

# **Binaural Hearing**

HST.723 Neural Coding and Perception of Sound

© Andrew J. Oxenham (2005)

# Binaural Hearing

## Why two ears?

- *Locating sounds in space*: acuity is poorer than in vision by up to two orders of magnitude, *but* extends in all directions. Role in alerting and orienting?
- *Separating sound sources*: signals from two receivers can be compared and operated on to enhance the reception of one sound source over another. The same applies to signals in a reverberant environment.

## Definitions

- Monaural: Sound presented to one ear
- Binaural: Sound presented to both ears
  - Diotic: Sounds identical at the two ears
  - Dichotic: Sounds different at the two ears

Figure removed due to copyright considerations.  
Please see: Blauert, J. *Spatial Hearing: The Psychophysics of Human Sound Localization*. Cambridge, MA: MIT Press, 1983.

# Spatial cues

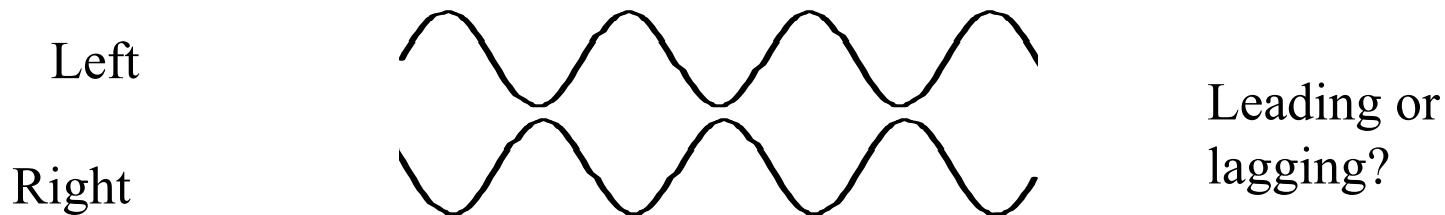
Two primary cues for determining the azimuth of a sound:

a) Interaural time differences (ITDs)

$c \approx 340$  m/s; max. path difference  $\approx 0.23$  m

$\Rightarrow$  Max. natural ITD  $\approx 680$   $\mu$ s.

For sinusoids, this can be expressed as a phase difference. As the maximum delay approaches a lag of  $180^\circ$ , the phase difference becomes ambiguous. This occurs for  $f > 1.5$  kHz.



b) Interaural intensity differences (IIDs – also known as ILDs).

The head starts to create an acoustic shadow at frequencies above about 500 Hz. The effect is most apparent above 2 kHz.

# Sound at the two ears

Figures removed due  
to copyright reasons.

Interaural time differences

Interaural level differences

(From Moore, 1997)

# Localization of sinusoids in the horizontal plane

“Duplex theory” (Rayleigh, 1907):

$f < 1500$  Hz: ITDs dominant (ILDs are small)

$f > 1500$  Hz: ILDs dominant (ITDs cannot be perceived and would be ambiguous)

Figure removed due to copyright reasons.

Stevens and Newman (1936):

From Moore (1997)

Localization worst around 2-3 kHz.

Front-back reversals common, especially below 2 kHz. Can be explained by the “cone of confusion.”

## Minimum Audible Angles (MAA) for sinusoids, first measured by Mills (1958).

### Points to note:

- $1^\circ$  corresponds to an ITD of about  $10 \mu\text{s}$ . (!)
- MAA is a function of frequency and reference azimuth.
- The dip in performance around 1.5 kHz may be explained in terms of the duplex theory: neither mechanism is operating efficiently in that region.
- Measurements could not be made with the reference at  $90^\circ$  at any frequency: Due to cone of confusion.

Figure removed due to copyright reasons.

# Lateralization of sinusoids

ITDs and ILDs can be manipulated independently when sounds are presented over headphones, but the sound generally appears to come from inside the head. Lowest JNDs: 10  $\mu$ s ITD; 0.5-1 dB ILD.

Localization: Free field

Lateralization: Headphones

Figure removed due to copyright considerations.

Please see: Blauert, J. *Spatial Hearing: The Psychophysics of Human Sound Localization*. Cambridge, MA: MIT Press, 1983.

Results for a 600-Hz tone. *No effect* of ITD on lateralization is observed above about 1.5 kHz, if sound is gated on and off slowly.

Note that lateralization occurs for a 600-Hz tone, even though large ILDs do not occur naturally at this frequency.

# Lateralization of complex sounds

ITDs much greater than found naturally can be discriminated and can produce lateralization.

## Role of envelope fluctuations

- ITDs at high frequencies can be discriminated, and do cause lateralization, in complex sounds.
- Henning (1974): ITD discrimination of a 3900-Hz tone, amplitude modulated with a 300-Hz modulator, was almost as good as ITD discrimination of a 300-Hz tone. Listeners were not sensitive to the phase (delay) of the carrier; even slightly different carrier frequencies produced lateralization.

*Ongoing ITDs in the stimulus envelope can be detected at high frequencies.*

Figure removed due to copyright considerations.  
Please see: Blauert, J. *Spatial Hearing: The Psychophysics of Human Sound Localization*.  
Cambridge, MA: MIT Press, 1983.



# Time-intensity trades in lateralization

- Can an ITD and an ILD “cancel” each other out?

Figure removed due to copyright considerations.

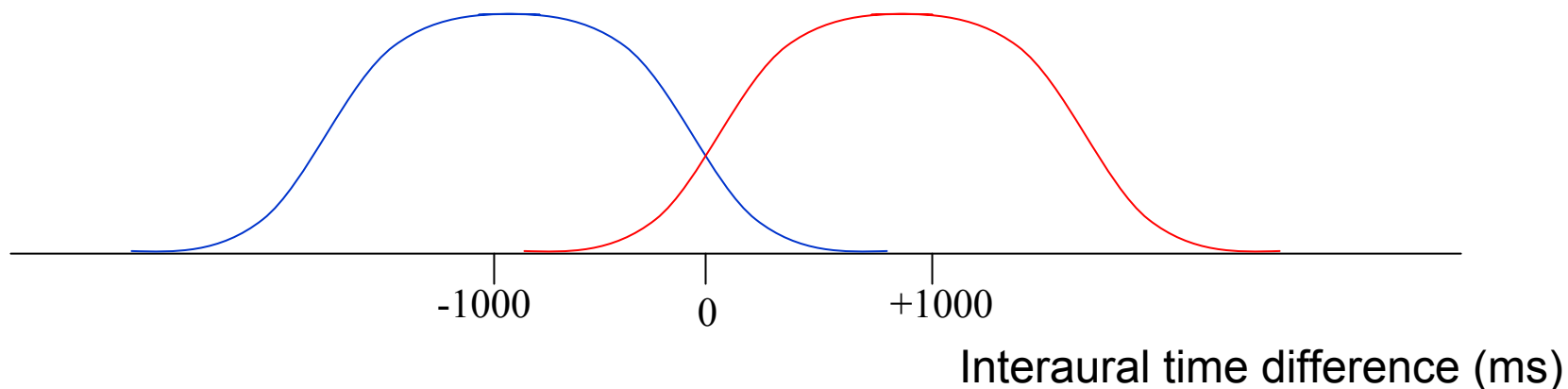
Please see: Blauert, J. *Spatial Hearing: The Psychophysics of Human Sound Localization*.  
Cambridge, MA: MIT Press, 1983.

# Possible underlying mechanisms

Jeffress (1948) coincidence counter to account for ITD sensitivity: Neural spatial “map”.

Figure removed due to copyright reasons.

May apply to birds (or at least barn owls). A different arrangement may apply to mammals (McAlpine et al., 2001).



# Single-mechanism hypotheses

Ability to “trade” time and intensity differences led to one-mechanism proposals:

- Level-dependent latency: Increased level leads to decreased neural latency, transforming ILD to ITD.
- Neural inhibition: Earlier signal leads to inhibition in contralateral pathways, transforming ITD to ILD.

Evidence against single-mechanisms hypotheses:

- ITDs cannot compensate for ILDs of 25 dB or more.
- Time-intensity trades do not evoke the same percept as a diotic stimulus: image is more diffuse or may separate into two separate images.
- Sensitivity to ongoing ITDs in envelope cannot be explained by simple inhibition.

# Resolving the Cone of Confusion: Hearing Height

- The cone of confusion can be resolved using head movements. But, for broadband sounds, position in the median plane can sometimes be determined even without head movements.
- Due primarily to direction-dependent filtering effects of the head and pinnae. The filter between the source and the eardrum is referred to as the Head-Related Transfer Function (HRTF).

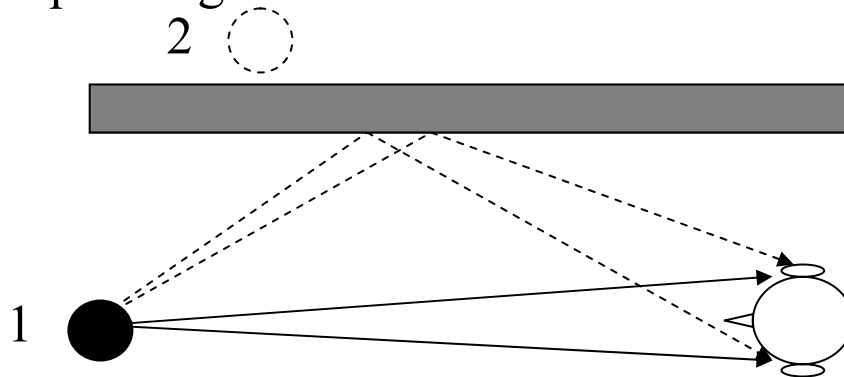
Figure removed due to copyright considerations.

Please see: Blauert, J. *Spatial Hearing: The Psychophysics of Human Sound Localization*.  
Cambridge, MA: MIT Press, 1983.

- A realistic simulated HRTF can result in “externalization” of sounds over headphones. Important for “virtual reality” applications.

# Listening in rooms: Echo suppression and the precedence effect

- In many environments, the direct path of a sound to our ears is only one of many. Yet we mostly hear only one source at a location usually corresponding to the that of the source.



This situation can be modeled using 2 loudspeakers (at 1 & 2)

Three time regions identified (times valid for clicks):

- 1) Summing localization ( $< 1$  ms delay): two events fused: perceived location is a weighted sum of the two.
- 2) Precedence effect (ca. 1-5 ms delay): Only one sound perceived: direction of first sound dominant.
- 3) Echo threshold ( $> 5$  ms delay): Two sounds heard.

# **Build-up and Breakdown of the Precedence Effect**

The lagging click of a click pair with a delay of 8 ms is initially clearly heard. However, after a few repetitions (say, 4 per second), the lagging click is no longer perceived as a separate event (Freyman et al., 1991). The auditory system seems to have “adjusted” to the new acoustic environment.

Changing the delay, direction, or spectrum of a lagging click in a series will suddenly render it audible. After a few more repetitions, it will “disappear” again (Clifton, 1987). The auditory system is alerted to a change in acoustic environment.

(The precedence effect, or “law of first wavefront” is used in the design of PA systems to make the amplified sound appear to come from the source.)

# Binaural Masking Level Differences

BMLDs can be as much as 15 dB for a tone in broadband noise, and even more for a tone in narrowband noise. The effect is robust and can also be demonstrated for speech in noise.

Figures removed due to copyright reasons.

From Moore (1997)

# Motion Perception

Binaural system is often referred to as being “sluggish”.

Using continuously varying ITDs or ILDs, Blauert (1972) found the maximum rate at which movement could be followed “in detail” was less than 3 Hz.

Above about 10-20 Hz, no movement is perceived at all – just a steady blur (Grantham & Wightman, 1979).

Similar time constants govern BMLDs in a changing environment (Kollmeier & Gilkey, 1990):

Figures removed due  
to copyright reasons.



# Spatial Hearing and Auditory Grouping

- Empirical results seem contradictory:

## *Evidence for the influence of spatial cues*

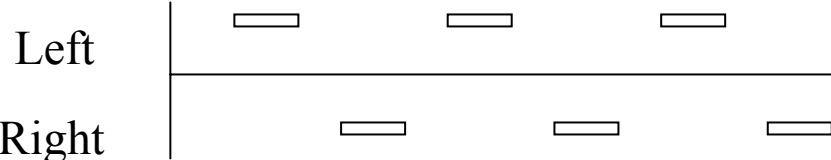
- Benefit of different ITDs (Darwin et al., 1999) or perceived spatial separation (Freyman et al., 1999) for speech in speech.
- Monophonic melodies hard to follow across the ears (van Noorden, 1975; Bregman, 1990).

## *Evidence against the influence of spatial cues*

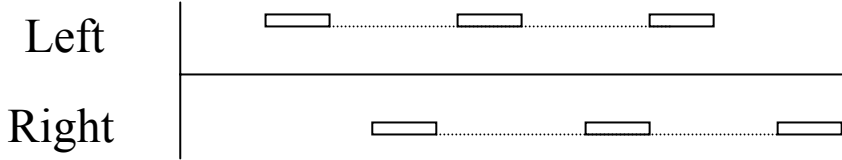
- Simultaneous vowels cannot be segregated using ITDs (Culling and Summerfield, 1995).
- Spatial differences are ignored in concurrent melodies are ignored (Deutsch, 1975; Tchaikovsky).

# Spatial Organization

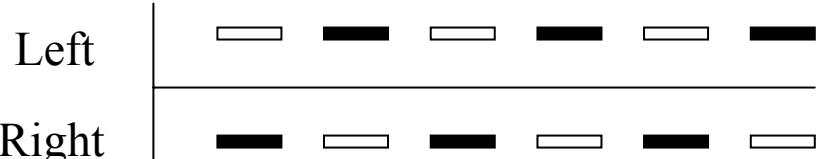
played as:



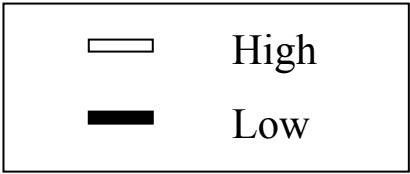
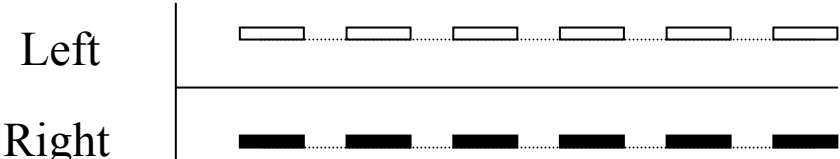
heard as:



played as:



heard as:



# **Spatial cues may help to form auditory streams, but not objects**

- In sequential streams of isolated tones, spatial cues play a large role in assigning individual tones to streams.
- However, when more than one tone is present at one time, with other cues (frequency proximity, common onset, etc.) conflicting, spatial information does not play a strong role in forming objects.
- Makes intuitive sense if one considers the ambiguity of spatial cues in a reverberant environment.

# References

- Blauert, J. (1972). "On the lag of lateralization caused by interaural time and intensity differences," *Audiol.* **11**, 265-270.
- Bregman, A. S. (1990). Auditory Scene Analysis: The Perceptual Organisation of Sound (MIT Press, Cambridge, MA).
- Clifton, R. K. (1987). "Breakdown of echo suppression in the precedence effect," *J Acoust Soc Am* **82**, 1834-1835.
- Darwin, C. J., and Hukin, R. W. (1999). "Auditory objects of attention: the role of interaural time differences," *J. Exp. Psychol. Hum. Percept. Perform.* **25**, 617-629.
- Deutsch, D. (1975). "Two-channel listening to musical scales," *J. Acoust. Soc. Am.* **57**, 1156-1160.
- Freyman, R. L., Clifton, R. K., and Litovsky, R. Y. (1991). "Dynamic processes in the precedence effect," *J Acoust Soc Am* **90**, 874-884.
- Freyman, R. L., Helfer, K. S., McCall, D. D., and Clifton, R. K. (1999). "The role of perceived spatial separation in the unmasking of speech," *J. Acoust. Soc. Am.* **106**, 3578-3588.
- Grantham, D. W., and Wightman, F. L. (1979). "Detectability of a pulsed tone in the presence of a masker with time-varying interaural correlation," *J. Acoust. Soc. Am.* **65**, 1509-1517.
- Henning, G. B. (1974). "Detectability of interaural delay in high-frequency complex waveforms," *J. Acoust. Soc. Am.* **55**, 84-90.
- Kollmeier, B., and Gilkey, R. H. (1990). "Binaural forward and backward masking: evidence for sluggishness in binaural detection," *J. Acoust. Soc. Am.* **87**, 1709-1719.
- Mills, A. W. (1958). "On the minimum audible angle," *J. Acoust. Soc. Am.* **30**, 237-246.
- Rayleigh, L. (1907). "On our perception of sound direction," *Phil. Mag.* **13**, 214-232.
- Stevens, S. S., and Newman, E. B. (1936). "The localization of actual sources of sound," *Am. J. Psychol.* **48**, 297-306.
- van Noorden, L. P. A. S. (1975). "Temporal coherence in the perception of tone sequences," Ph.D. Thesis, Eindhoven University of Technology.