverberation conditions
Echoes, reverberation and background sound: perception of the environmental context
Spatial hearing fundamentally involves perception of the location of a sound source at a point in space (azimuth, elevation, distance).

But a sound source simultaneously reveals information about its environmental context.
- reverberation
- image size & extent
Effect of delay time for a single echo

Image shift $\rightarrow$ Image broadening $\rightarrow$ Echo

Approximate delay time to left channel (msec)

Sound examples: 5. stereo echo - 6. monaural echo

Relative to the reference condition, spatially separated echoes create spatial percepts; non-spatially separated echoes create timbral effects.
Early and late reverberant sound fields

7. Audio examples:
- direct sound
- direct w/ 1st, 2nd order ERs
- direct with full auralization
Early and late reverberant sound fields

8. audio examples: normal and 0.25 speed impulse response
Echo thresholds

- Sensitivity can increase as much as 10 dB if echoes occur at different locations
- Late reverberation can decrease sensitivity
- Sensitivity increases with increasing time delay
Although thresholds for reverberation are relatively low, background noise (e.g., NC 35) can mask the reverberant decay.
Distance perception: amplitude cues

- The inverse square law states that sound decays 6 decibels per doubling of distance in a reflection-free environment.

9. sound example
Distance perception: amplitude cues

However, “half-as-loud” corresponds to a 10 dB reduction in level with distance

10. sound example
Distance perception: reverberant ratio cues

An increase in reverberant level indicates movement into the diffuse sound field.
Concert Hall reverberation physical-perceptual parameters

- Reverberance (*reverberation time, strength*)
- Apparent source width (ASW) (*interaural cross-correlation*)
- Envelopment (*spatial diffusion of reflections from all around*)
- Clarity (*ratio of first 50-80 ms of early sound to late sound*)
- Warmth (*ratio of bass frequency RT to mid-band RT*)
Cognitive cues; multisensory cues
Cognitive cues to distance perception

![Graph showing judged vs. actual position for shouting and whispering]
Auditory localization can be influenced or biased by cognitive mapping
Influence of visual, vibratory cues

Helicopter fly-overs

Explosions & crashes
Summary

• ILD, ITD differences and lateralization

• HRTF spectral changes for 3D imagery

• Binaural versus monaural influence of echoes

• Effects of reverberation on perception of the environmental context

• Cues to auditory distance

• Cognitive and multisensory cues