Hearing II
Perceptual Aspects

Overview of Topics
Chapter 6 in Chaudhuri
- Intensity & Loudness
- Frequency & Pitch
- Auditory Space Perception

Intensity & Loudness

- **Loudness** is the subjective perceptual quality of sound related to pressure or intensity
- Loudness also depends on frequency of sound, among other things

Absolute Detection Thresholds

- The pressure at which loudness is greater than 0 (i.e., absolute threshold) varies according to Hz
- Max. sensitivity at 2-5 kHz, also where human speech has most energy
Fundamental Concept: Psychometric vs. Psychophysical Functions

- **Psychometric:**
  - Function of physical quantity & behaviour
  - Shows one threshold
- **Psychophysical Function:** Function of two physical quantities
  - Shows how threshold on one physical quantity varies as a function of the other

Loudness

- Other factors affecting loudness:
  - Binaural vs. Monaural presentation: Threshold is 6 dB lower with two ears (spatial integration)
  - Duration of sound: Threshold increases below 200 ms (temporal integration)
Fundamental Concept: Spatial & Temporal Integration

- Any sensor (biological or artificial) must integrate signals across a certain range of space and time
- Longer/larger integration provides greater sensitivity but lower acuity (details can be lost)
- Recall spatial integration in somatosensory systems

Difference Thresholds

- As with absolute thresholds, difference thresholds vary with Hz
- That is, the Weber fraction varies with Hz
- It also varies with intensity, only remaining constant across moderate intensities

Off-Frequency Spread of Activation

- Generally, k drops with intensity
- Greater intensity = larger travelling wave on the basilar membrane
- \( \therefore \) excitation in neurones whose CFS are adjacent to that of the sound’s frequency

Questions?

- What is the minimal audibility curve?
- What is the dynamic range of the auditory system?
- What is a psychometric function? What is psychophysical function?
Noise Masking Experiments

- A type of psychophysical experiment examining how one sound affects perception of another
- Two basic types:
  - Tonal masking: How much does one pure tone impair perception of another?
  - Noise masking: How much does aperiodic sound impair perception of tones?

Two basic methods:
- Constant Target Hz, Varying Masker Hz (typically used with tonal masking)
- Varying Target Hz, Constant Masker Hz (typically used with noise masking)
- Together, have revealed important things about how auditory system functions

Tonal Masking

- Typically, a test tone of fixed frequency and amplitude is chosen, say 1000 Hz at 40 dB
- Then a series of masking tones of different frequencies are presented
- For each mask, the observer adjusts its intensity (dB) until it just drowns out the test tone

Results shown for 1000 Hz test tone
- Most effective mask is at same frequency
- Note the asymmetry of off-frequency masks
Noise Masking

- A constant noise mask is used
- Typically, this will be narrow-band white noise, meaning an aperiodic sound with frequencies in a certain limited range
- Example: Noise with energy at 410±45 Hz. We say this has a centre frequency of 410 Hz and a bandwidth of 90 Hz.

17

Noise Masking

- The effect of the noise mask on the absolute threshold for pure tones is measured
- This is done for test tones across the frequency spectrum
- Threshold is most elevated for test tones near the frequency of the noise
- Again, asymmetry is seen off-frequency

18

Noise Masking

- We can subtract the baseline from the masked audibility curve
- This gives us the threshold elevation produced by the noise mask
- Threshold elevation shows asymmetry. The mask is more effective against higher Hz.

19

Noise Masking

Red: Baseline minimal audibility curve (no mask)
Green: Masked minimal audibility curve (410±45 Hz noise mask)
Masking Asymmetries

- Why do masking asymmetries occur?
- Because of asymmetries in how the basilar membrane responds to pure tones
- Specifically, the travelling wave builds up gradually then collapses suddenly

This means the responses to test and masking tones will be...

...more similar if the mask is lower in frequency than the test

...more different if the mask is higher in frequency than the test

Questions?

- If a masking tone is 500 Hz in frequency:
  - What Hz of target tone will it most effectively mask?
  - Which will it mask more: A 400 Hz tone, or a 600 Hz tone?
  - What is the reason for the asymmetry in masking?

Frequency & Pitch

- Relate to one another in a complex way.
- As frequency increases, one aspect of pitch, called tone height, increases linearly
- But another aspect, tone chroma, changes in a circular fashion
Height & Chroma:
Do-Re-Mi...Do-Re-Mi

| C4 | C5 | C6 |
| D4 | D5 | D6 |
| E4 | E5 | E6 |
| F4 | F5 | F6 |
| G4 | G5 | G6 |
| A4 | A5 | A6 |
| B4 | B5 | B6 |

Height & Chroma:
Do-do-do, Re-re-re

| C4 | C5 | C6 |
| D4 | D5 | D6 |
| E4 | E5 | E6 |
| F4 | F5 | F6 |
| G4 | G5 | G6 |
| A4 | A5 | A6 |
| B4 | B5 | B6 |

Harmony & Disharmony

| C4 | C5 | C6 |
| D4 | D5 | D6 |
| E4 | E5 | E6 |
| F4 | F5 | F6 |
| G4 | G5 | G6 |
| A4 | A5 | A6 |
| B4 | B5 | B6 |

Auditory Localization

- Auditory space - surrounds an observer and exists wherever there is sound
- Locations in auditory space are defined by:
  - Azimuth coordinates - position left to right
  - Elevation coordinates - position up and down
  - Distance coordinates - position from observer
Sound Localization Accuracy

On average, people can localize sounds directly in front of them more accurately than those to the side or behind them.

In vision, source cues are present on the retina.

In audition, source cues are NOT present on the cochlea.
Questions

• What are the three coordinates used to describe sound location?
• In what directions are humans most accurate in judging sound location? Least?

Cues for Sound Localization

• Binaural Cues:
  • Interaural Time Difference (ITD)
  • Interaudial Intensity Difference (IID)
• Monaural cue:
  The Head-Related Transfer Function (HRTF)

Cues for Sound Location

• Binaural cues - location cues based on the comparison of the signals received by the left and right ears
• Interaural time difference (ITD) - difference between the times sounds reach the two ears
• When distance to each ear is the same, there are no differences in time
• When the source is to the side of the observer, the times will differ

Interaural Time Differences
Binaural Cues

- Interaural Intensity Difference (IID) difference in sound pressure level reaching the two ears
- Reduction in intensity occurs for high frequency sounds for the far ear
- The head casts an acoustic shadow
- This effect doesn’t occur for low frequency sounds, which diffract around the head

Interaural Intensity Difference

The higher the frequency of the sound, the greater the intensity (i.e., level) difference between the two ears when sounds are coming from the side.

A Given ITD/IID...

...Cannot, on its own, tell you where a sound is coming from
...Can only tell you it is coming from somewhere on a particular cone of points
...We call these cones of confusion
Monaural Cue for Sound Location: HRTF

- As sounds encounter the head, they are modified by its structures (bones, muscles, etc.), esp. the pinna.
- Some frequencies are reduced in amplitude, others, due to resonance and constructive interference, are increased.
- The pattern of increases and decreases is called the Head-Related Transfer Function (HRTF).
- Importantly, the HRTF differs depending on the elevation of the sound source.

Thus, the head and pinna leave a unique “frequency fingerprint” on sounds.

Consider how each frequency element in the bassoon’s sound would be affected by the HRTF for different elevations.
Hofmann Experiment on Judging Elevation

- IID and ITD are not effective for judgments on elevation since in many locations they may be zero
- Experiment by Hofmann investigating spectral cues
  - Listeners were measured for performance locating sounds differing in elevation
  - They were then fitted with a mould that changed the shape of their pinnae

Results of Hofmann Exp on Elevation Judgments

- Right after the moulds were inserted, performance was poor
- After 19 days, performance was close to original performance
- Once the moulds were removed, performance stayed high
- This suggests that there might be two different sets of neurones—one for each set of HRTF cues

Questions

- What are the main binaural cues for sound location? For which coordinates do they work?
- Which binaural cue works for high frequencies but not low ones?
- What is the HRTF? How does it help sound localization?
Methods for Measuring Sound Location Accuracy

- Two general experimental methods are used for measuring accuracy of sound localization.
- Free-field presentation - sounds are presented by speakers located around the listener's head in a dark sound-proof room.
- Headphone presentation with applied IIDs, ITDs and HRTFs.

Free Field Presentation

- Sounds are presented by speakers located around the listener's head in a dark room.
- Listener indicates location by pointing or by giving azimuth and elevation coordinates.
- Advantage: Highly naturalistic stimulus properties.
- Disadvantage: Expensive equipment, some lack of control over exact sound contents.

Headphone Presentation of Sounds

- Advantage: Experimenter has precise control over sounds. Also, much cheaper.
- Disadvantage: Cues from the pinna (HRTF) are eliminated, which results in the sound being internalized.
- Sound can be externalized by measuring each subject's HTRF and applying it to the presented sounds.
Questions

- What are the two main methods of testing localization accuracy? What are their advantages and disadvantages?
- Which binaural cue supersedes the others for low frequency sounds?
- What did Hoffman’s experiment on pinna shape show?

Physiological Representation of Auditory Space

- Interaural time-difference detectors: neurones that respond to specific ITDs
- Found in auditory cortex and at the first nucleus (superior olivary) in the system that receives input from both ears
- Topographic map: Neural structure that responds to locations in space (not to be confused with tonotopic)

Topographic Maps

- Barn owls have neurones in the mesencephalicus lateralis dorsalis (MLD) that respond to locations in space
- Mammals have similar maps in the subcortical structures, such as the inferior colliculus
- These neurones have receptive fields for sound location

The receptive field for location (the rectangles) of three topographic neurones in the owl’s MLD.
The Auditory Cortex

• Even though there are topographic maps in subcortical areas of mammals, there is no evidence of such maps in the cortex.
• Instead, panoramic neurones have been found that signal location by their pattern of firing.
• In general, the where stream shows more specific neural responses for location the further upstream one goes in the cortex.

Panoramic Neurons

Panoramic neurones fire different patterns of bursts depending on the direction from which sound is coming.

Questions

• What does it mean to say that some animals have a topographic map of sound locations in their cortices?
• What is a panoramic neurone?

Auditory Scene Analysis
Identifying Sound Sources

- Auditory Scene - the array of all sound sources in the environment
- Auditory Scene Analysis - process by which sound sources in the auditory scene are separated into individual perceptions
- This does not happen at the cochlea since simultaneous sounds will be together in the pattern of vibration of the basilar membrane

Auditory Scene Analysis

- Aspects of Auditory Scene Analysis
  - Sound localization
  - Segregation of sound signal into those coming from individual sources.
  - Grouping of separate sounds into those coming from a given source

Fundamental Concept: Gestalt Heuristics

- Recall that a heuristic is a rule of thumb used by a system (e.g., the brain)
- A gestalt heuristic is such a rule applied to organizing sensory inputs to determine (e.g.):
  - Which elements of a scene belong to which objects
  - Which elements represent edges between an object and its background
**Fundamental Concept: Gestalt Heuristics**

- Many gestalt heuristics have been proposed.
- Examples include the *proximity principle* and the *similarity principle*.
- Obviously, these are only partially defined.

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>B</td>
<td>B</td>
</tr>
<tr>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>B</td>
<td>B</td>
</tr>
</tbody>
</table>

**Principles of Auditory Grouping**

- Like visual stimuli, sound stimuli tends to be perceptually organized according to the following heuristics:
  - Proximity: A single sound source comes from one location.
  - Smoothness: A single sound source changes (location, pitch, etc.) in a smooth & continuous way.
  - Similarity: A single sound source produces sounds of similar timbre.
  - etc., etc.

**Melodic Channeling**

- Experiment by Deutsch
- Stimuli were two sequences alternating between the right and left ears.
- Listeners perceive two smooth sequences by grouping the sounds by similarity in pitch.
- This demonstrates the perceptual heuristic that sounds with the same frequency come from the same source, which is usually true in the environment.
Deutch's Melodic Channeling Experiment

http://tinyurl.com/6pc66or

(a) Stimuli presented to listener's left ear (blue) and right ear (red). N.B.: notes presented to each ear jump up and down.

(b) What listener hears. Although the notes in each ear jump up and down, listener perceived smooth sequence of notes.

Questions

• Name some Gestalt principles that apply to auditory scene analysis.
• Name some characteristics of sound by which sounds from the environment are grouped.

Wessel's Timbre Illusion

http://tinyurl.com/7vfy2ao

Auditory Grouping

• Proximity in time - sounds that occur in rapid succession usually come from the same source
• This principle was illustrated in auditory streaming
• Good continuation - sounds that stay constant or change smoothly are usually from the same source
Good Continuation

- Demonstration by Bregman et al.
  - Tones were presented interrupted by gaps of silence or by noise
  - In the silence condition, listeners perceived that the sound stopped during the gaps
  - In the noise condition, the perception was that the sound continued behind the noise

Auditory Grouping: Effects of Experience

- Experiment by Dowling
  - Used two interleaved melodies ("Three Blind Mice" and "Mary Had a Little Lamb")
  - Listeners reported hearing a meaningless jumble of notes
  - But listeners who were told to listen for the melodies were able to hear them by using melody schema

Schema-driven Grouping

http://tinyurl.com/7sbqgqm
Questions

• What gestalt principle does Bregman’s demo show?

• What does the melody schema experiment by Dowling imply about the role of top-down and bottom-up processing in the organization of sound?

Direct & Indirect Sound

• Direct sound - reaches the listeners’ ears straight from the source

• Indirect sound - is reflected off of environmental surfaces and then to the listener

• Amount & timing of echoes can give a sense of the distance to a sound source

Interactions Between Vision and Sound

• Visual capture or the ventriloquist effect - an observer perceives the sound as coming from the seen location rather than the source for the sound

• Experiment by Sekuler et al.
  • Balls moving without sound appeared to move past each other
  • Balls with an added “click” appeared to collide
Questions

- Where are you most likely to experience indirect sound?
- What does Sekular's experiment with the bouncing/non-bouncing balls demonstrate?