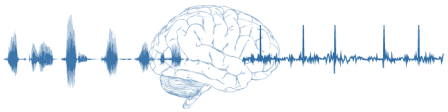


Primer on Auditory Processing

Mounya Elhilali
Department of Electrical & Computer Engineering
Johns Hopkins University
mounya@jhu.edu

601.467/667 Introduction to Human Language Technology

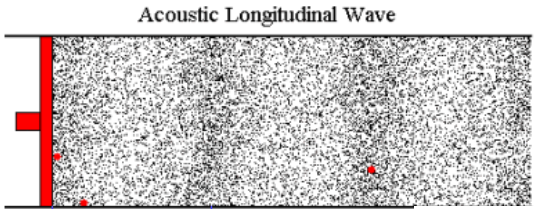
1

Speech as waves

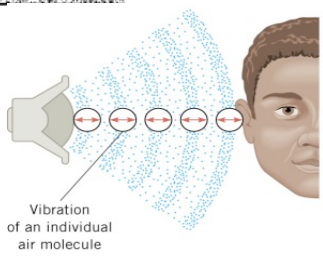
2

Sound is a wave

- Sound is a mechanical wave caused by a vibrating source
- The vibrating source that causes the matter around it to move

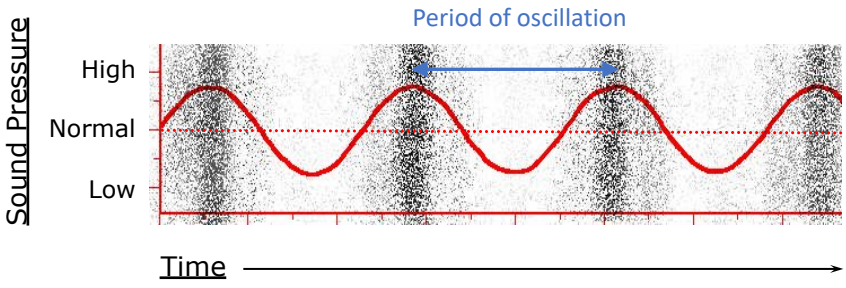


- No sound is produced in a vacuum
  - Matter (air, water, earth) must be present
- Individual air molecules do not move the wave. A given molecule vibrates forth about a fixed location.



3

Sound waves

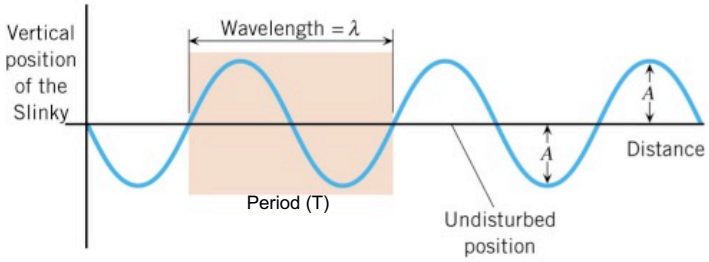


- Motion air particles do not travel, they oscillate around a point in space
- The rate of oscillation is called frequency (f)
  - ✓ denoted in cycles per second (cps) or hertz (Hz).

5

5

Physical Dimensions of Sound



Amplitude
- Height of a cycle

Frequency (F)
- Cycles per second

Wavelength (λ)
- Distance traveled by one cycle

6

6

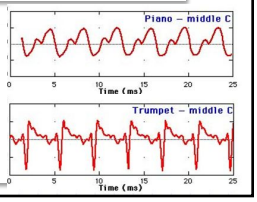
Perceptual dimensions of Sound

Physical Properties of Sound — **Perceptual Dimensions**

Amplitude/Intensity ⟷ Loudness

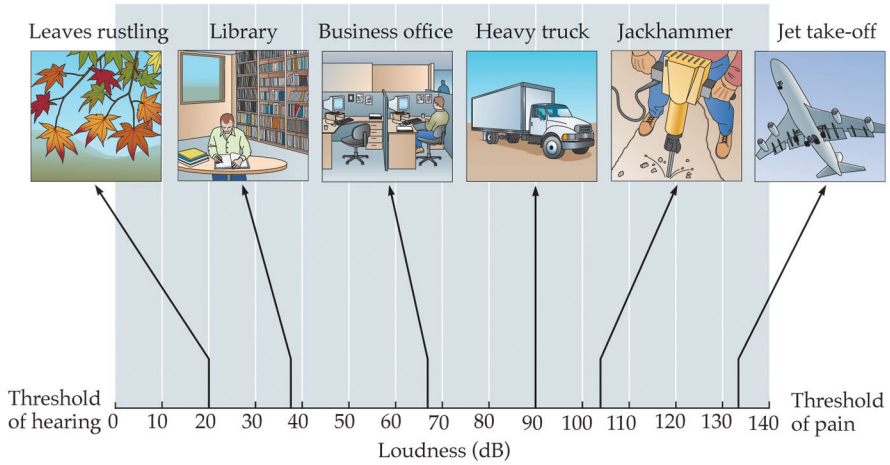
Frequency ⟷ Pitch

Complexity (frequency content & time) ⟷ Timbre



7

Sounds in the environment

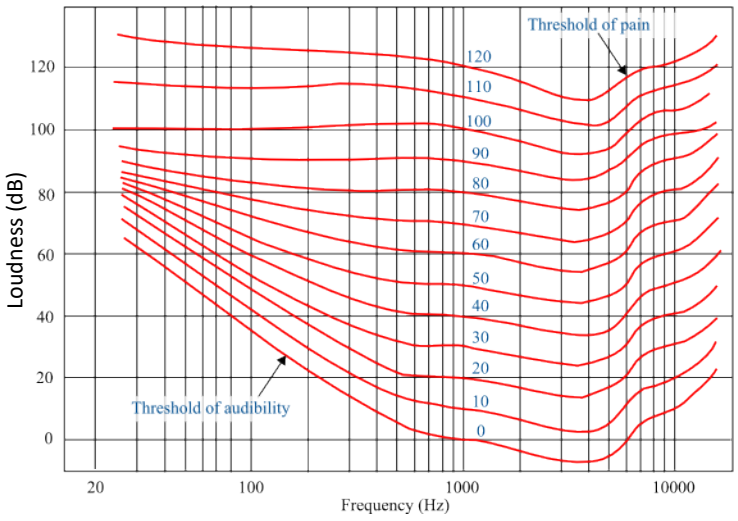


Note: Listening to loud music will gradually damage your hearing!

8

8

Equal Loudness Curves/Contours

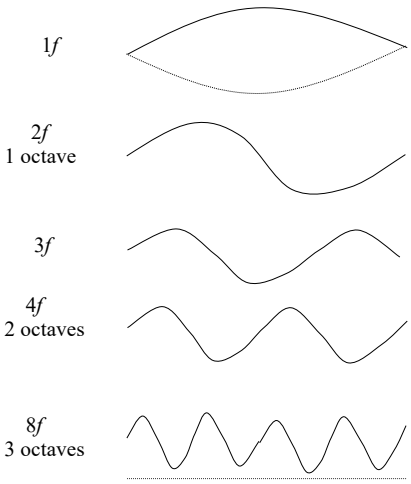


Each contour represents equally-perceived tones

9

9

Pitch



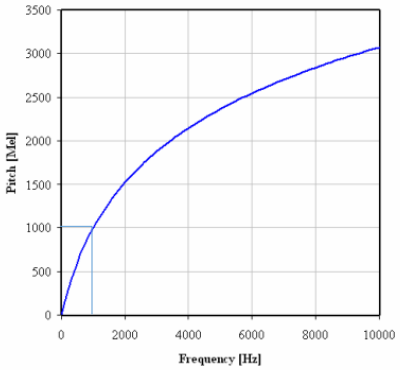
- At first approximation, the pitch of a simple periodic signal is determined by its frequency.
- Most oscillators (guitar string, vocal chords) naturally oscillate at a fundamental frequency (F_0) as well as its integer multiples (called harmonics/partials/overtones).
- The pitch of a *complex* period signal is often determined by its *fundamental* frequency (F_0)

10

10

a Pitch scale

- Perceptual scale of pitch: mel scale
  - *How far in frequency do we have to be in order to feel a tone as doubled in pitch?*



➡ It's a relative scale, based on pitch comparisons

$$m = 2595 \log_{10}\left(1 + \frac{f}{700}\right)$$

✓ Mel-scaling is used in signal processing to build filters that approximate human pitch perception (MFCC)

11

11

Masking

- Hearing phenomenon
  - When the perception of one sound is affected by presence of *another* sound
  - one sound being *masked* by another
- Term masking is used to describe effects of noise and interference in sound perception
- We experience masking everyday

12

12

Masking

13

13

How do we perceive sounds?

17

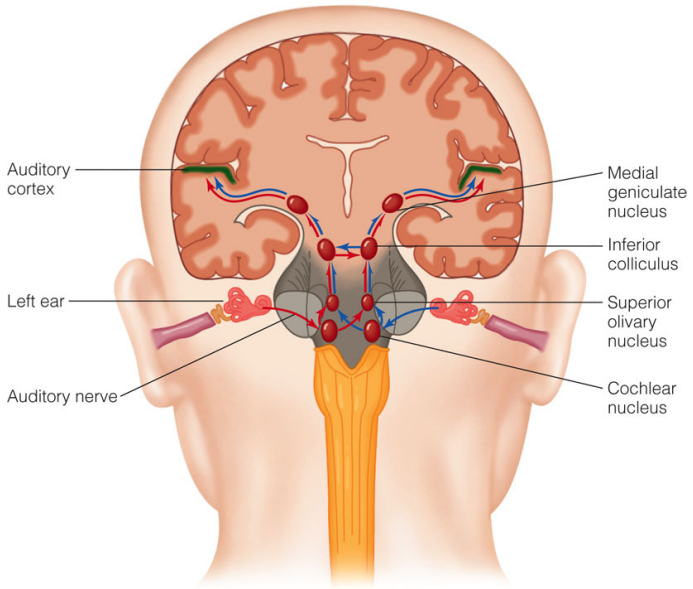
The auditory system

- Two major components in the auditory system
  - The peripheral auditory organs (the ear)
    - Converts sounds pressure into mechanical vibration patterns, which then are transformed into neural firings
  - The auditory nervous system (the brain)
    - Extracts perceptual information in various stages

18

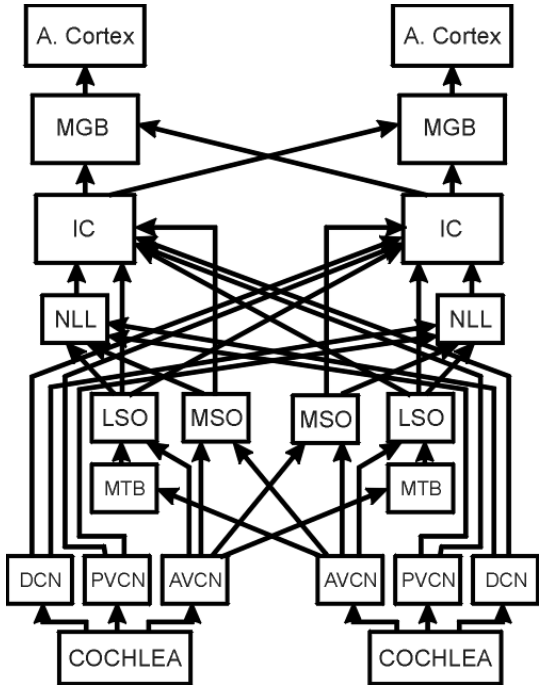
18

Auditory Pathway



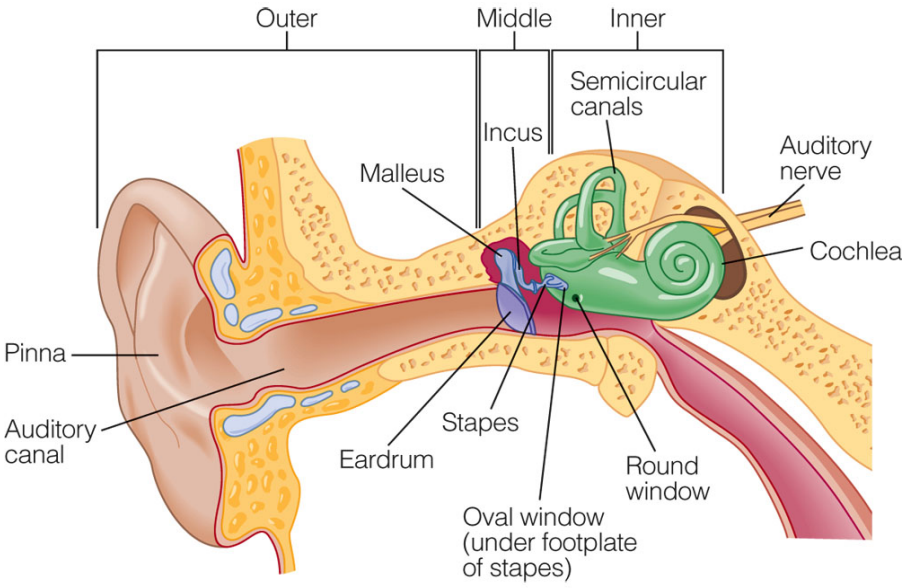
19

19



20

20

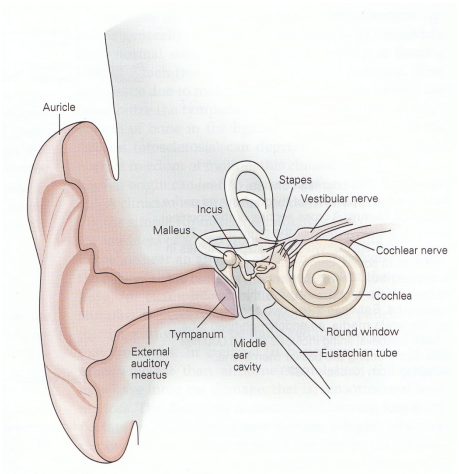


21

21

The ear

- The ear is the organ of hearing
- It changes sound pressure waves from the outside world into a signal of nerve impulses sent to the brain.
- It consists of 3 components:
  - Outer ear
  - Middle ear
  - Inner ear

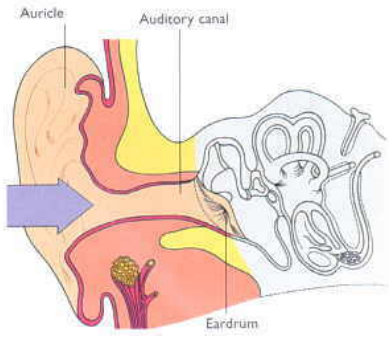


22

22

Organ of hearing
outer ear

- The external ear plays the role of an acoustic antenna,
- It diffracts and focuses sound waves (pinna), while the ear canal acts as a resonator => amplifies sounds in 2-5 kHz range
- The end of the canal has an eardrum which vibrates with sound

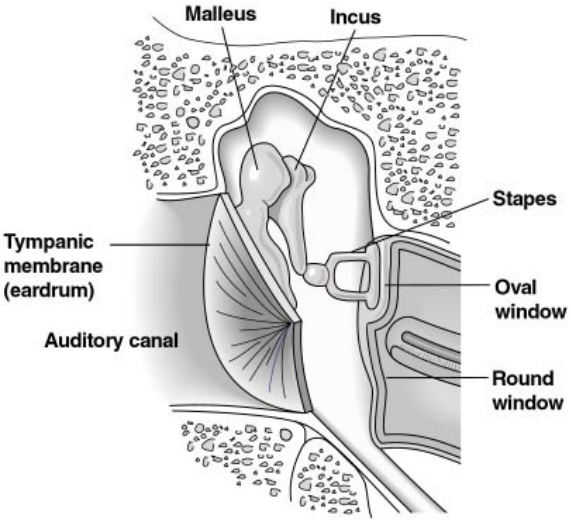


23

23

Organ of hearing
middle ear

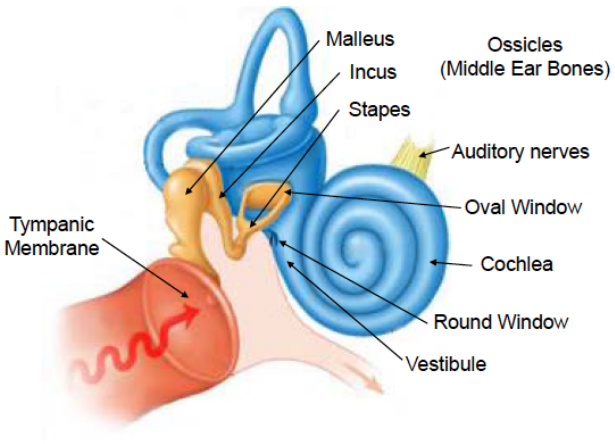
- Eardrum (or tympanic membrane) vibrations cause mechanical motion of the small bones of the middle ear (malleus, incus & stapes) [3 smallest bones in the human body]
- The middle ear acts as an impedance adapter to adjust energy difference between air environment and fluid environment



24

Organ of hearing inner ear

- Cochlea translates physical vibrations into electrical signals for the brain to process
- Cochlea acts a frequency analyzer of sound signals

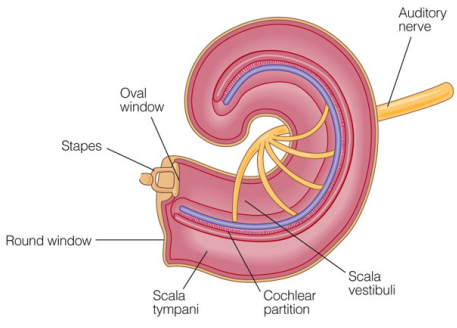


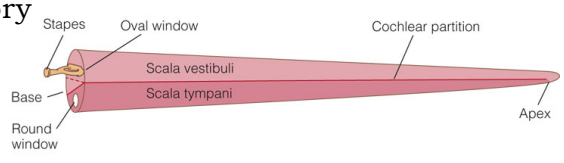
25

25

The Cochlea

- The cochlea is the inner ear organ that converts sound waves into neural signals.
- The neural signals are passed to the brain via the auditory nerve.

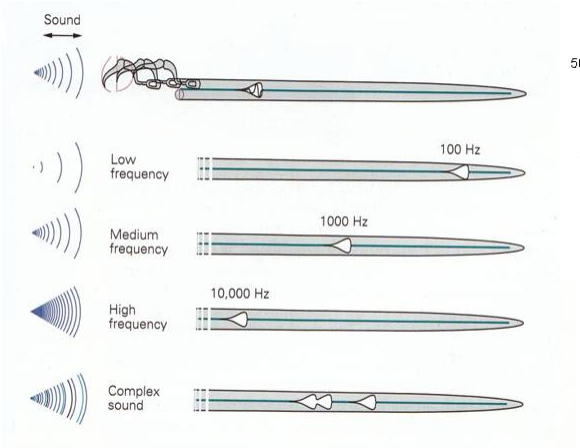


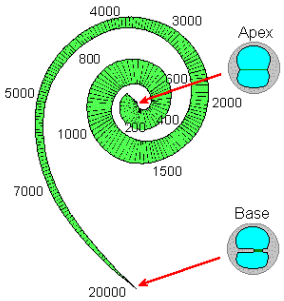


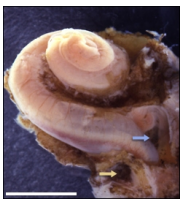
26

26

Cochlea as frequency analyzer

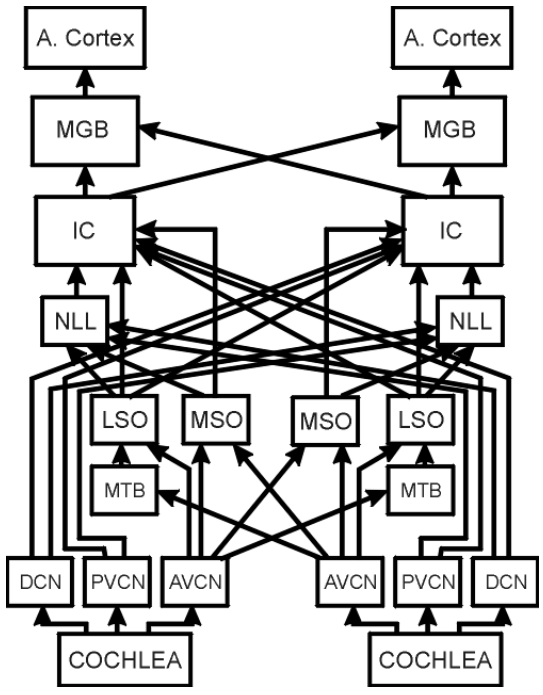






28

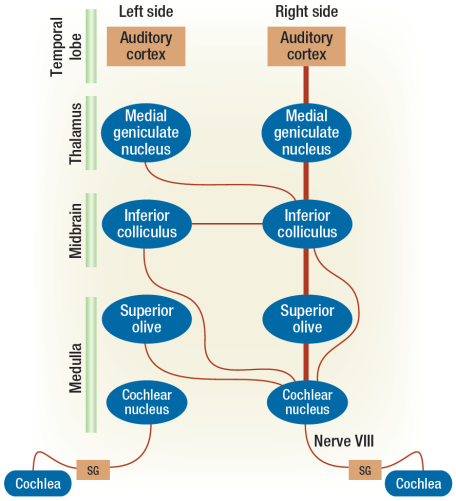
28



32

32

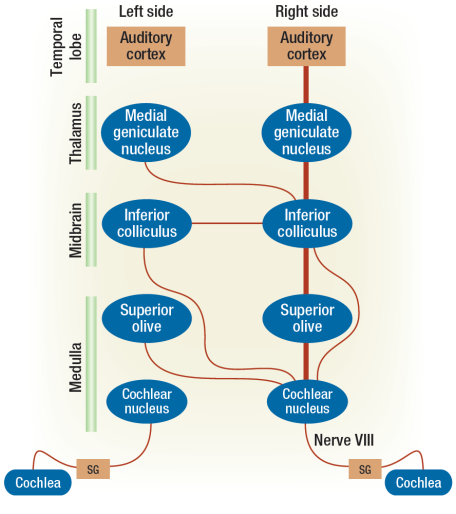
Ascending pathway



- Very complex. Just some major pathways shown.
- Extensive binaural interactions
- General principle:
  - ✓ Increasing complexity of responses (like vision, touch)

33

Ascending pathway



FUNCTION

Identify and process complex sounds

Principle relay to cortex

Form full spatial map

Locate sound sources in space

Start sound feature processing

Sound sensor / periphery

34

Tonotopy

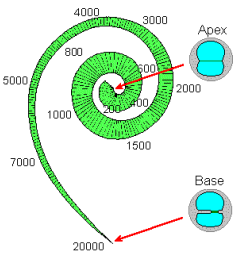
- Tonotopic map:
  - topographic organization (spatial arrangement) of where sound is processed
  - Derived from Greek tono/topos = place of tones
- Most nuclei along auditory pathway from cochlea to A1 are tonotopically organized (inherit *cochleotopy* from periphery)

35

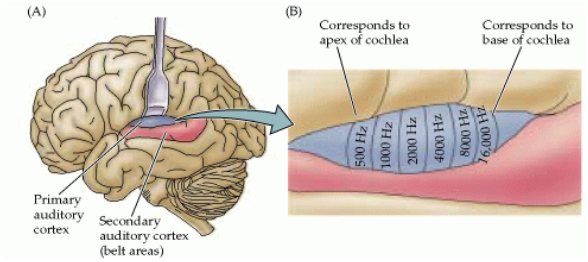
35

Auditory tonotopy

- Adjacent cells in A1 form a frequency-map, similar to the one observed in the cochlea.



Cochlea

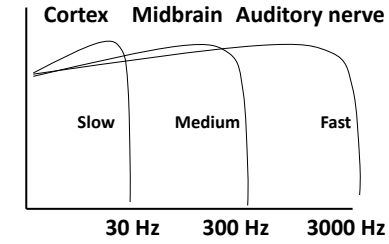


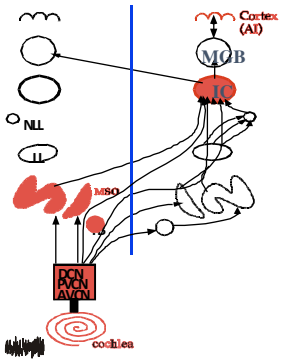
A1

36

36

Encoding speech modulation beyond the cochlea

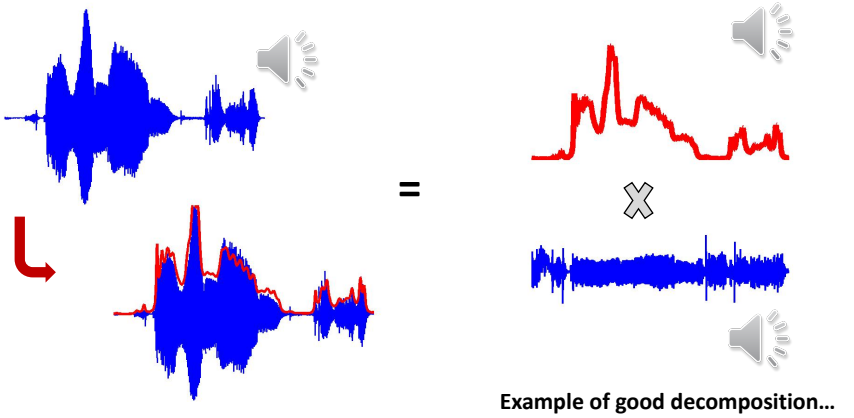




39

Speech carries information at multiple levels

- Any speech signal can be separated into two signals.



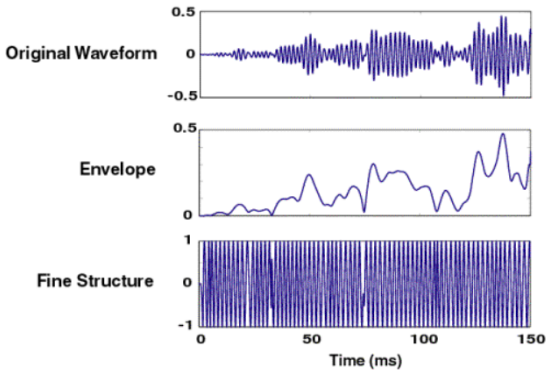
Example of good decomposition…
A non-trivial task

40

40

Speech carries information at multiple levels

- Any speech signal can be separated into two signals.
  - The envelope is the amplitude of the sound
  - The fine structure is the detailed waveform, without its envelope



41

41