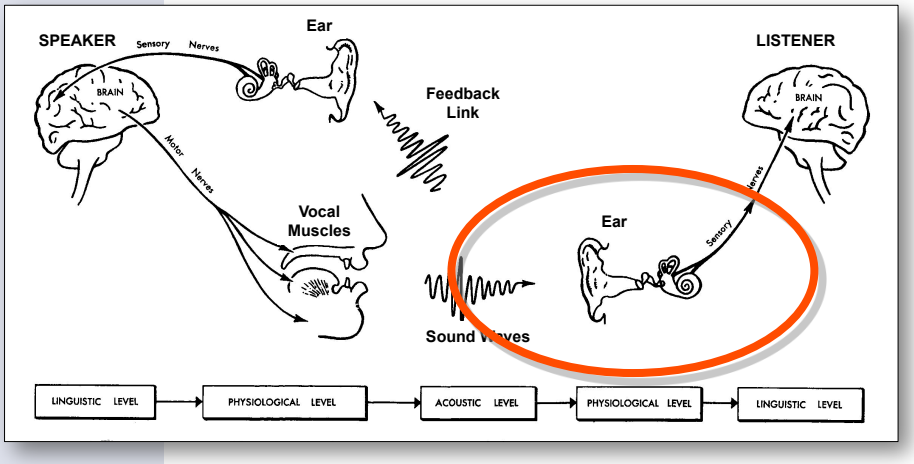
COM3502-4502-6502
SPEECH PROCESSING

Lecture 4
Hearing

The University Of Sheffield.

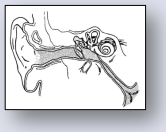
The Speech Chain



Taken from: Denes, P. B., & Pinson, E. N. (1973). ***The Speech Chain: The Physics and Biology of Spoken Language***: New York: Anchor Press.

The University Of Sheffield.

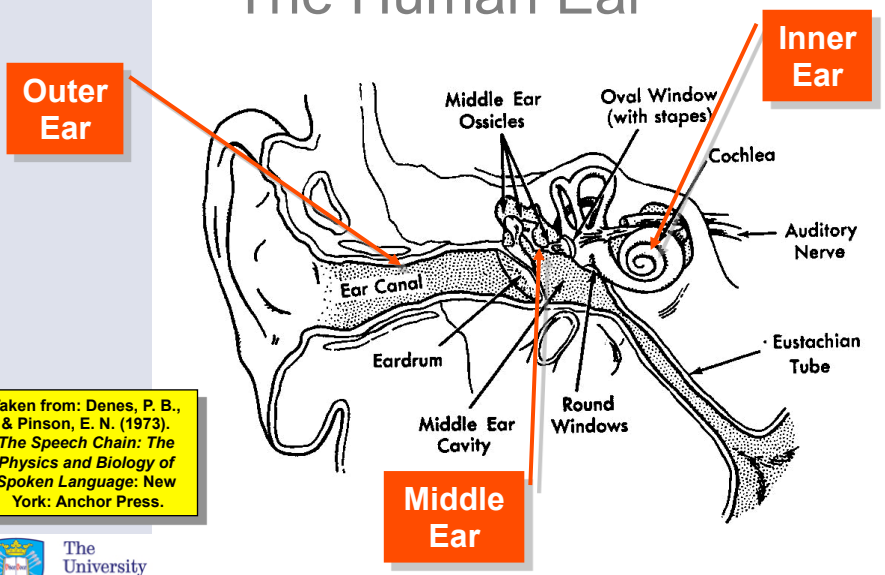
The Human Ear



- The auditory system has evolved for …
  - acoustic sensing
  - sound localisation
  - communication
- The ear is more *general purpose* than the articulatory system
- Its main function is frequency analysis
- The main percepts are …
  - pitch
  - loudness
  - timbre

The University Of Sheffield.

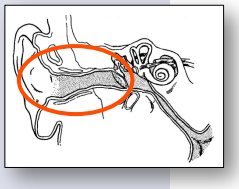
The Human Ear



Taken from: Denes, P. B., & Pinson, E. N. (1973). *The Speech Chain: The Physics and Biology of Spoken Language*: New York: Anchor Press.

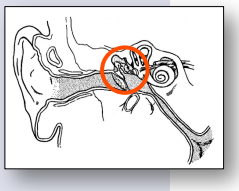
The University Of Sheffield.

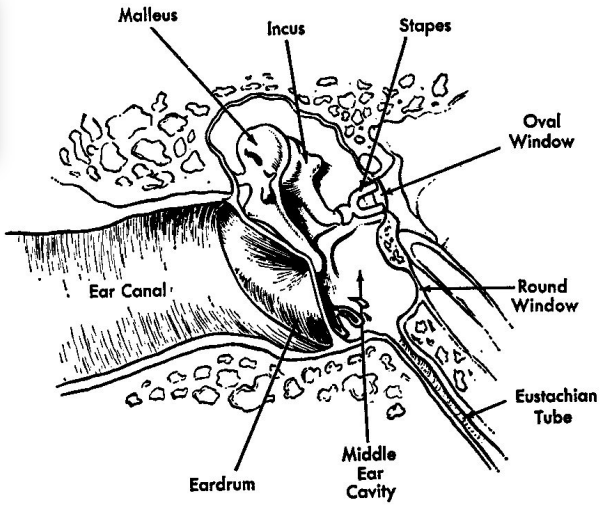
The Outer Ear



- The 'pinna' protects the entrance to the ear canal, and its shape makes it directionally sensitive at high frequencies
- The external canal - 'meatus' - is a tube (*~2.7 cm long, ~0.7 cm in diameter*) that leads from the pinna to the middle ear
- The meatus terminates at the cone shaped 'tympanic membrane' (*eardrum*)
- Sound waves entering the ear impinge upon the eardrum and cause it to vibrate

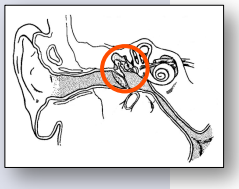
The Middle Ear





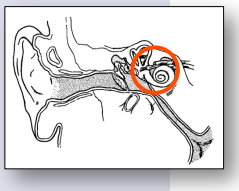
Taken from: Denes, P. B., & Pinson, E. N. (1973). *The Speech Chain: The Physics and Biology of Spoken Language*: New York: Anchor Press.

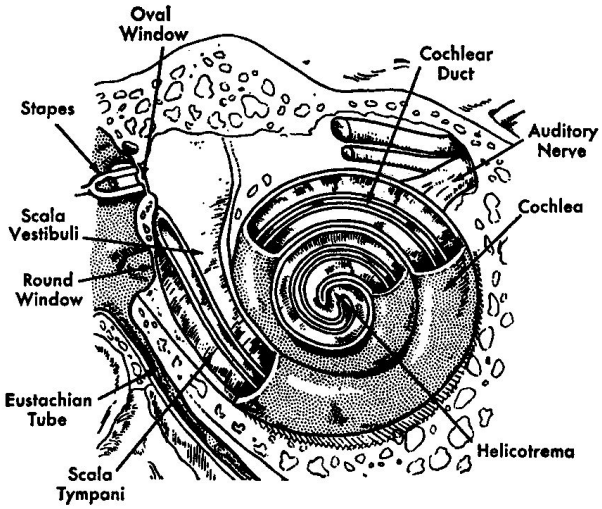
The Middle Ear



- The middle ear transforms the vibration of the eardrum into oscillations of the liquid in the inner ear by vibrating the 'oval window'
- The necessary impedance matching (*between air and liquid*) is achieved by a group of bones - the 'ossicles' - acting as a system of mechanical levers
- The pressure at the oval window is ~35x greater than that arriving at the eardrum
- This mechanical amplification allows us to hear sounds 1000x weaker than otherwise
- Muscles attached to the ossicles protect the inner ear from potential damage due to high sound levels

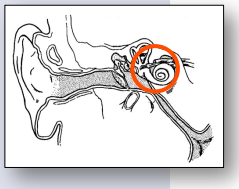
The Inner Ear





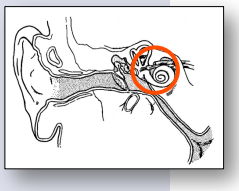
Taken from: Denes, P. B., & Pinson, E. N. (1973). *The Speech Chain: The Physics and Biology of Spoken Language*: New York: Anchor Press.

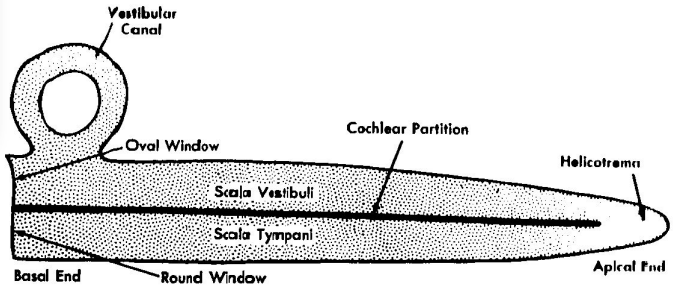
The Inner Ear



- The transformation from mechanical vibrations to electrical nerve impulses ('*neural transduction*') takes place in the snail-like structure of the 'cochlea'
- The cochlea is ~35 mm long and is filled with a colourless liquid called 'perilymph'
- The cochlea is divided into two regions along its length by a membrane structure called the 'cochlea partition' (*a channel filled with a liquid called 'endolymph'*)
- The cochlea partition is bounded by …
  - the 'basilar membrane'
  - 'Reissner's membrane'

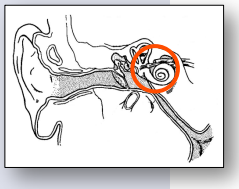
The Cochlea (*unwound*)

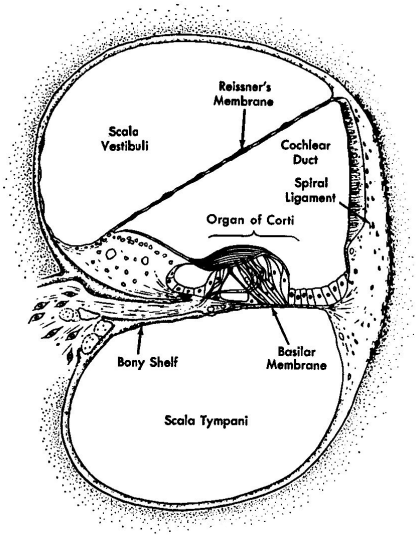




Taken from: Denes, P. B., & Pinson, E. N. (1973). *The Speech Chain: The Physics and Biology of Spoken Language*: New York: Anchor Press.

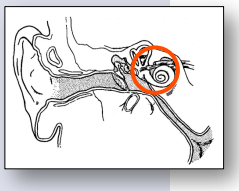
Cochlear Cross-Section





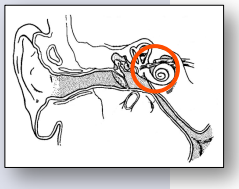
Taken from: Denes, P. B., & Pinson, E. N. (1973). *The Speech Chain: The Physics and Biology of Spoken Language*: New York: Anchor Press.

Action of the Cochlea



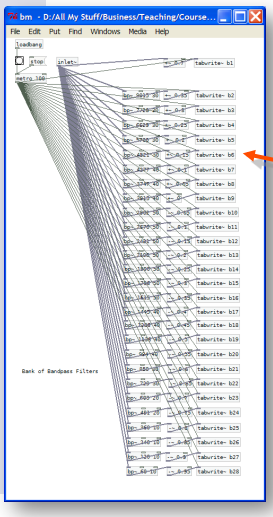
- The mechanical properties of the basilar membrane determine how the cochlea responds to sound
- Vibrations entering at the oval window set up travelling waves which lead to peaks of energy at different places along the cochlea depending on the frequency
- The vibration is nearest the oval window for high-frequency sounds
- The 'organ of corti' transform the mechanical movements into electrochemical pulses by bending the outer hair cells (*of which there are ~25,000*)
- These actions are equivalent to a bank of 'bandpass filters'

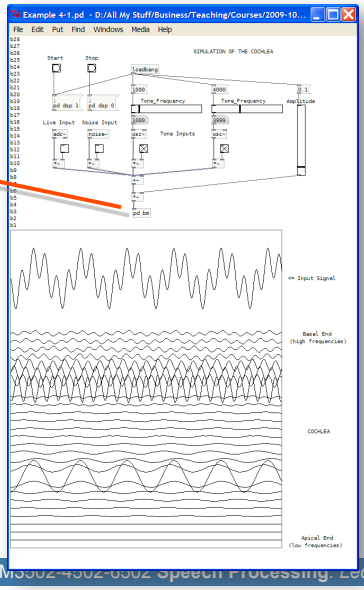
Action of the Cochlea



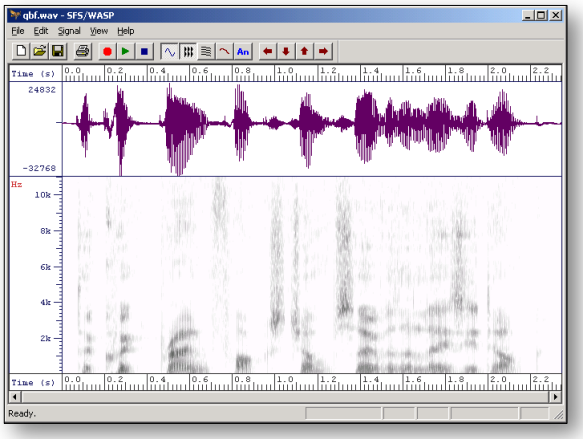
https://youtu.be/dyenMluFaUw

Demo: Cochlear Simulation



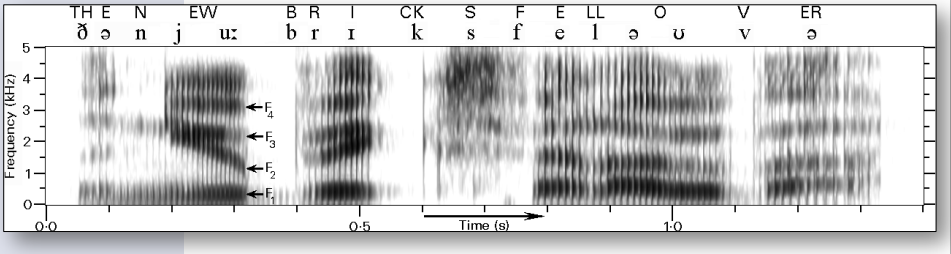


Demo: Real-Time Spectrogram



http://www.phon.ucl.ac.uk/resource/sfs/rtgram/

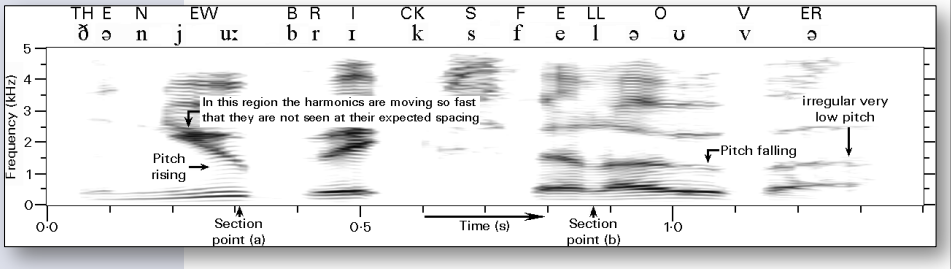
Wideband Speech Spectrogram



- Good time resolution
- Poor frequency selectivity

Taken from: Holmes, J. N., & Holmes, W. (2002). *Speech Synthesis and Recognition*: Taylor & Francis.

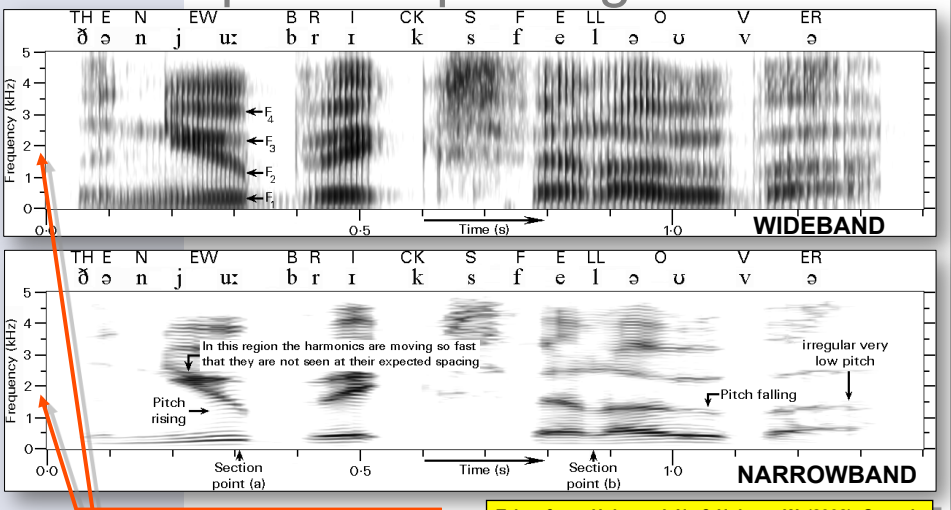
Narrowband Speech Spectrogram



- Good frequency selectivity
- Poor time resolution

Taken from: Holmes, J. N., & Holmes, W. (2002). *Speech Synthesis and Recognition*: Taylor & Francis.

Speech Spectrograms



- same frequency scales
- different frequency selectivities

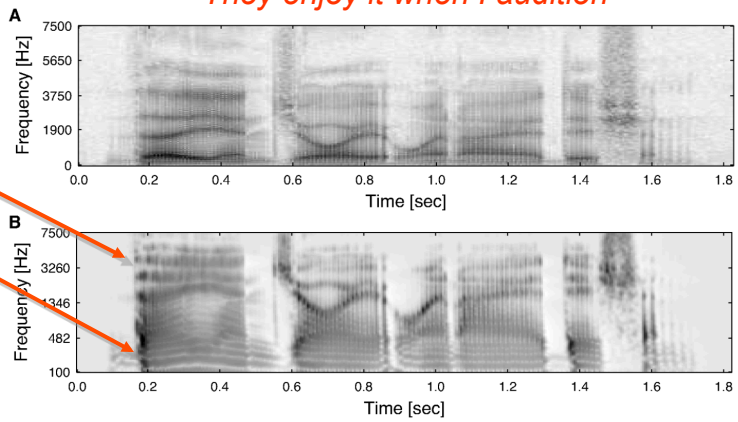
Taken from: Holmes, J. N., & Holmes, W. (2002). *Speech Synthesis and Recognition*: Taylor & Francis.

Spectrogram (A) *vs*. Cochleagram (B)

"They enjoy it when I audition"

Wideband spectrogram at high frequencies

Narrowband spectrogram at low frequencies



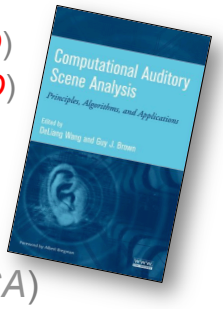
Note difference in frequency scales: spectrogram is linear, cochleagram is approximately logarithmic

Question …

What's the point of having two ears?

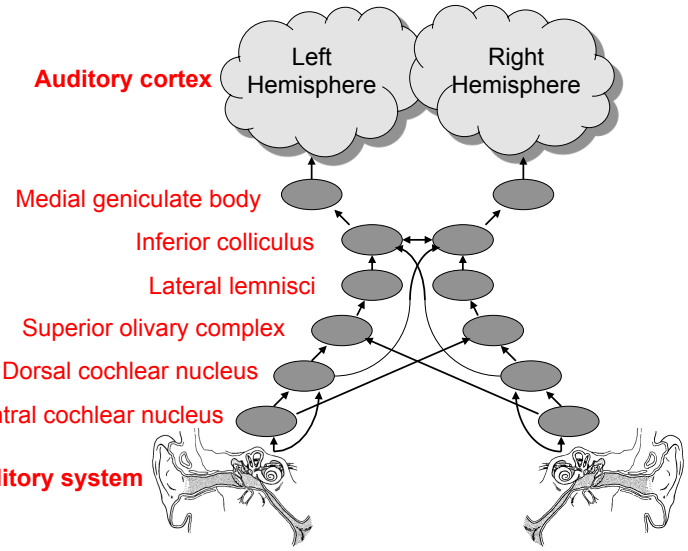
Binaural Auditory Processing

- Sound localisation
  - to direct (*visual*) attention
- Possible mechanism:
  - inter-aural time differences (*ITD*)
  - inter-aural level differences (*ILD*)
- Listening in complex acoustic environments
- 'Auditory Scene Analysis' (*ASA*)



The University Of Sheffield.

Auditory Pathways

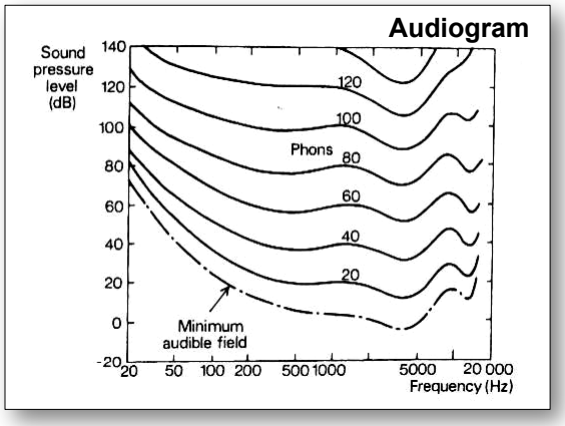


The University Of Sheffield.

Sensitivity/Selectivity of Hearing

- The human ear varies in sensitivity and selectivity for sounds with different …
  - loudness
  - frequency components
- These effects are studied in 'auditory psychophysics'

Loudness Sensitivity

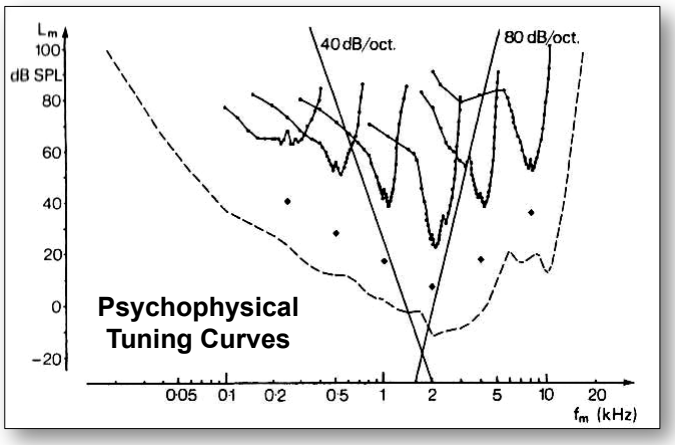


Taken from: Holmes, J. N., & Holmes, W. (2002). *Speech Synthesis and Recognition*: Taylor & Francis.

Frequency Selectivity

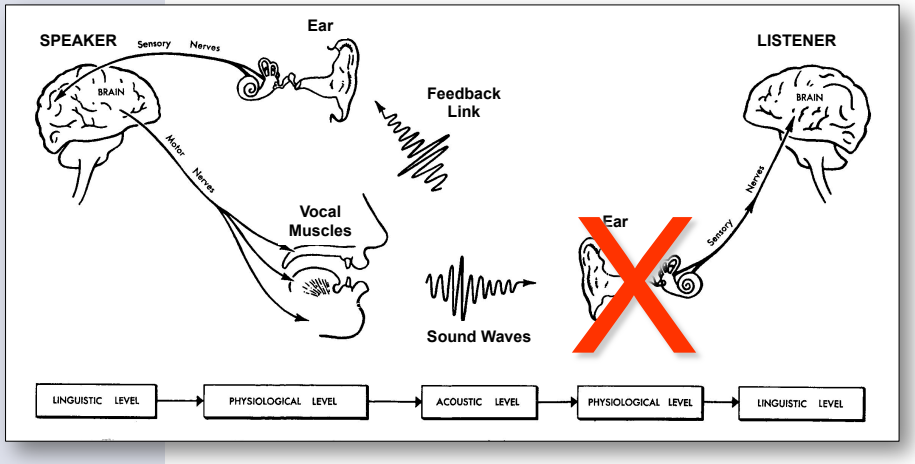
- The frequency range of human hearing lies between ~20 Hz and ~20,000 Hz
- The ability to discriminate energy varies according to frequency ...
  - 0.2 Hz at 100 Hz
  - 1.5 Hz at 2000 Hz
  - 30 Hz at 12,000 Hz
- Low frequency sounds can mask higher frequency sounds because of the overlap between 'auditory filters'
- The bandwidth over which masking operates is termed the 'critical band'
- The shapes of the auditory filters are revealed by deriving 'psychophysical tuning curves'

Frequency Selectivity

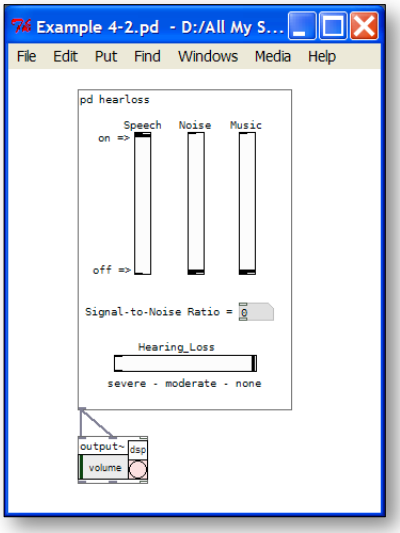


Taken from: Holmes, J. N., & Holmes, W. (2002). *Speech Synthesis and Recognition*: Taylor & Francis.

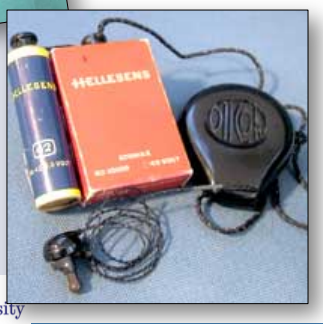
Hearing Impairment



Demo: Hearing Loss



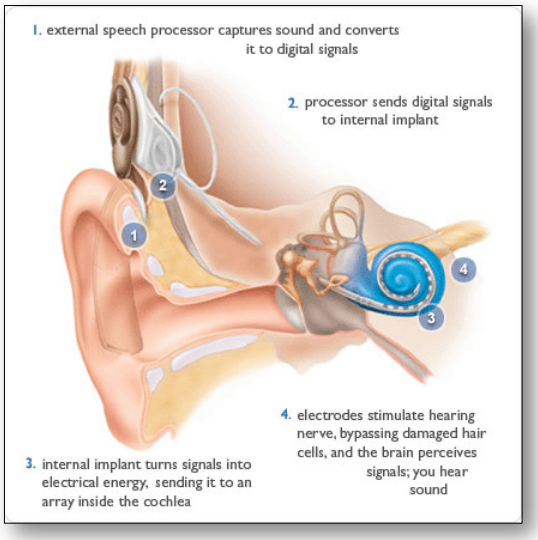
Hearing Aids



The University Of Sheffield.

Cochlear Implants

Cochlear™ Hear now. And always



The University Of Sheffield.

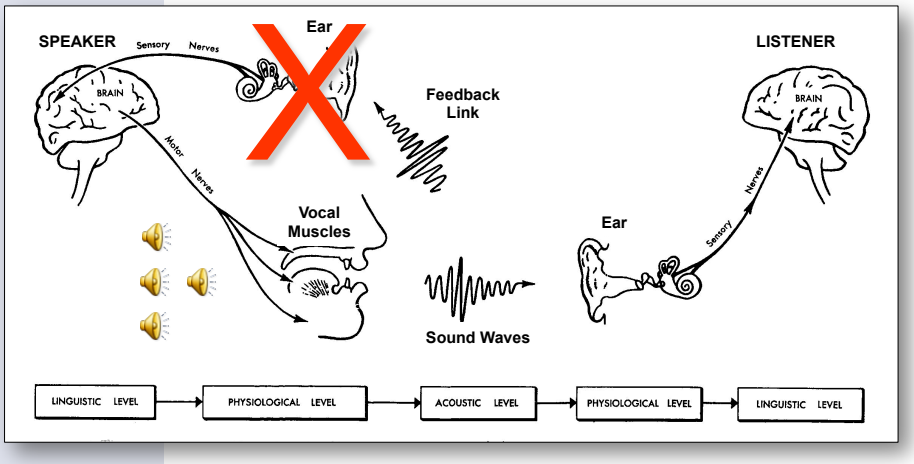
Cochlear Implants

http://www.actiononhearingloss.org.uk

Cochlear™ Hear now. And always

https://youtu.be/HTzTtlVnHRM

Speaking Also Depends on Hearing



This lecture has covered …

- Outer, middle and inner ear
- Action of the cochlea
- Spectrographic analysis
- Binaural processing
- Auditory psychophysics
- Hearing impairment

Any Questions ?

Next time …

The Nature of Speech

The University Of Sheffield.