CS 414 – Multimedia Systems Design

Lecture 2 – Multimedia Characterization & Auditory Perception

Klara Nahrstedt

Spring 2008

Administrative

- Form Groups for MPs
  - Deadline January 20 to email TA

Multimedia Characteristics

- Multimedia means an integration of continuous media (audio/video) and discrete media (text/graphics/images) through which digital information can be conveyed to the user in an appropriate way

Media Classification

- Perception Medium
- Representation Medium
- Presentation Medium
- Storage Medium
- Transmission Medium
- Information Exchange Medium

Medium Definition

- Representation Values
  - Continuous representation values (e.g., electron-magnetic waves
  - Discrete representation values (e.g., character of text)
- Representation Space
  - Visual representation space (e.g., paper, screen)
  - Acoustic representation space (e.g., stereo)

Representation Dimensions in Representation Spaces

- Spatial dimensions
  - Two dimensional (2D graphics)
  - Three dimensions (3D image, holography)
- Temporal dimensions
  - Time independent/Discrete Media (document)
  - Time dependent/Continuous Media (movie)

Data Stream Characterization

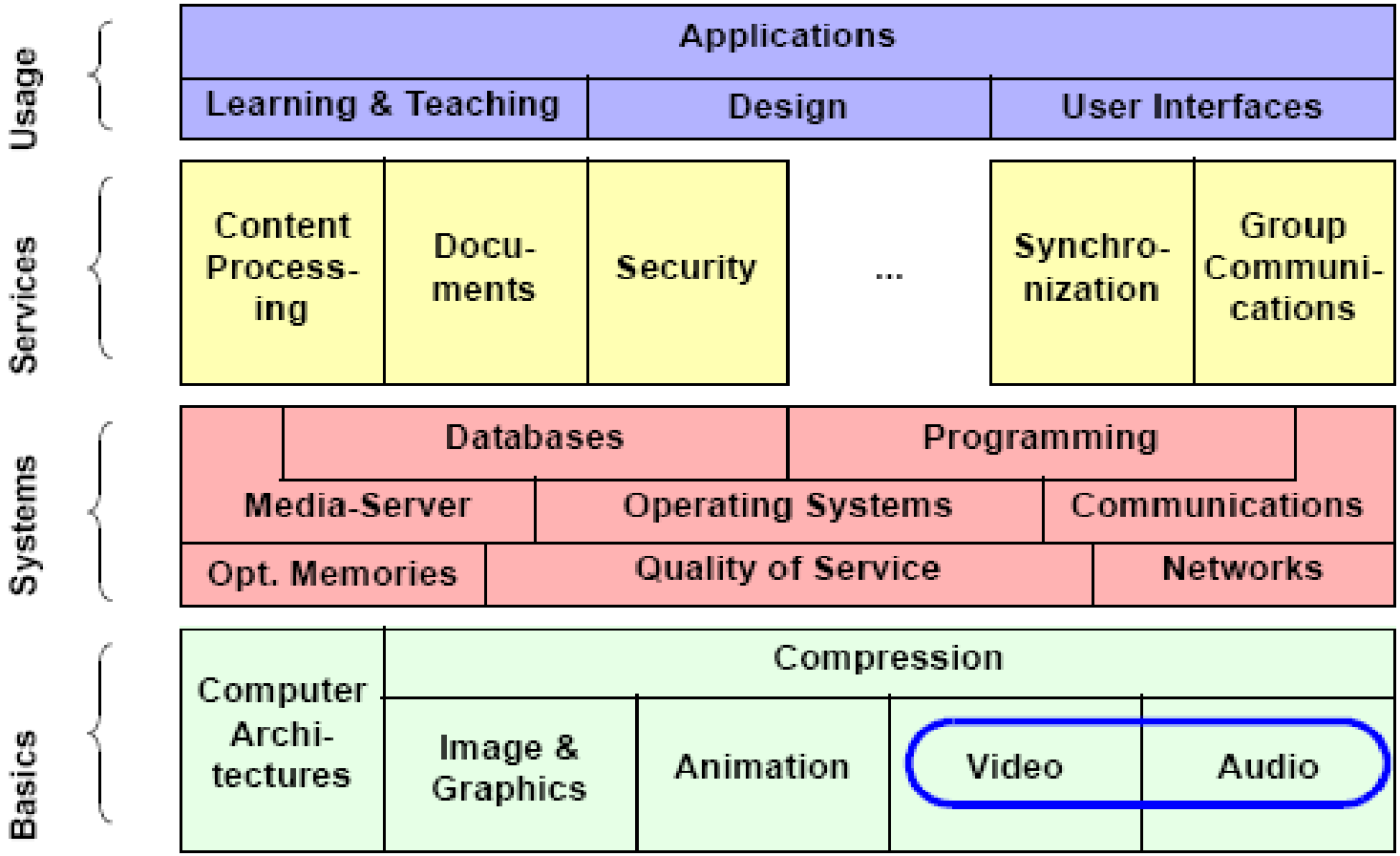
- Asynchronous transmission
  - No timely restrictions
- Synchronous transmission
  - Defines maximum end-to-end delay on each packet of data stream
- Isochronous transmission
  - Defines maximum and minimum end-to-end delay on each packet of data stream

Data Stream Characteristics

- Time Intervals
  - Strongly periodic data streams
  - Weakly periodic data streams
  - Aperiodic data streams
- Packet Size
  - Strongly regular data streams
  - Weakly regular data streams
  - Irregular data streams
- Continuity
  - Continuous data streams
  - Discrete data streams

Logical Data Units

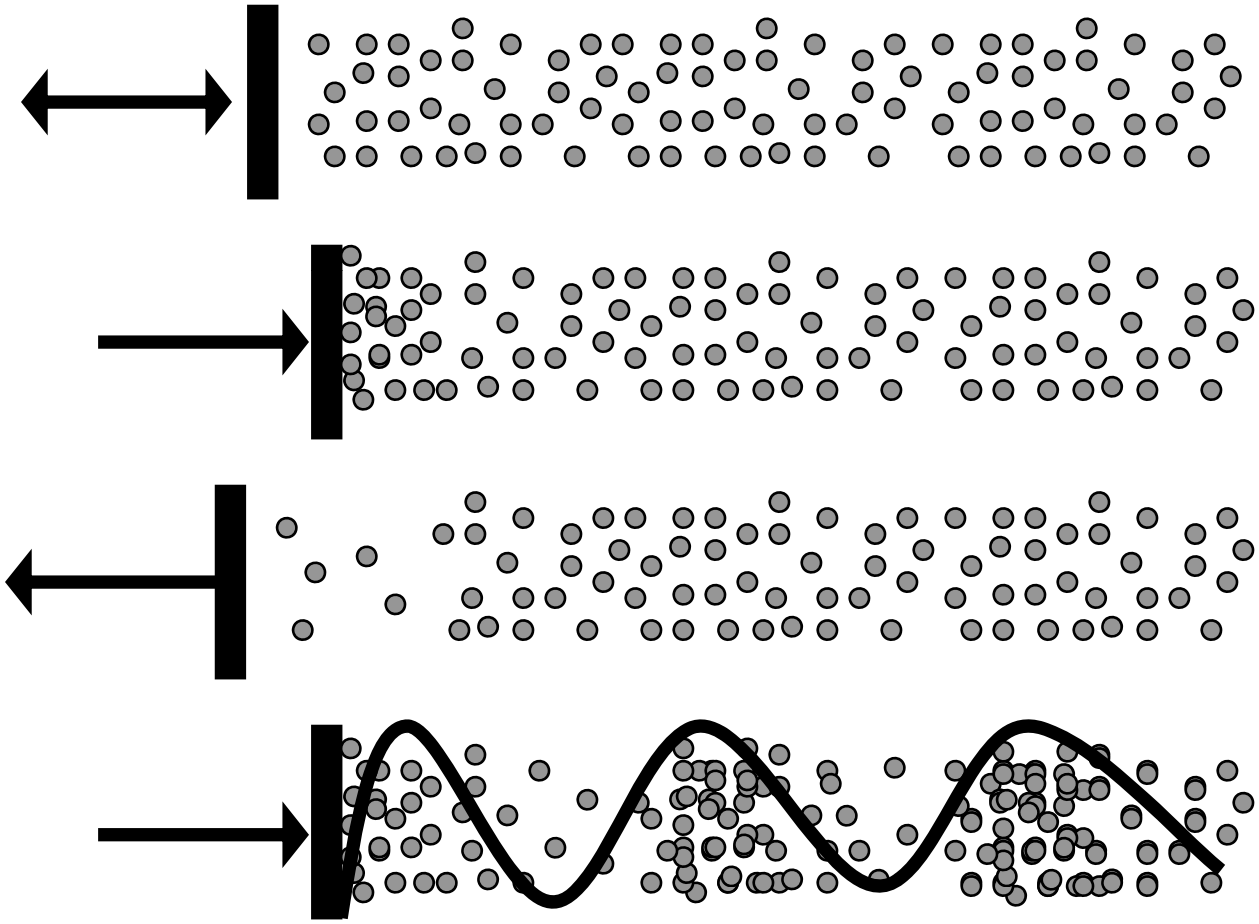
- Continuous media consist of time dependent sequence of LDUs
  - Example:
    - Symphony consists of sentences; sentences consist of notes, notes are sequences of samples
- Granularity of LDUs
  - pixel, image, video scene, movie
- Duration of LDUs
  - open LDU vs closed LDU


Usage
Services
Systems
Basics
Applications
Learning & Teaching
Design
User Interfaces
Content Processing
Documents
Security
...
Synchronization
Group Communications
Databases
Programming
Media-Server
Operating Systems
Communications
Opt. Memories
Quality of Service
Networks
Compression
Computer Architectures
Image & Graphics
Animation
Video
Audio

Auditory Perception

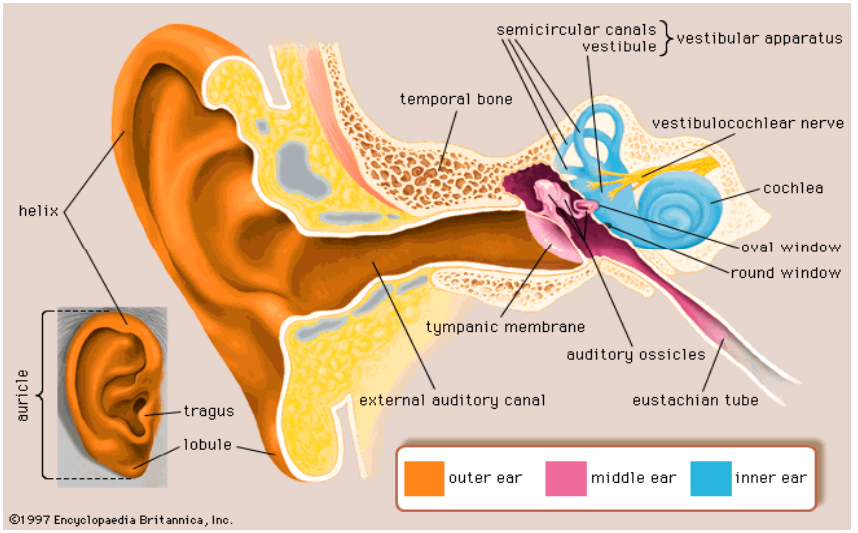
- Sound – physical phenomenon caused by vibration of material
- These vibrations trigger pressure wave fluctuations in the air
- Wave forms

Changes in Air Pressure

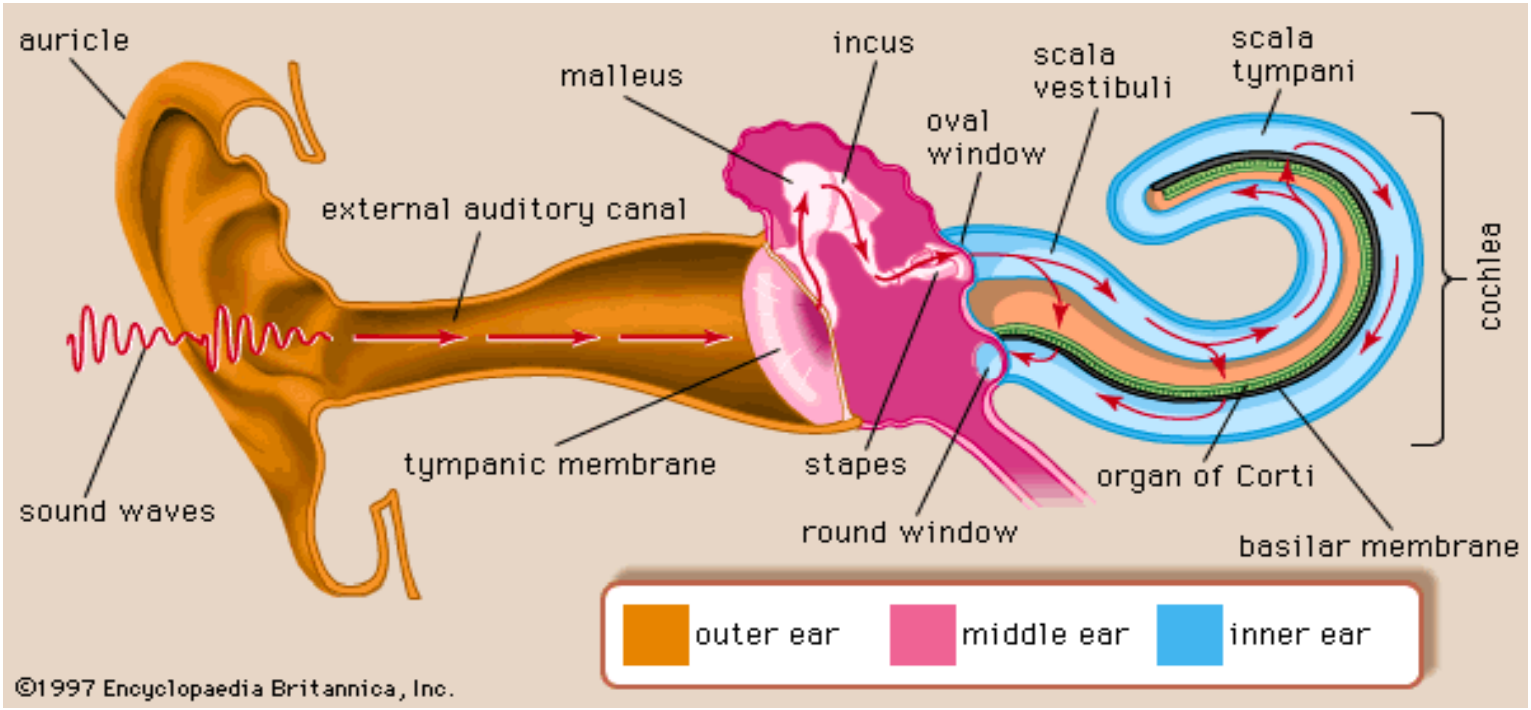


Auditory System

- Ears, parts of brain, and neural pathways
- Changes in pressure move hair-like fibers within the inner ear
- Movements result in electrical impulses sent to the brain

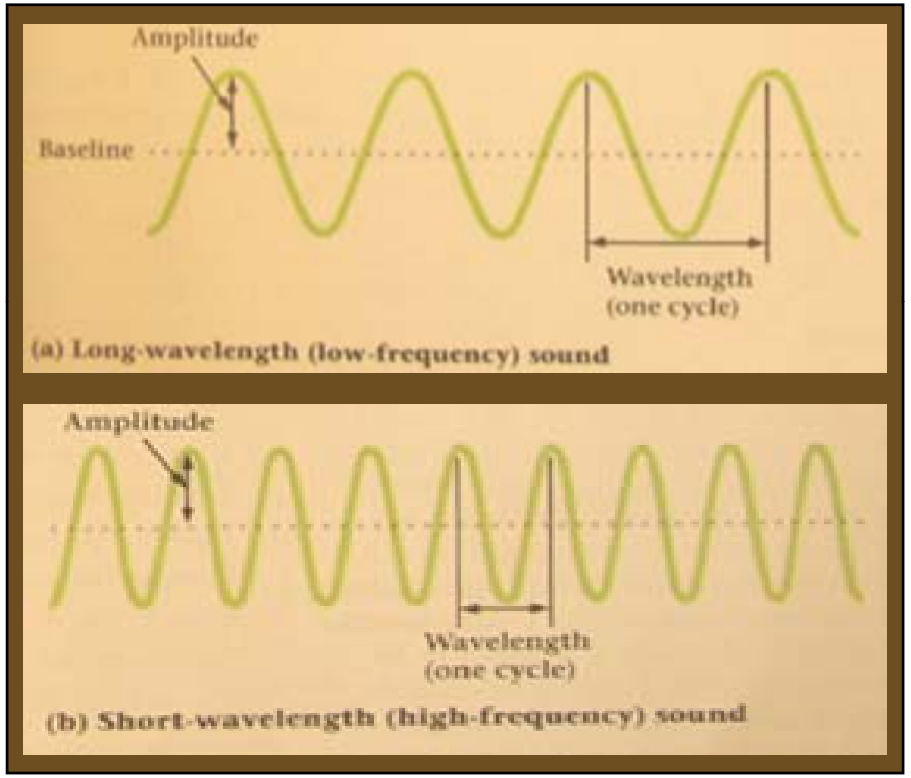


Process of Hearing (Transduction)



Physical Dimensions

- Amplitude
  - height of a cycle
  - relates to loudness
- Wavelength (*w*)
  - distance between peaks
- Frequency (λ)
  - cycles per second
  - relates to pitch
  - λw = velocity
- Most sounds mix many frequencies & amplitudes



Sound is repetitive changes in air pressure over time

Sound Perception and Psychoacoustics

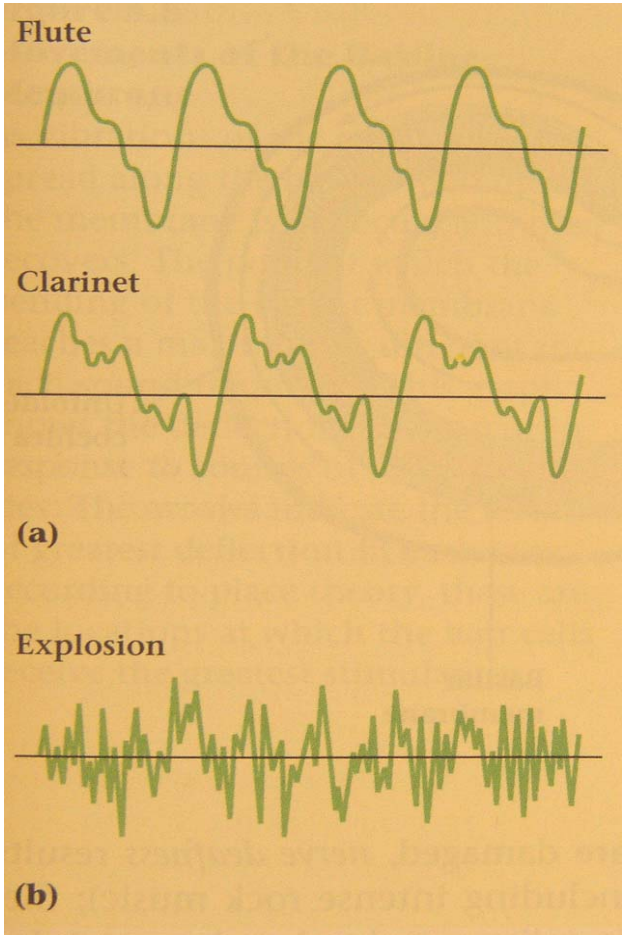
- Psychoacoustics
  - Study the correlation between the physics of acoustical stimuli and hearing sensations
  - Experiments data and models are useful for audio codec
- Modeling human hearing mechanisms
  - Allows to reduce the data rate while keeping distortion from being audible

Psychological Dimensions

- Loudness
  - higher amplitude results in louder sounds
  - measured in decibels (db), 0 db represents hearing threshold
- Pitch
  - higher frequencies perceived as higher pitch
  - hear sounds in 20 Hz to 20,000 Hz range

Psychological Dimensions (cont.)

- Timbre (tam-bre)
  - complex patterns added to the lowest, or *fundamental*, frequency of a sound, referred to as *spectra*
  - spectra enable us to distinguish musical instruments
- Multiples of fundamental frequency give music
- Multiples of unrelated frequencies give noise



Sound Intensity

- *Intensity* (I) of a wave is the rate at which sound energy flows through a unit area (A) perpendicular to the direction of travel

$$I = \frac{1}{A}\frac{\Delta E}{\Delta t} = \frac{P}{A}$$

P measured in watts (W), A measured in m^2

- *Threshold of hearing* is at 10^{-12} W/m^2
- *Threshold of pain* is at 1 W/m^2

Decibel Scale

- Describes intensity relative to threshold of hearing based on multiples of 10

$$dB = 10 \log \frac{I}{I_0}$$

I_0 is reference level = 10^{-12} W/m^2

Decibels of Everyday Sounds

Sound	Decibels
Rustling leaves	**10**
Whisper	**30**
Ambient office noise	**45**
Conversation	**60**
Auto traffic	**80**
Concert	**120**
Jet motor	***140***
Spacecraft launch	***180***

Interpretation of Decibel Scale

- 0 dB = threshold of hearing (TOH)
- 10 dB = 10 times more intense than TOH
- 20 dB = 100 times more intense than TOH
- 30 dB = 1000 times more intense than TOH

- An increase in 10 dB means that the intensity of the sound increases by a factor of 10

- If a sound is 10^x times more intense than another, then it has a sound level that is 10*x more decibels than the less intense sound

Loudness from Multiple Sources

- Use energy combination equation

$$L = 10 \log(10^{\frac{L1}{10}} + 10^{\frac{L2}{10}} + \ldots + 10^{\frac{LN}{10}})$$

where L_1, L_2, …, L_n are in dB

Exercises

- Show that the threshold of hearing is at 0 dB
- Show that the threshold of pain is at 120 dB

- Suppose an electric fan produces an intensity of 40 dB. How many times more intense is the sound of a conversation if it produces an intensity of 60 dB?

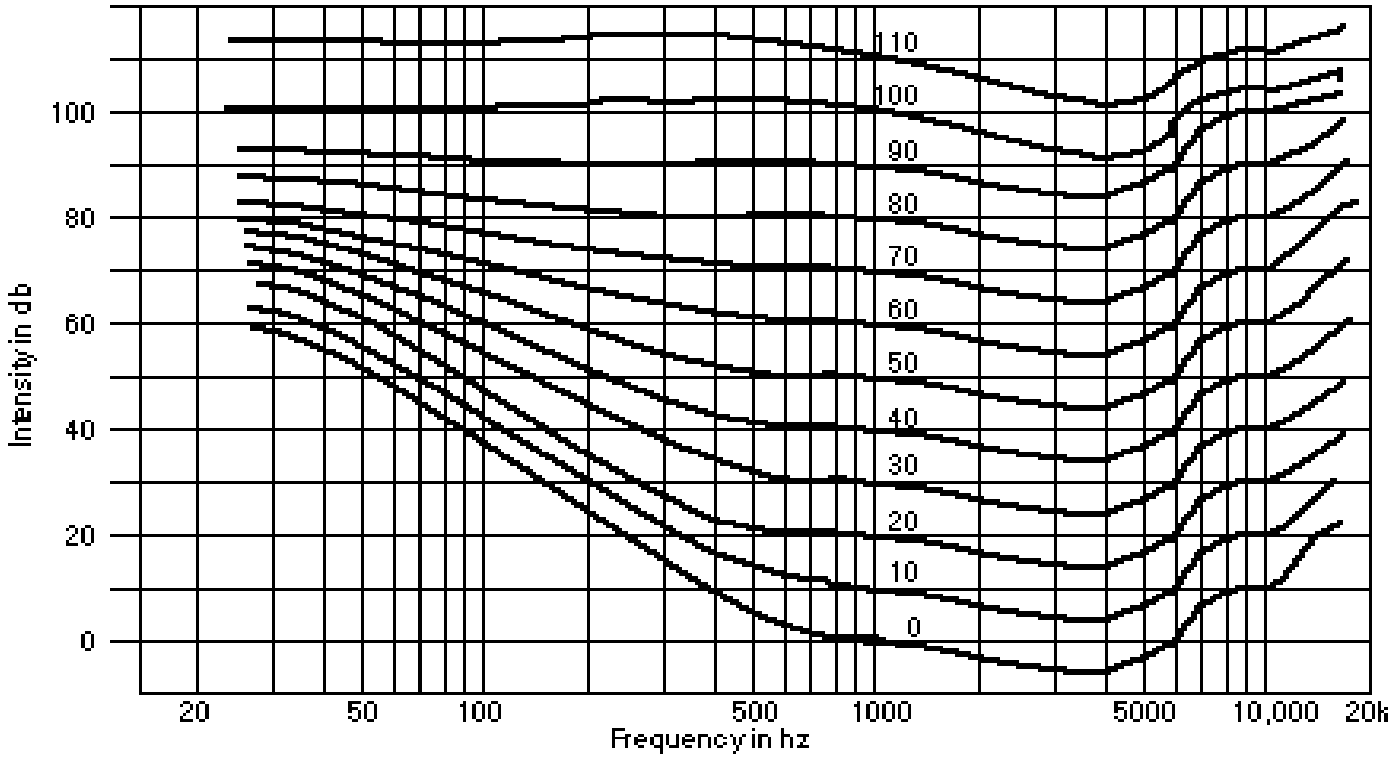
- One guitar produces 45 dB while another produces 50 dB. What is the dB reading when both are played?

- If you double the physical intensity of a sound, how many more decibels is the resulting sound?

Loudness and Pitch

- More sensitive to loudness at mid frequencies than at other frequencies
  - intermediate frequencies at [500hz, 5000hz]
- Perceived loudness of a sound changes based on the frequency of that sound
  - basilar membrane reacts more to intermediate frequencies than other frequencies

Fletcher-Munson Contours



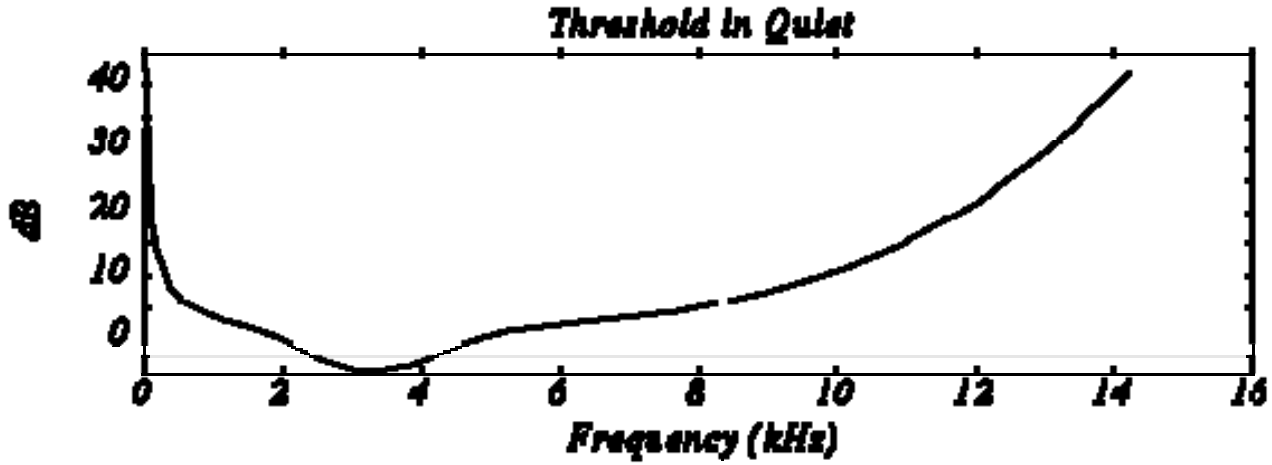
Each contour represents an equal perceived sound

Masking

- Perception of one sound interferes with another
- Frequency masking
- Temporal masking

Frequency Masking

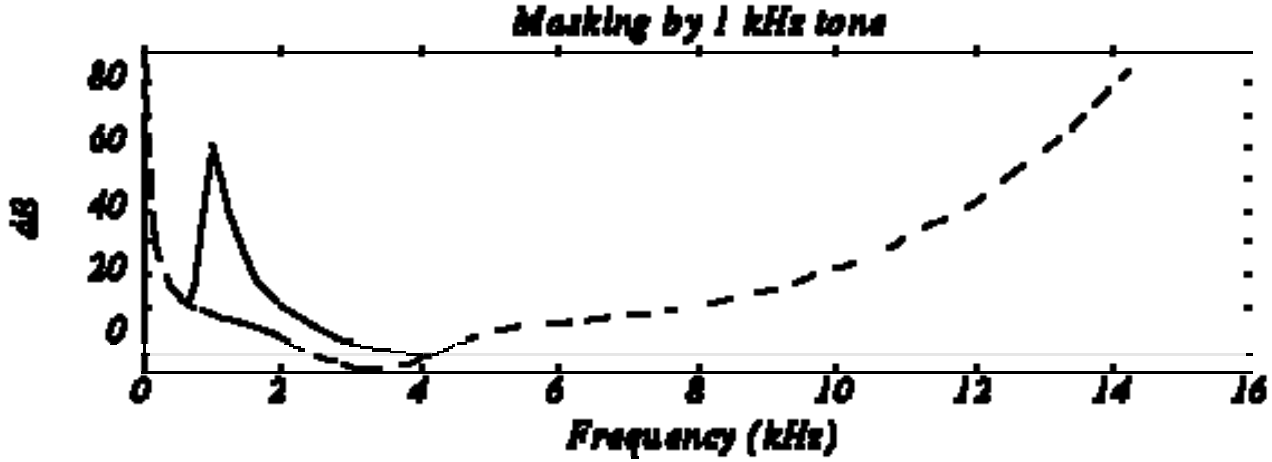
- Louder, lower frequency sounds tend to mask weaker, higher frequency sounds



From http://www.cs.sfu.ca/CourseCentral/365/

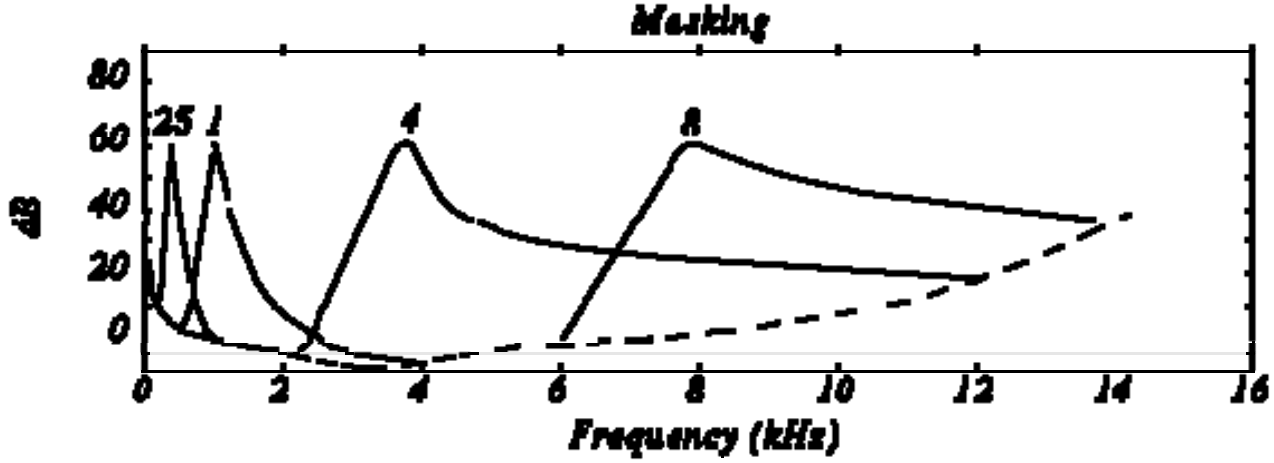
Frequency Masking

- Louder, lower frequency sounds tend to mask weaker, higher frequency sounds



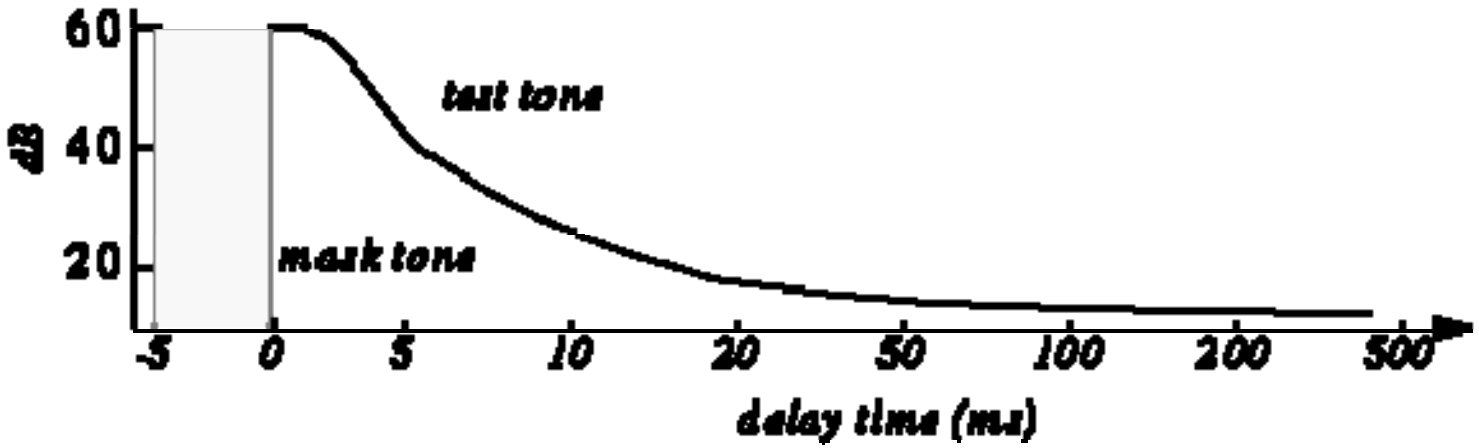
Frequency Masking

- Louder, lower frequency sounds tend to mask weaker, higher frequency sounds

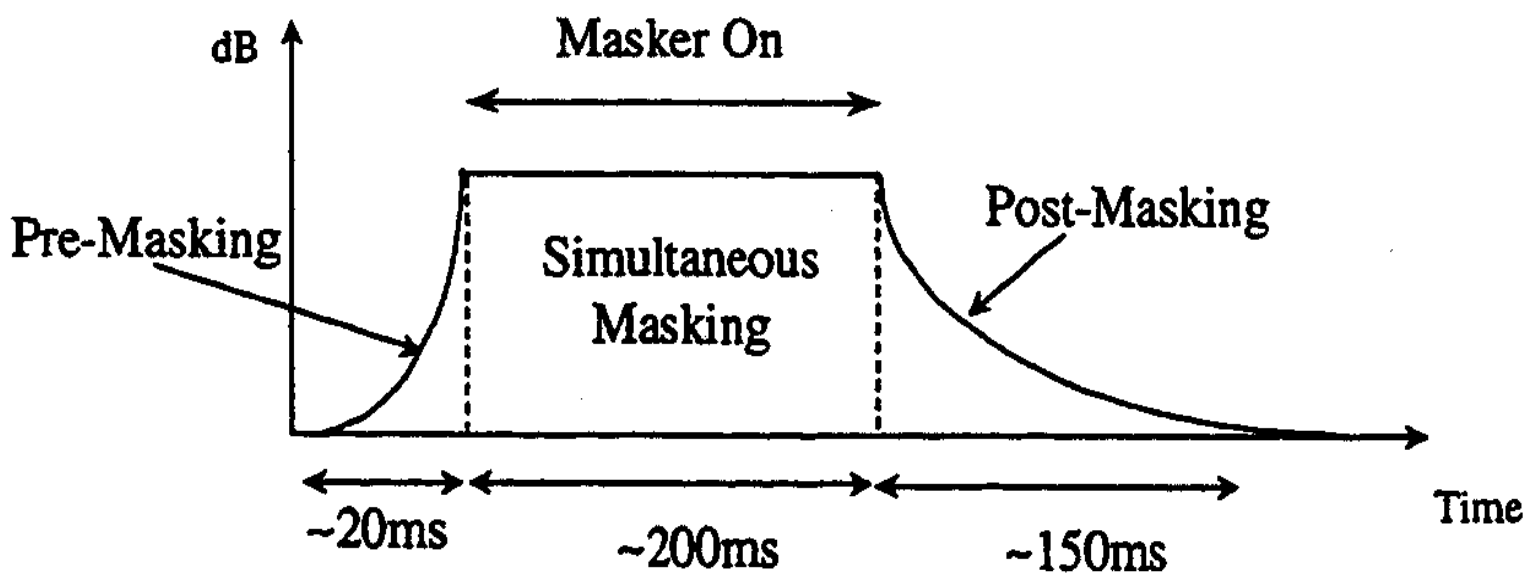


Temporal Masking

- When exposed to a loud sound, the human ear contracts slightly to protect delicate structures
- Causes louder sounds to overpower weaker sounds just *before* and just *after* it



Temporal Masking



Summary

- Auditory Perception is very important for understanding digital audio representation
- Psychoacoustic is used in MP3 audio compression