Acoustics, signals & systems for audiology

Week 6

Basic Psychoacoustic Phenomena

Basic Psychoacoustic Phenomena

- Absolute threshold
- Loudness
- Frequency selectivity
  - Auditory filters
  - Representation of spectral shape (timbre)
- Pitch
- Binaural hearing
  - Localisation of sound
  - Masking release

What is psychoacoustics?

- Psychophysics
  - Mapping the relationship between the physical/objective and perceptual/subjective world.
- Psychoacoustics — psychophysics of sounds.
- How does the loudness of a sound relates to its intensity?
  - loudness depends not only on intensity but also on frequency content
- Changing the fundamental frequency of a periodic sound from 100 to 200 Hz will not lead to the same perceived musical interval as a change from 800 Hz to 900 Hz.

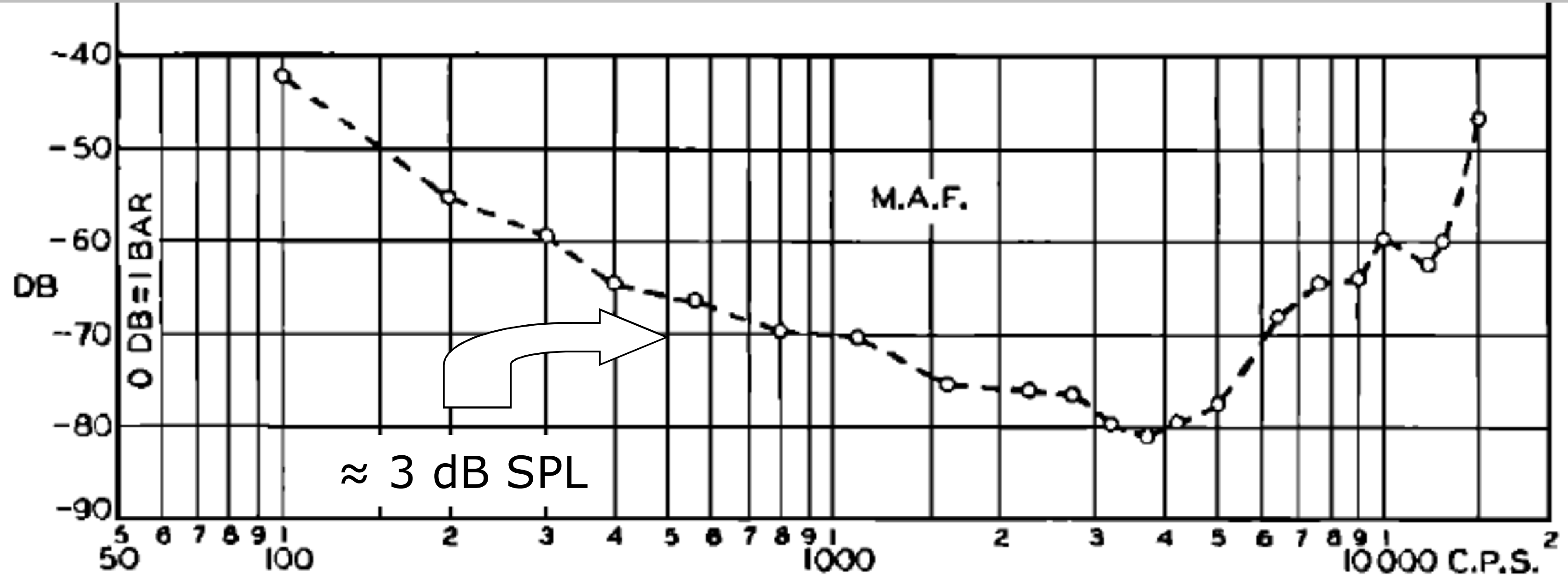
100-200 Hz 800-900 Hz 800-1600Hz

What is psychoacoustics?

- Terminology: Objective vs. subjective
  - intensity (W/m^2, Pa, dB SPL) vs. loudness
  - periodic/aperiodic vs. buzziness/noisiness
  - fundamental frequency (Hz) vs. pitch
  - spectral envelope/shape vs. timbre/quality/colour
- Much of psychoacoustics concerns abilities to …
  - detect
    - many HI people and CI users need higher levels to detect sounds
  - discriminate
    - many HI people and CI users need greater differences between stimuli to hear a difference between them
  - but limits on detectability and discriminability can also provide crucial data for developing models of auditory perception even in normal listeners

Sivian & White (1933) JASA

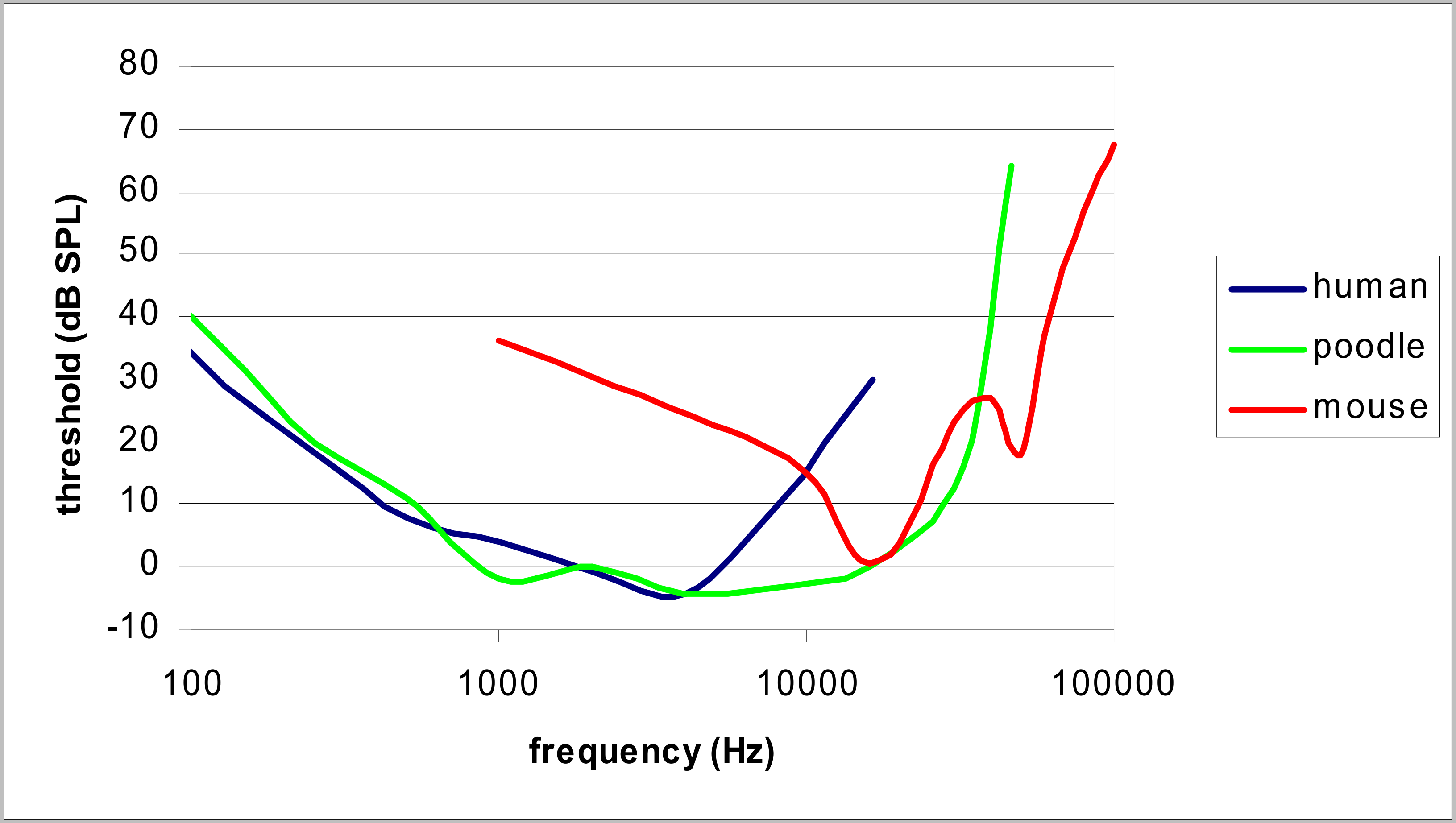
Sivian & White 1933



≈ 3 dB SPL

FIG. 3. *Monaural M.A.F., group A.*

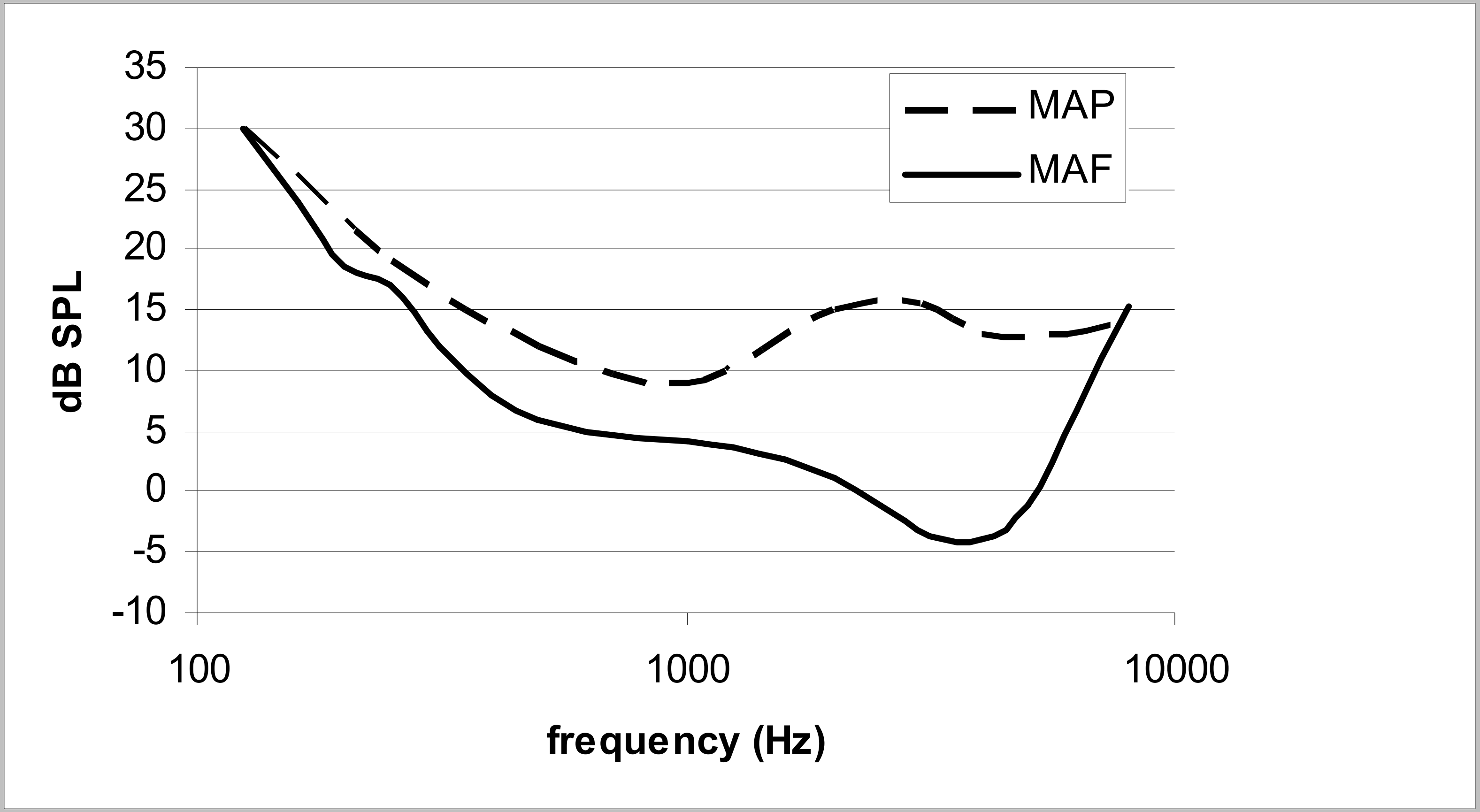
Thresholds for different mammals



Two ways to define a threshold once determined

- minimum audible field (MAF)
  - in terms of the intensity of the sound field in which the observer's head is placed
- minimum audible pressure (MAP)
  - in terms of the pressure amplitude at the observer's ear drum
  - often used with reference to headphones, and even more so, insert earphones
- MAF includes effect of head, pinna & ear canal

MAP vs. MAF
Accounting for the difference

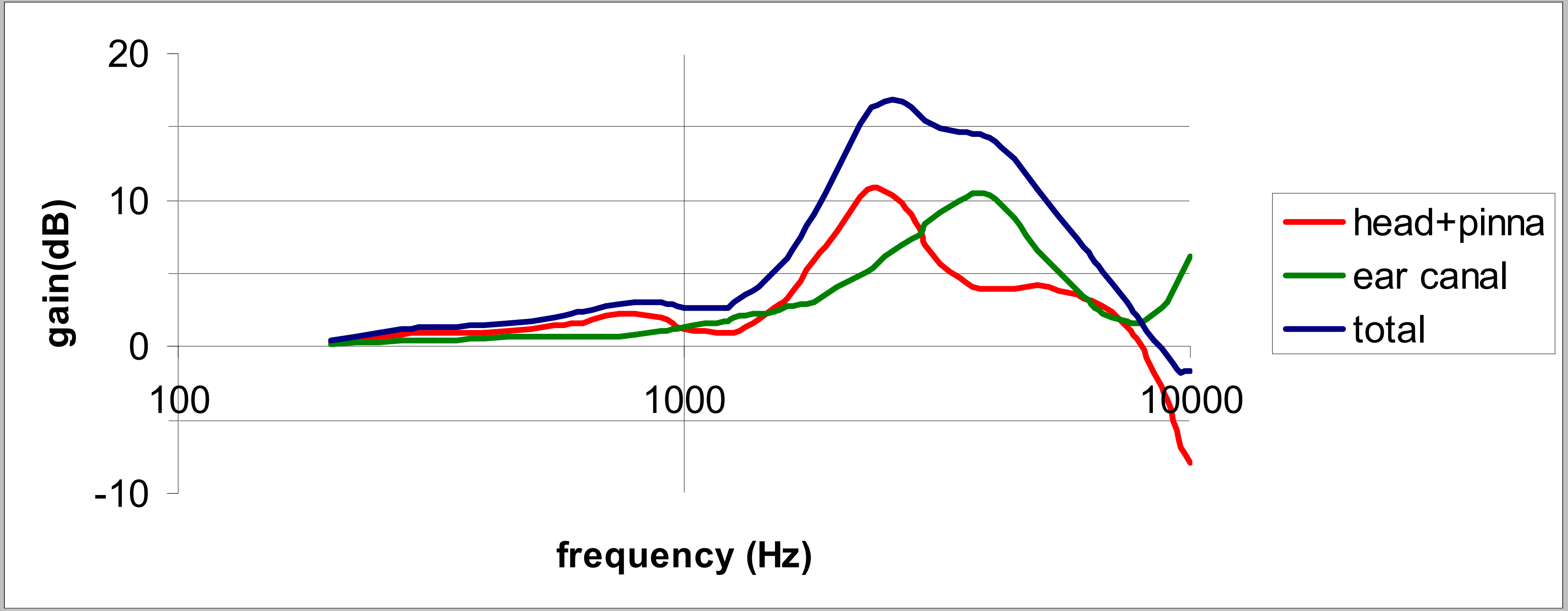


Frequency responses for:

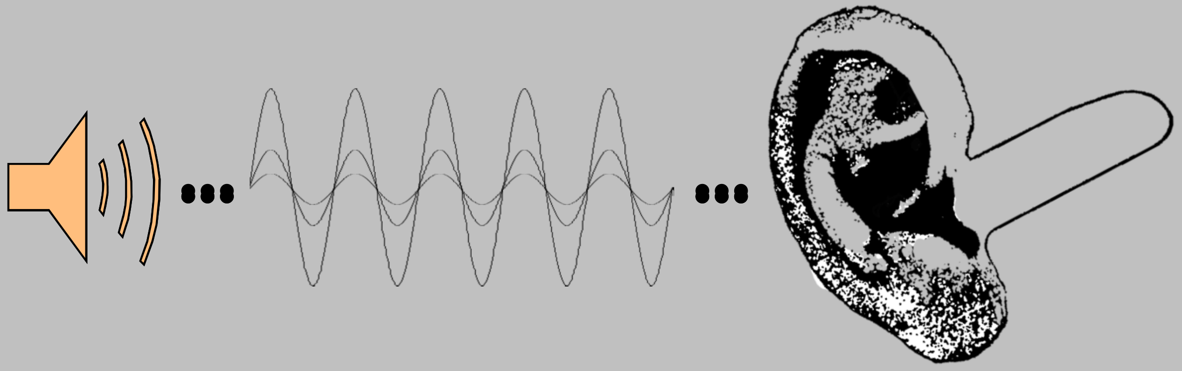
ear-canal entrance
free-field pressure

near the ear drum
ear-canal entrance

Total Effect:
near the ear drum
free-field pressure

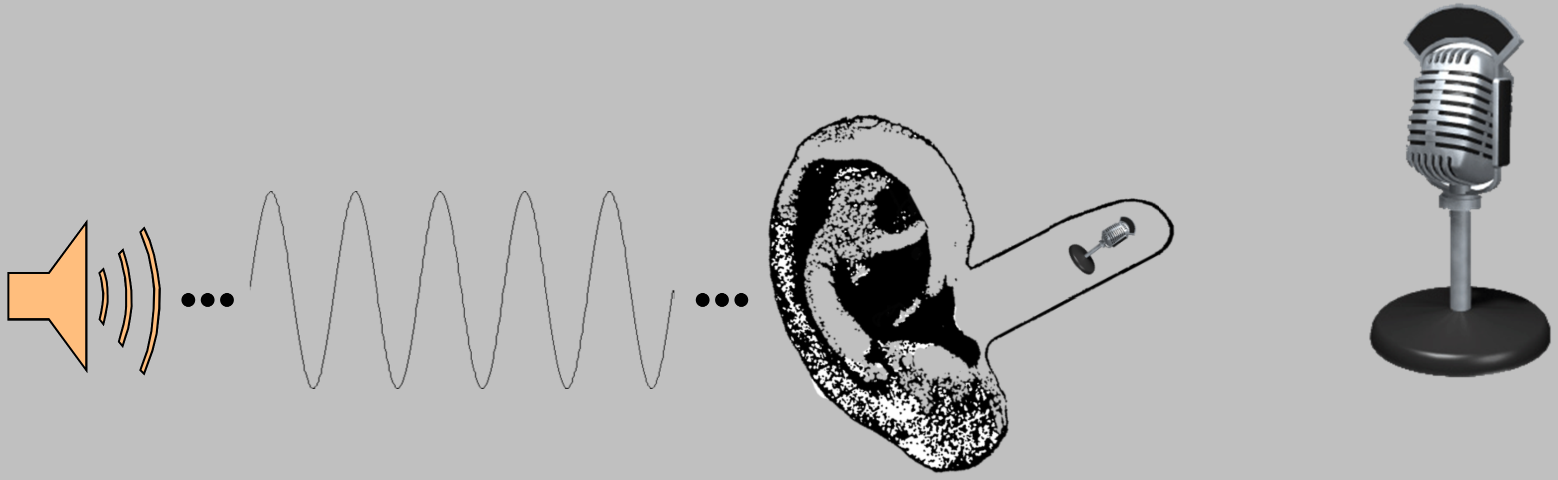


Determine a threshold for a 2-kHz sinusoid using a loudspeaker



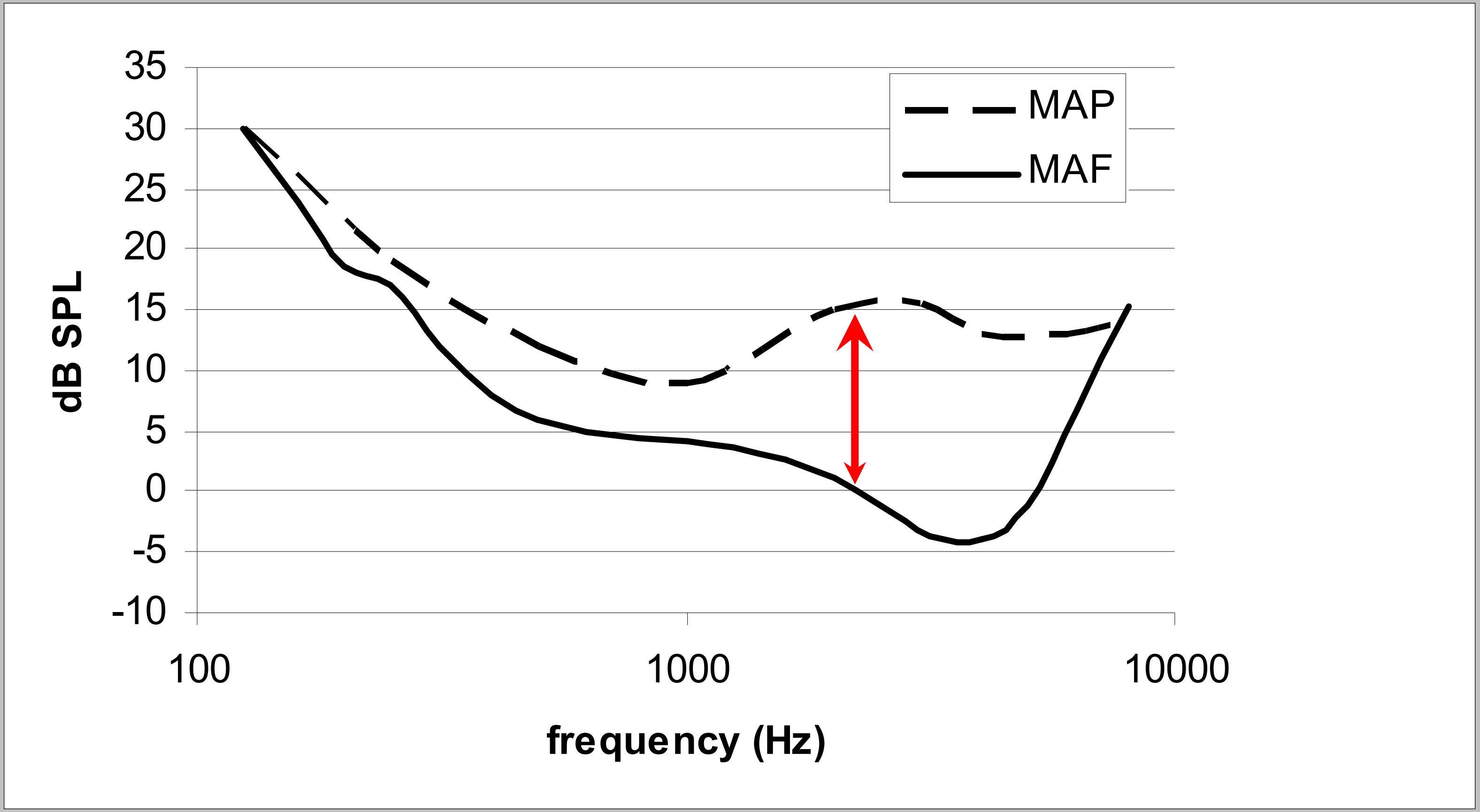
Now measure the sound level

at ear canal (MAP):
15 dB SPL



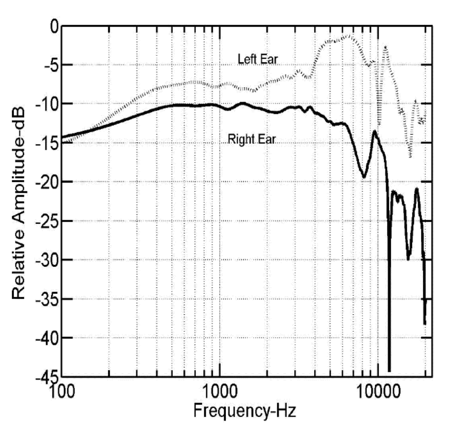
at head position (MAF):
0 dB SPL

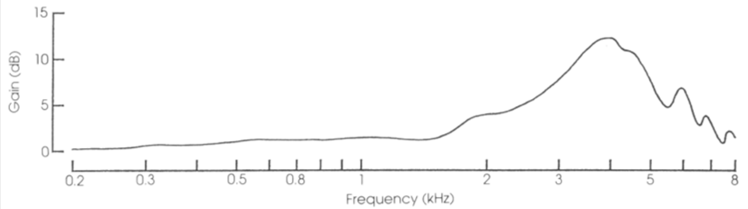
Accounting for MAP/MAF difference

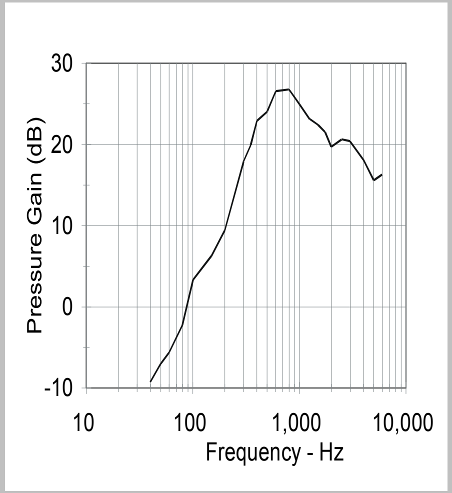


Accounting for the ‘bowl’

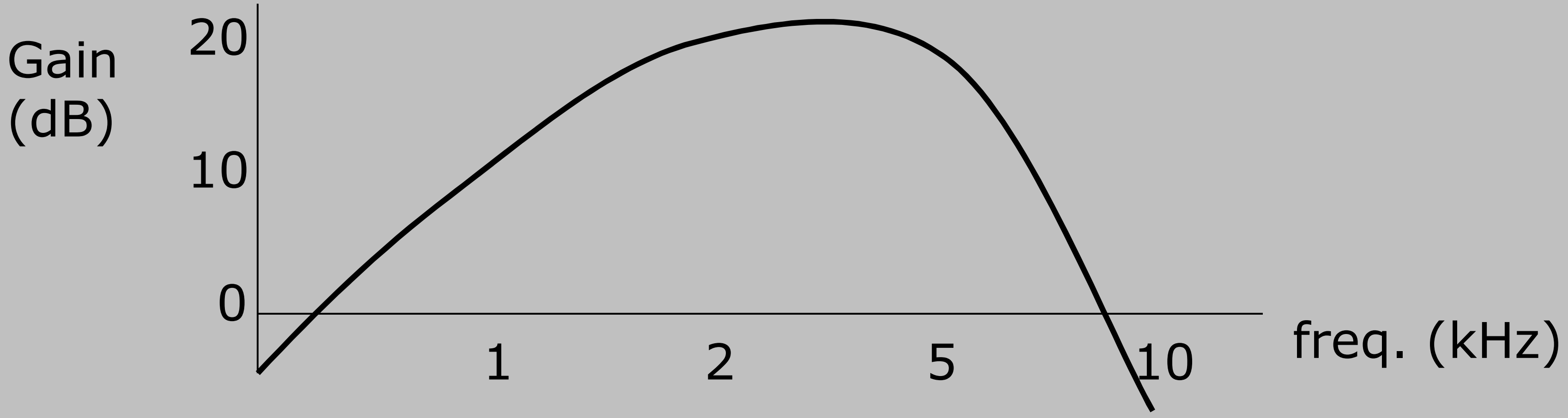
Combine head+pinna+canal+middle ear



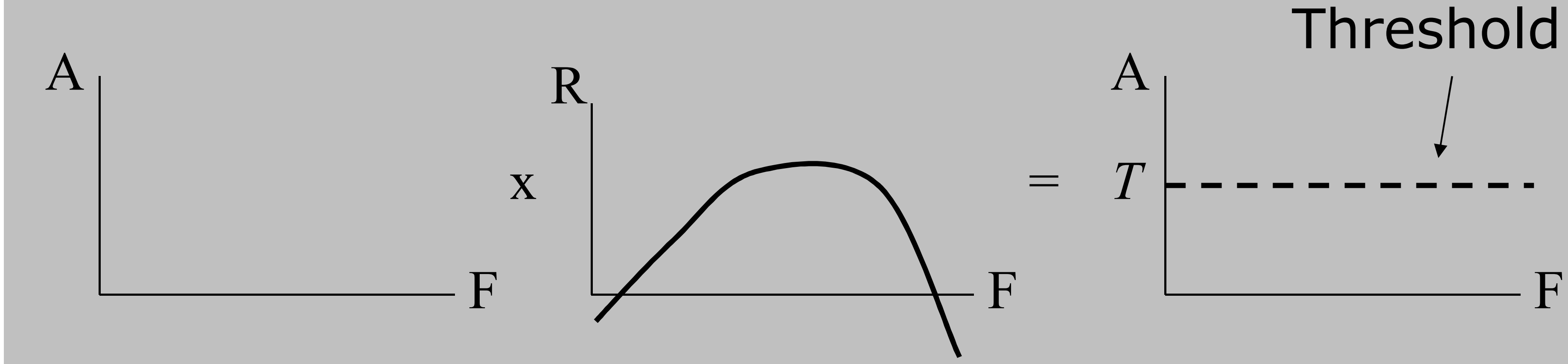




Overall

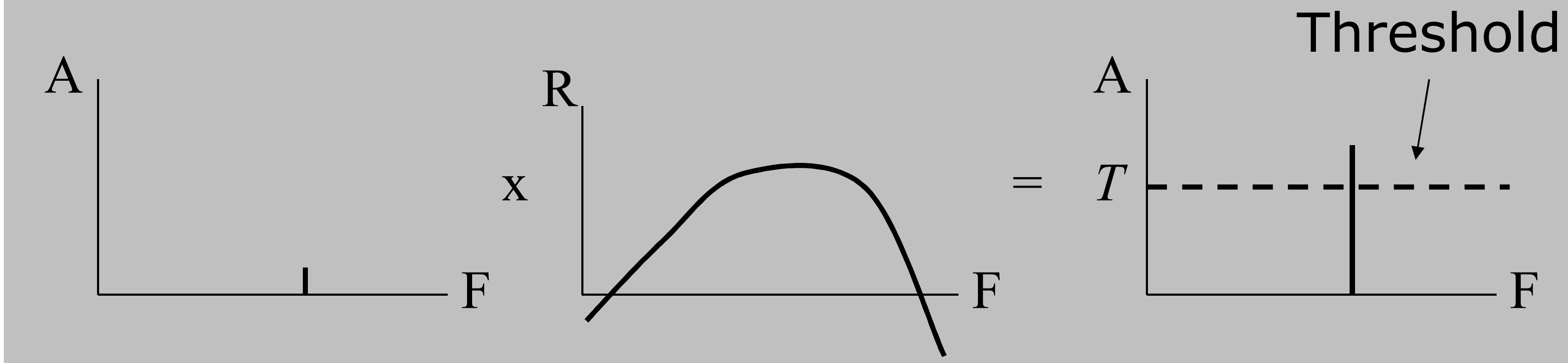


Detection of sinusoids in cochlea



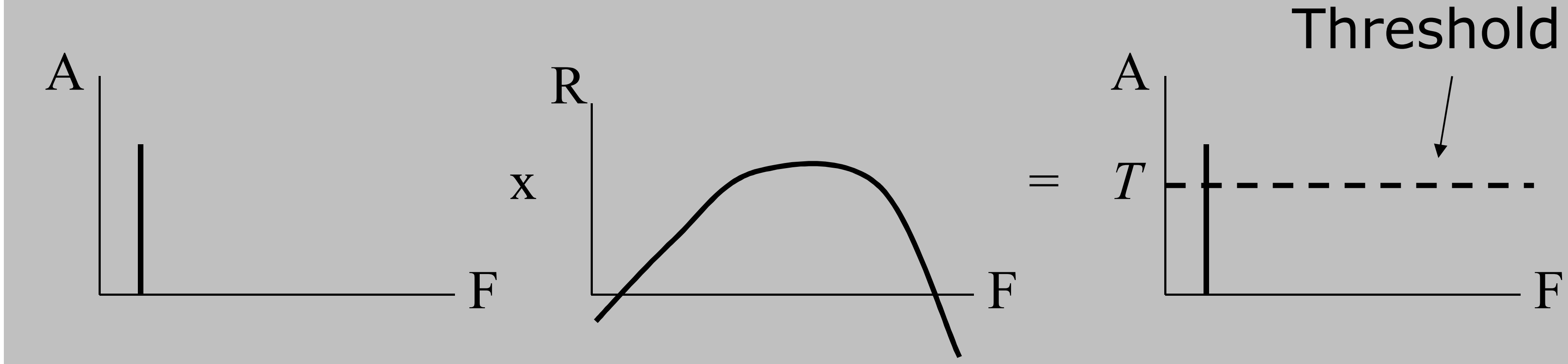
- How big a sinusoid do we have to put into our system for it to be detectable above some threshold?
- Main assumption: once cochlear pressure reaches a particular value, the basilar membrane moves sufficiently to make the nerves fire.

Detection of sinusoids in cochlea



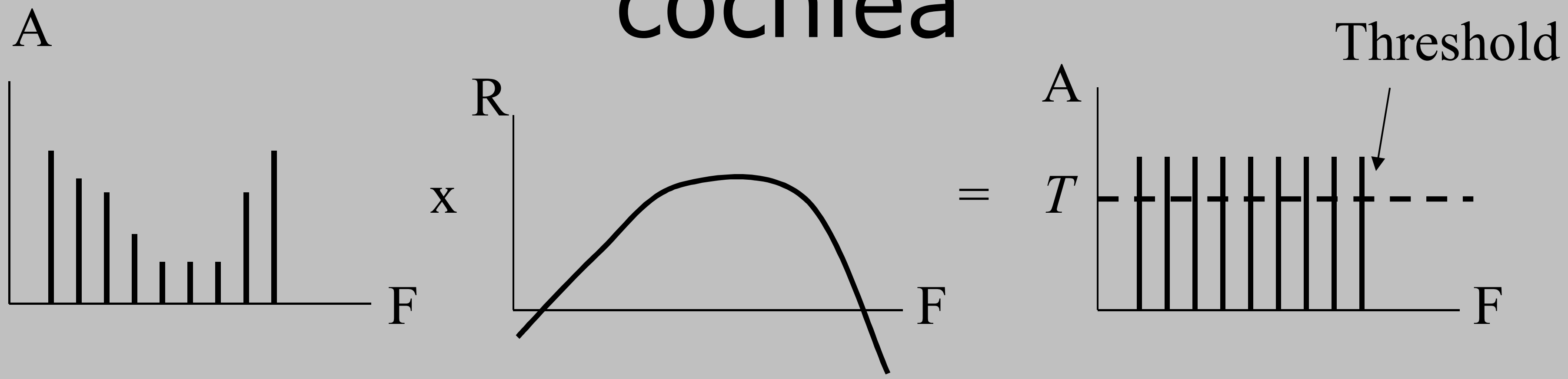
- A mid frequency sinusoid can be quite small because the outer and middle ears amplify the sound

Detection of sinusoids in cochlea



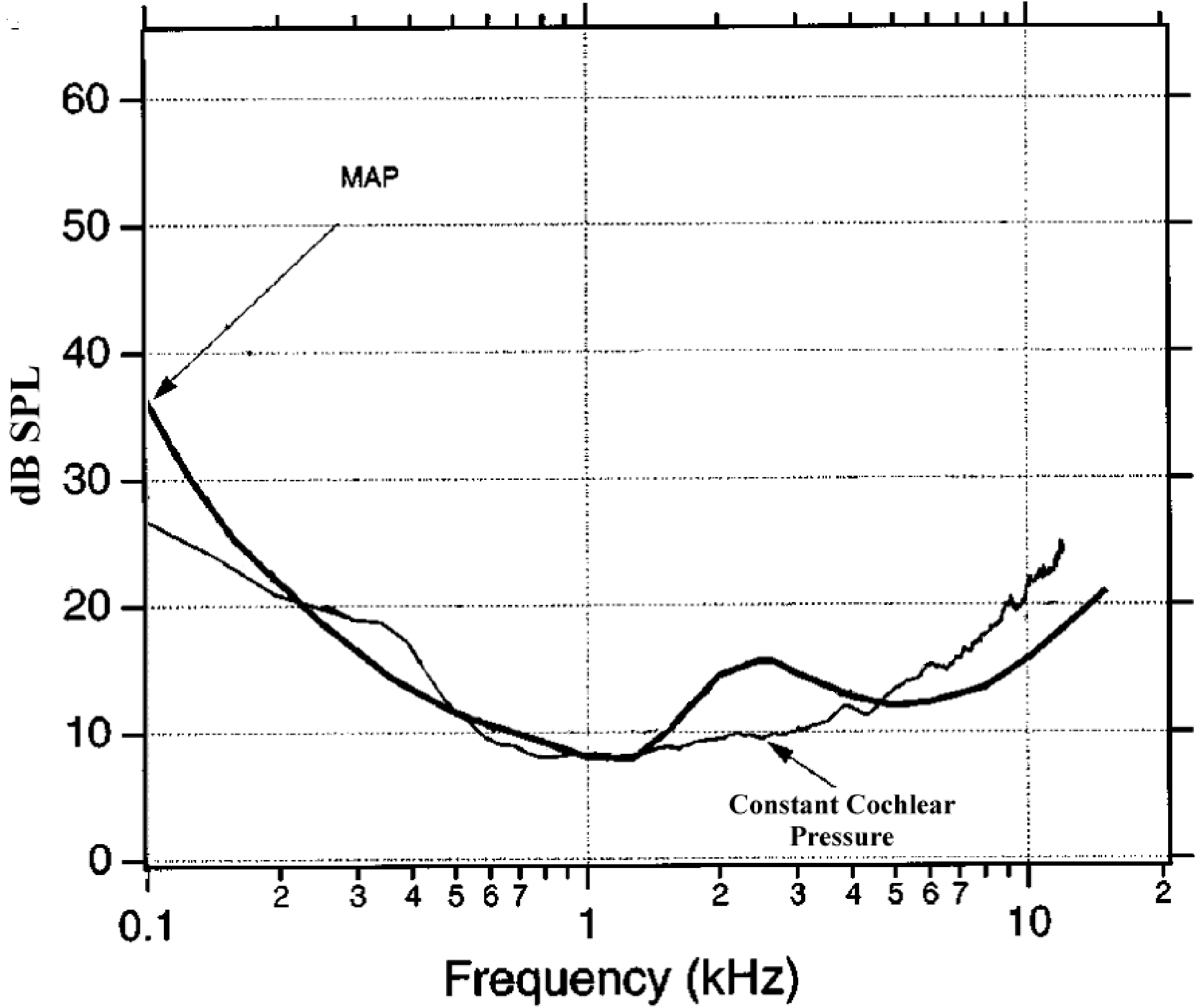
- A low frequency (or high frequency) sinusoid needs to be larger because the outer and middle ears do not amplify those frequencies so much

Detection of sinusoids in cochlea



- So, if the shape of the threshold curve is strongly affected by the efficiency of energy transfer into the cochlea …
- The threshold curve should look like this response turned upside-down: like a bowl.

Use MAP, and ignore contribution of head and ear canal



Much of the shape of the threshold curve can be accounted for by the efficiency of energy transfer into the cochlea

(from Puria, Peake & Rosowski, 1997)

Loudness

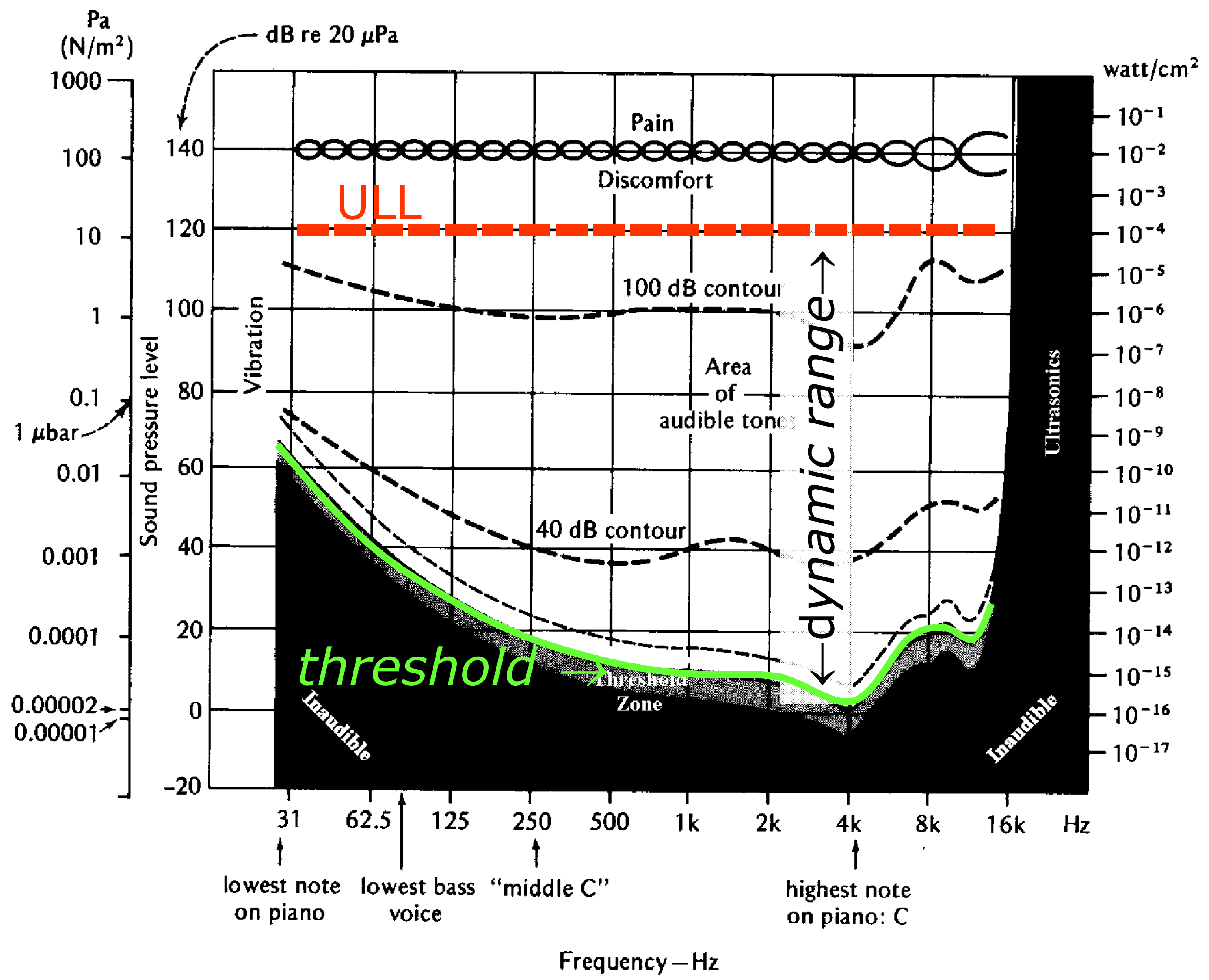
What determines how loud a sound is?

- Not just intensity

Loudness can also depend upon -

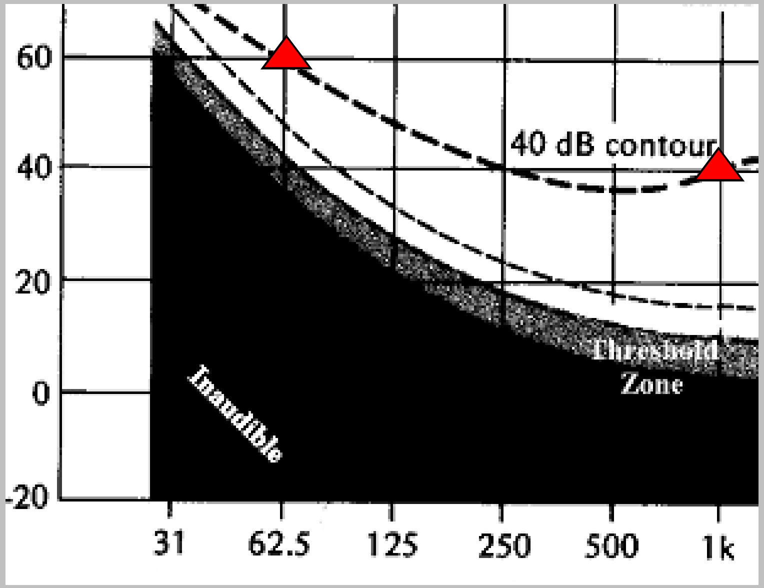
- Duration
  - Temporal integration (up to ~ 250 ms)
- Context
  - Loudness adaptation (over seconds or mins)
- Frequency

Loudness of supra-threshold sinusoids



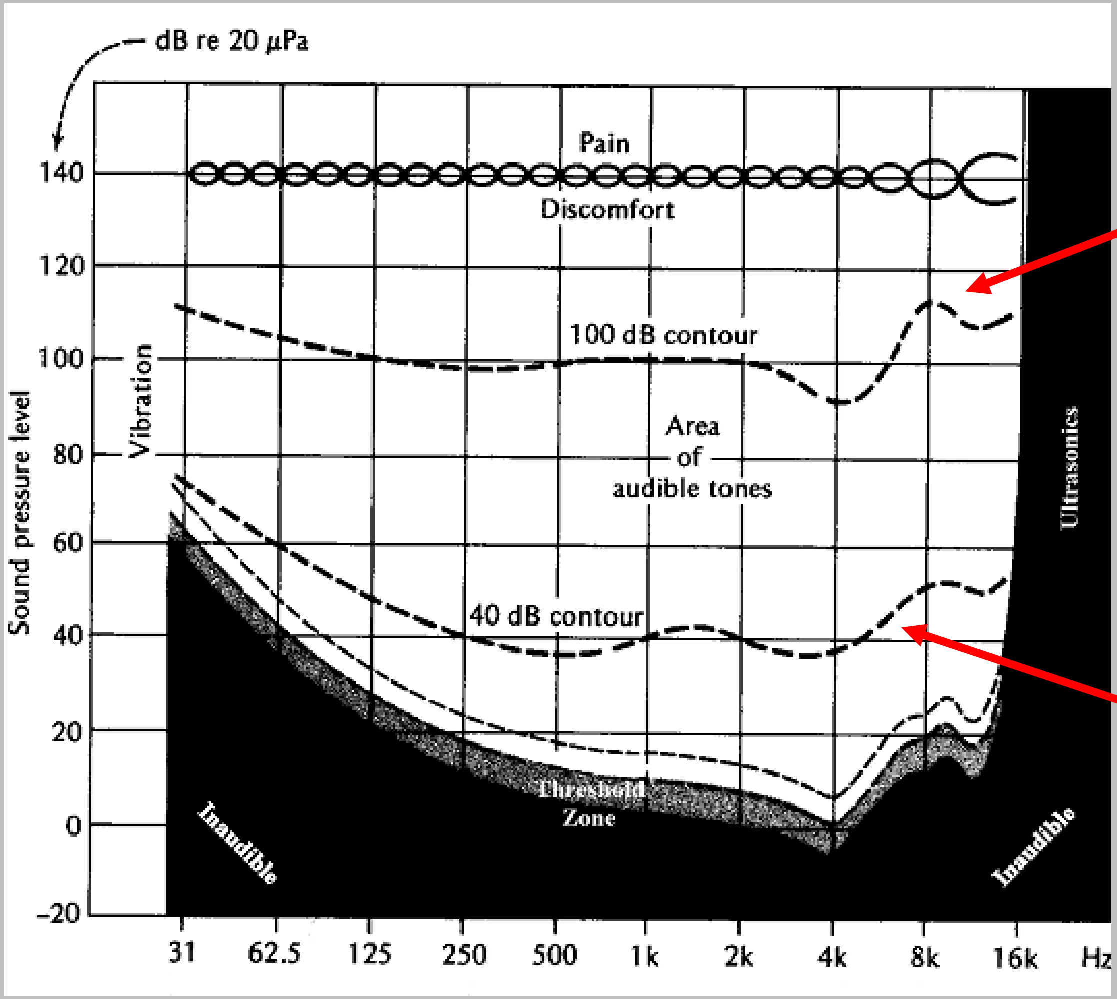
The Phon scale of loudness

- ‘A sound has a loudness of *X* phons if it is equally as loud as a sinewave of *X* dB SPL at 1kHz’



e.g. A 62.5Hz sinusoid at 60dB SPL has a loudness of 40 phons, because it is equally as loud as a 40dB SPL sinusoid at 1kHz

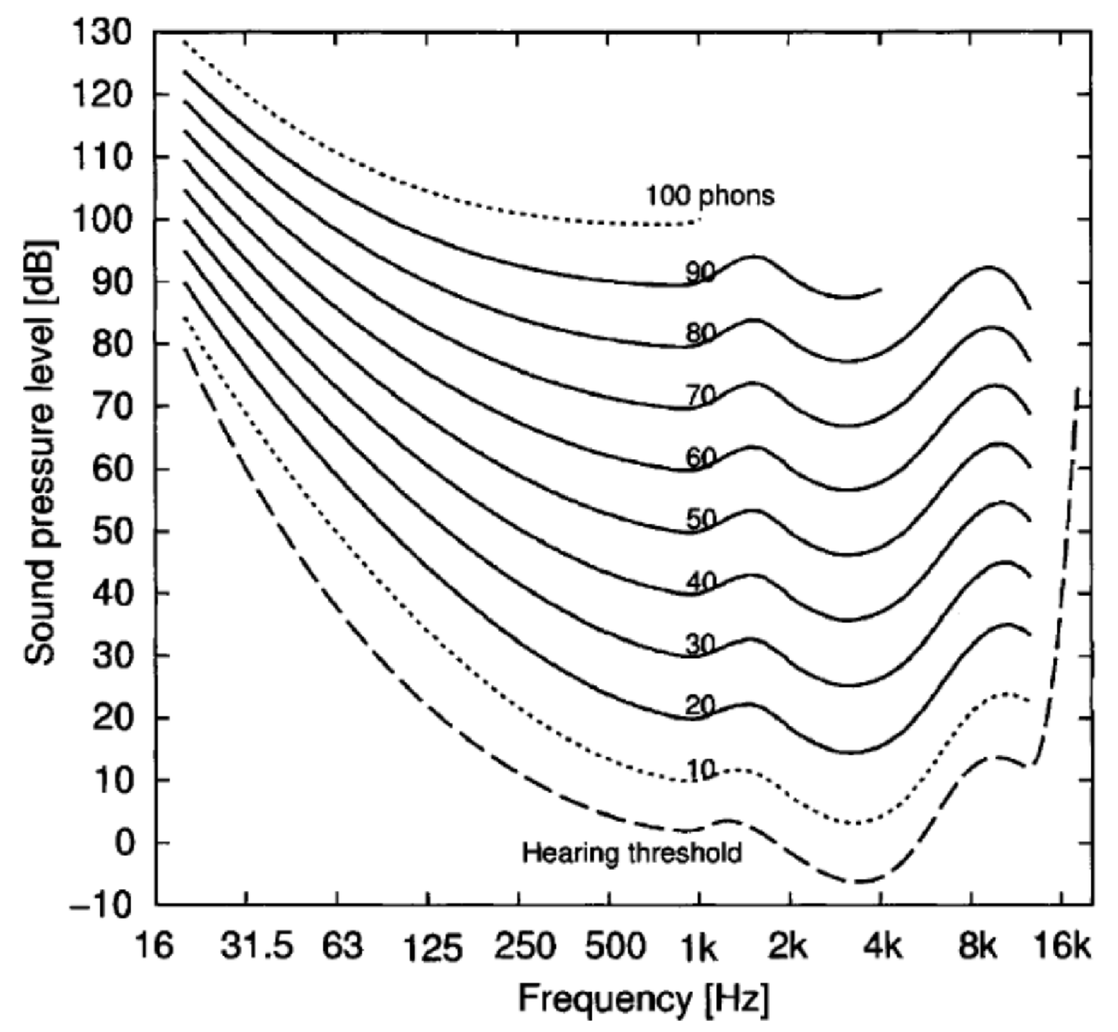
Equal loudness contours



Contour of tones equal in loudness to 100 dB SPL sinusoid @ 1kHz

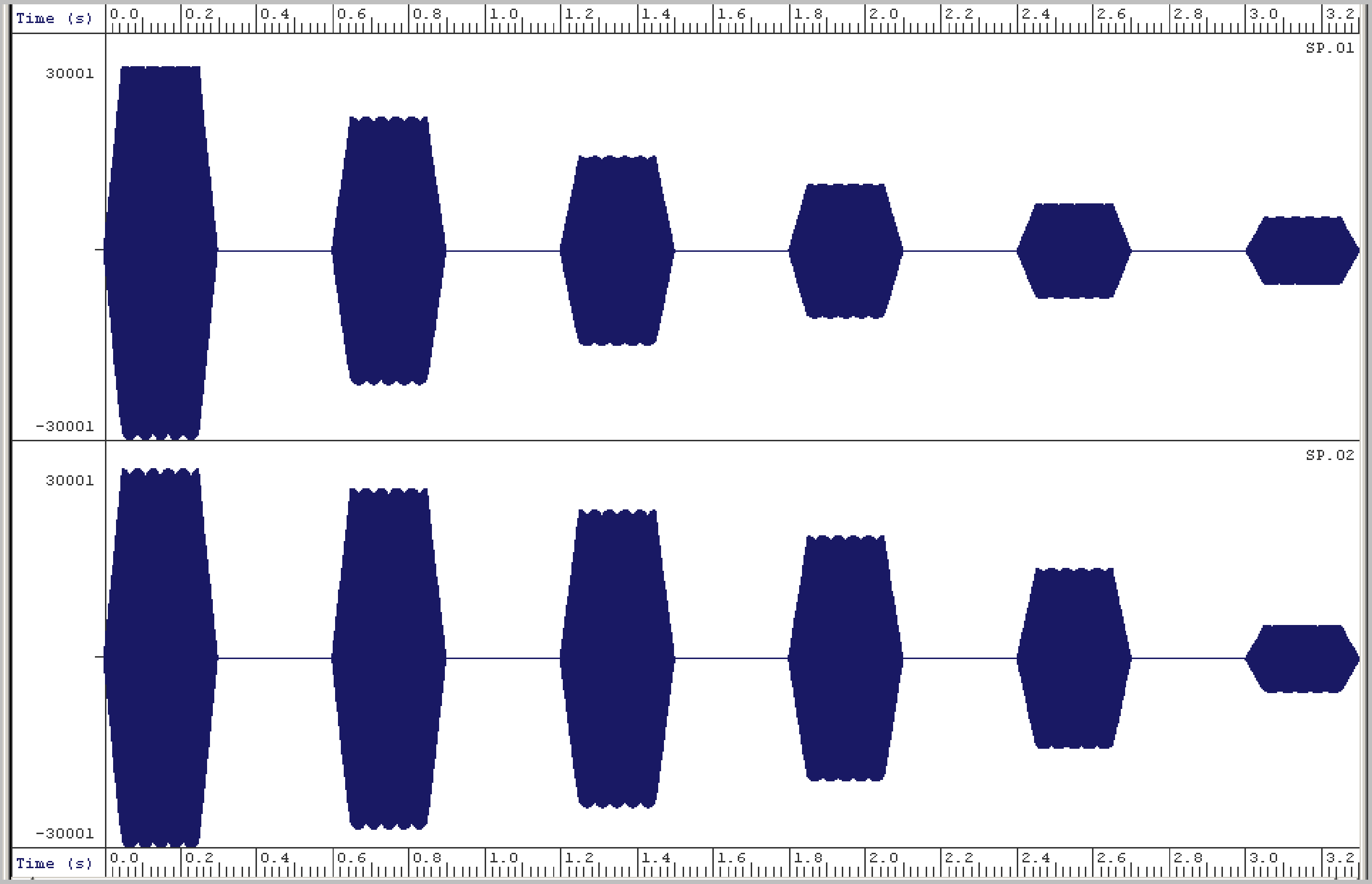
Contour of tones equal in loudness to 40 dB SPL sinusoid @ 1kHz

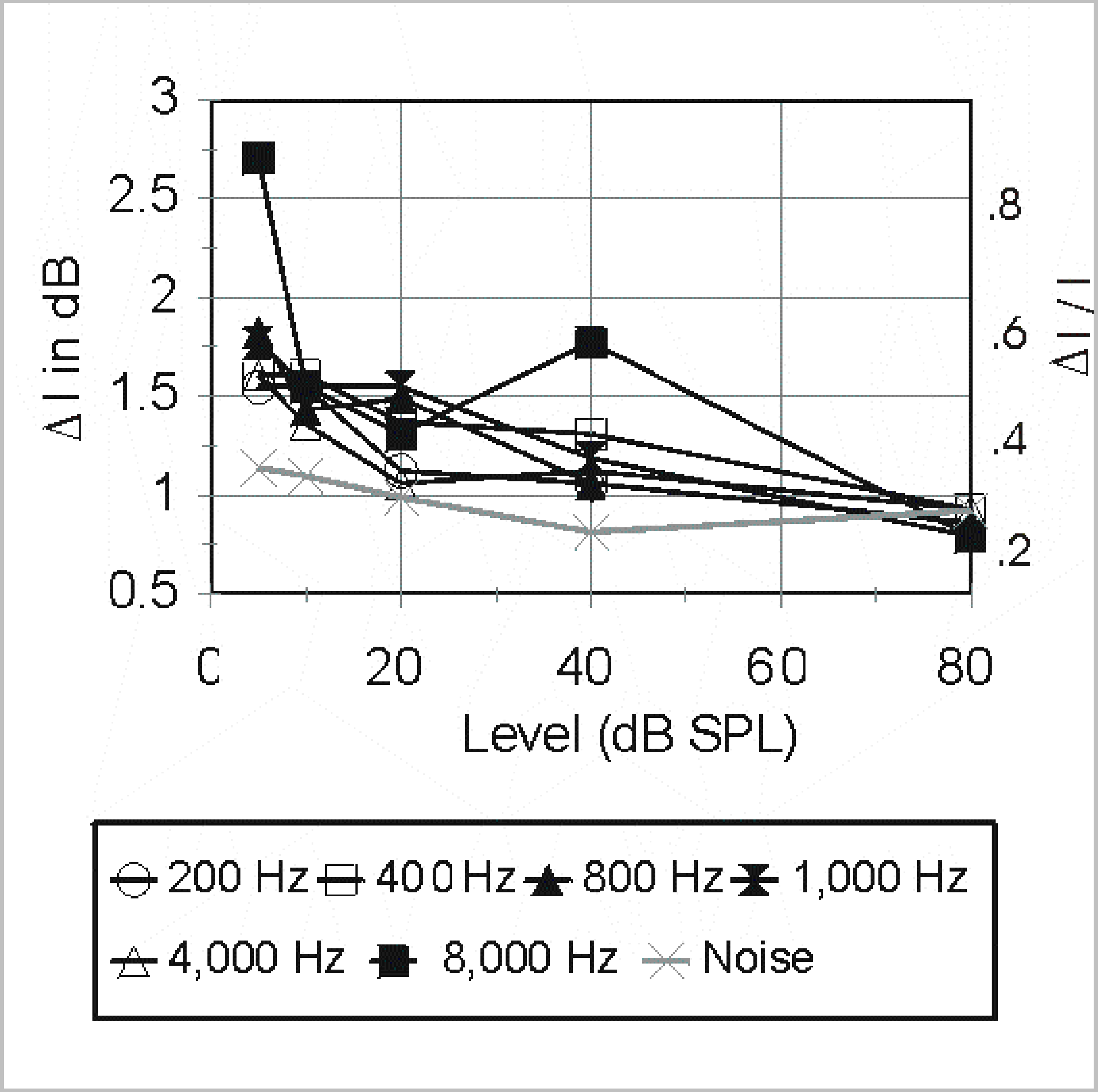
Contemporary equal loudness contours



From Suzuki & Takeshima (2004) JASA

Perceived loudness is (roughly) logarithmically related to pressure

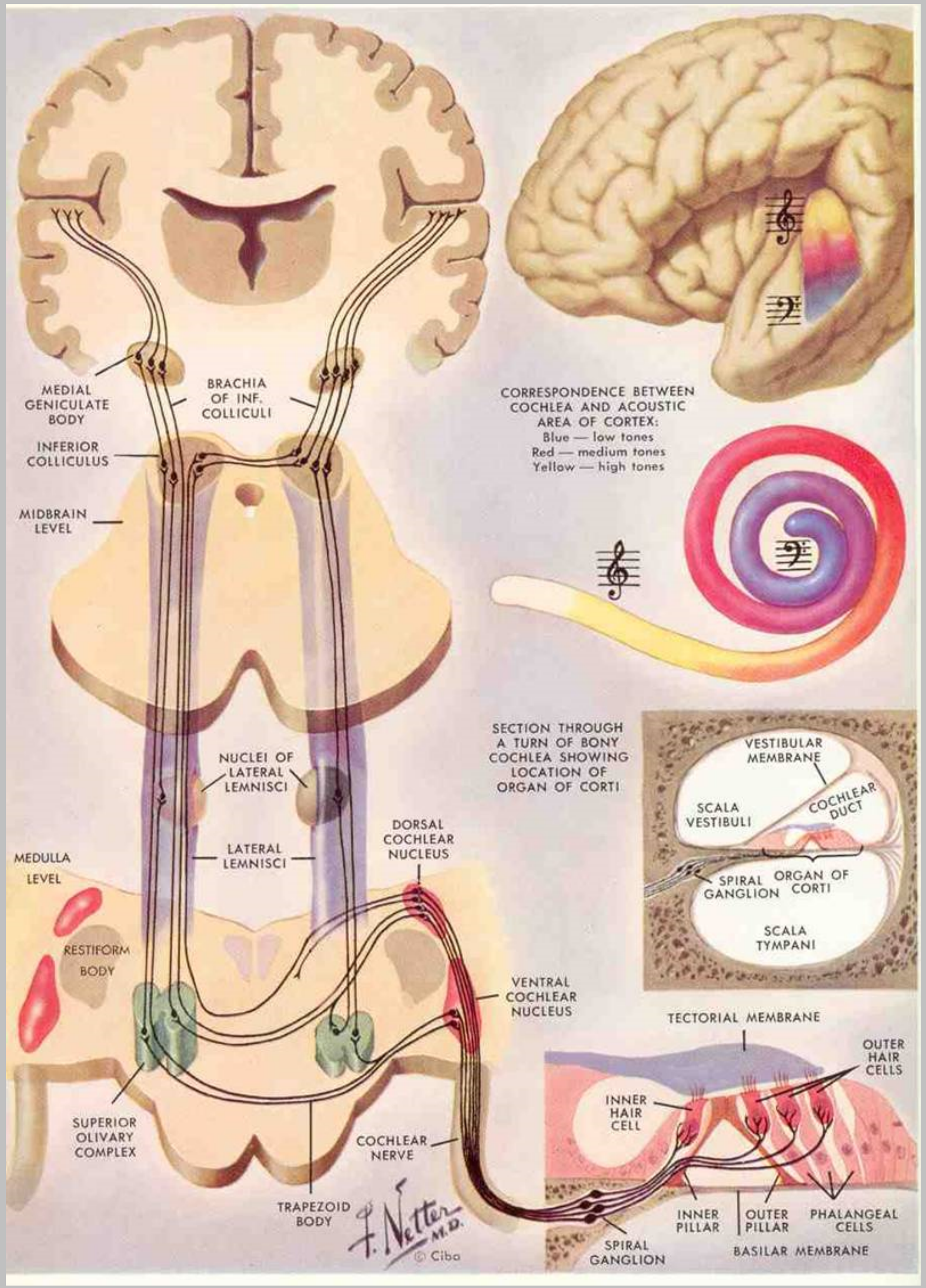




Just-noticeable differences (jnds) in intensity are roughly constant in dB

from Yost (2007)

Psychoacoustic reflections of frequency selectivity

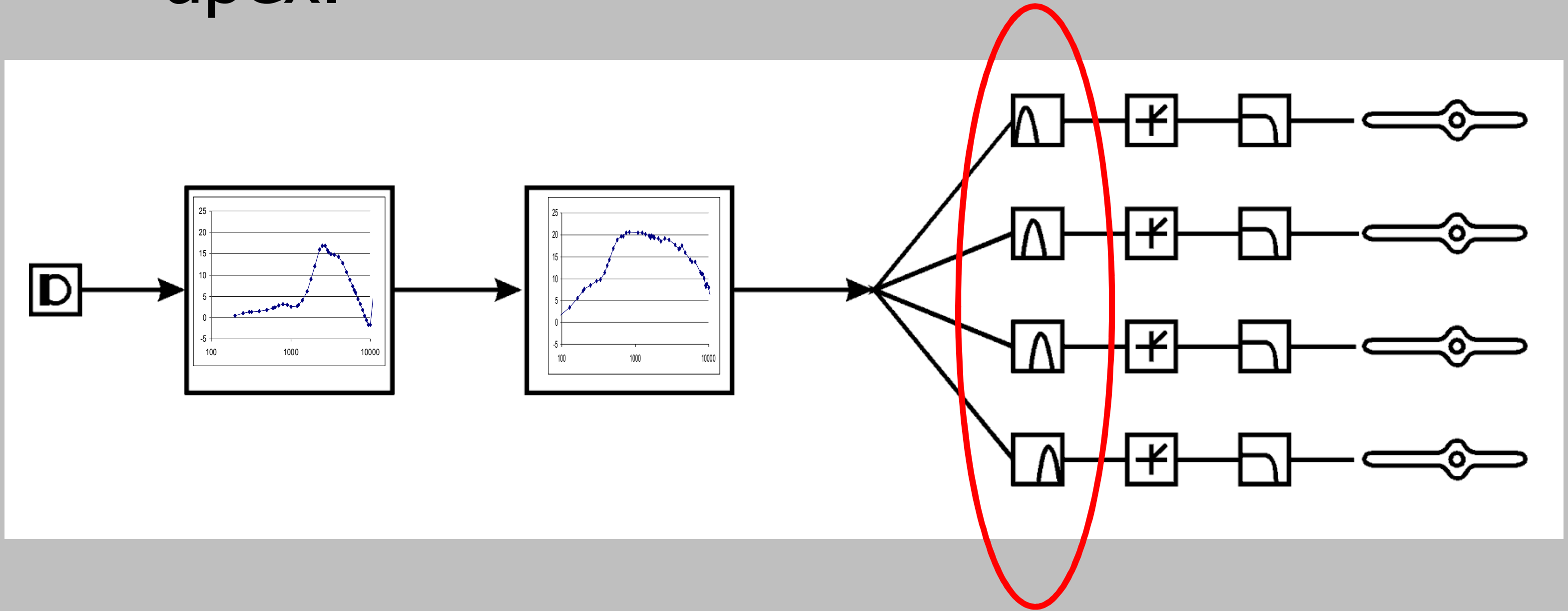


Frequency selectivity

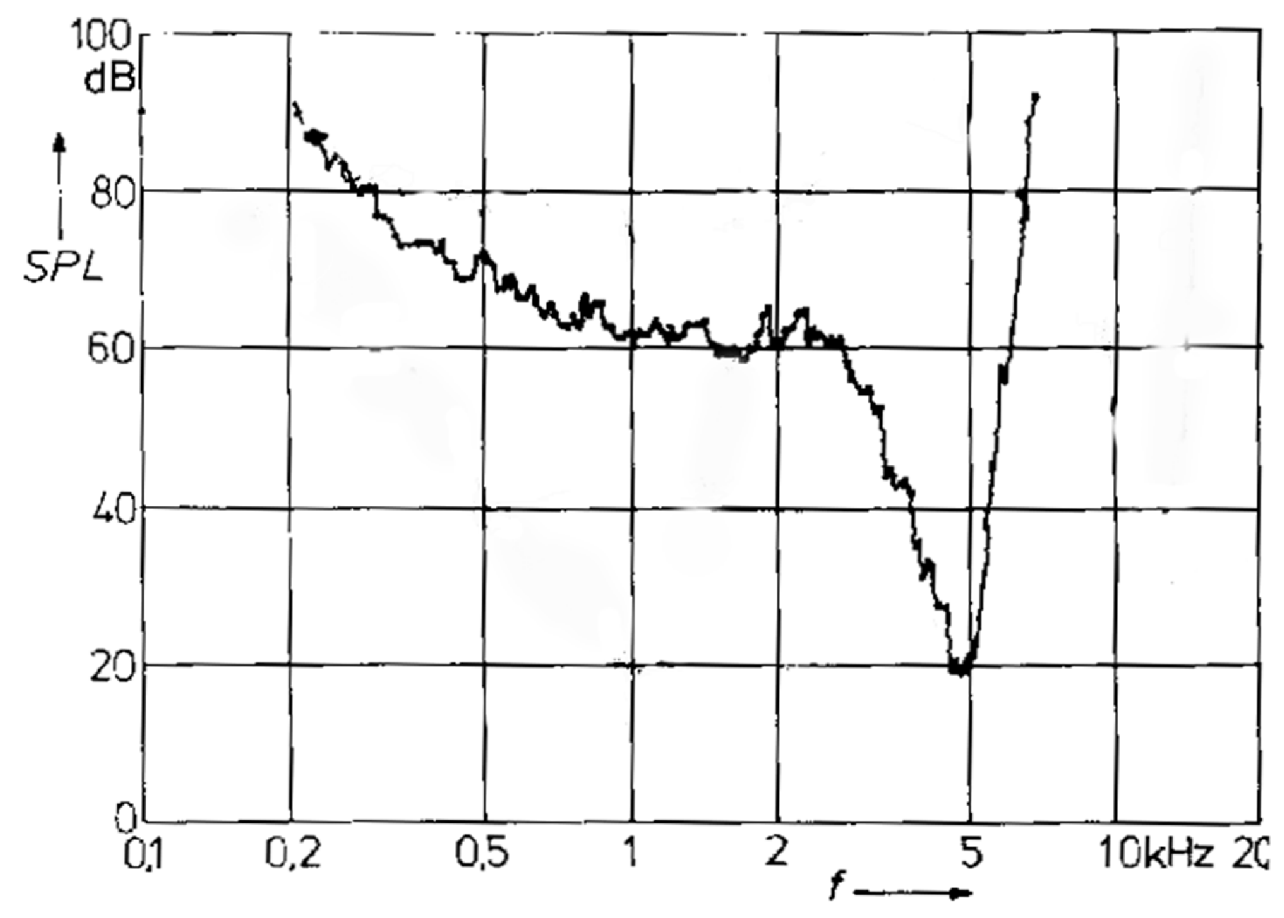
- **Auditory Filters**
  - The auditory periphery can be modelled as a bank of bandpass filters
- This splitting up of sounds into different frequency regions is fundamental to the auditory processing of complex real sounds, e.g. speech.

The filter bank analogy

- Imagine each afferent auditory nerve fibre has a bandpass filter attached to its input
  - centre frequencies decreasing from base to apex.

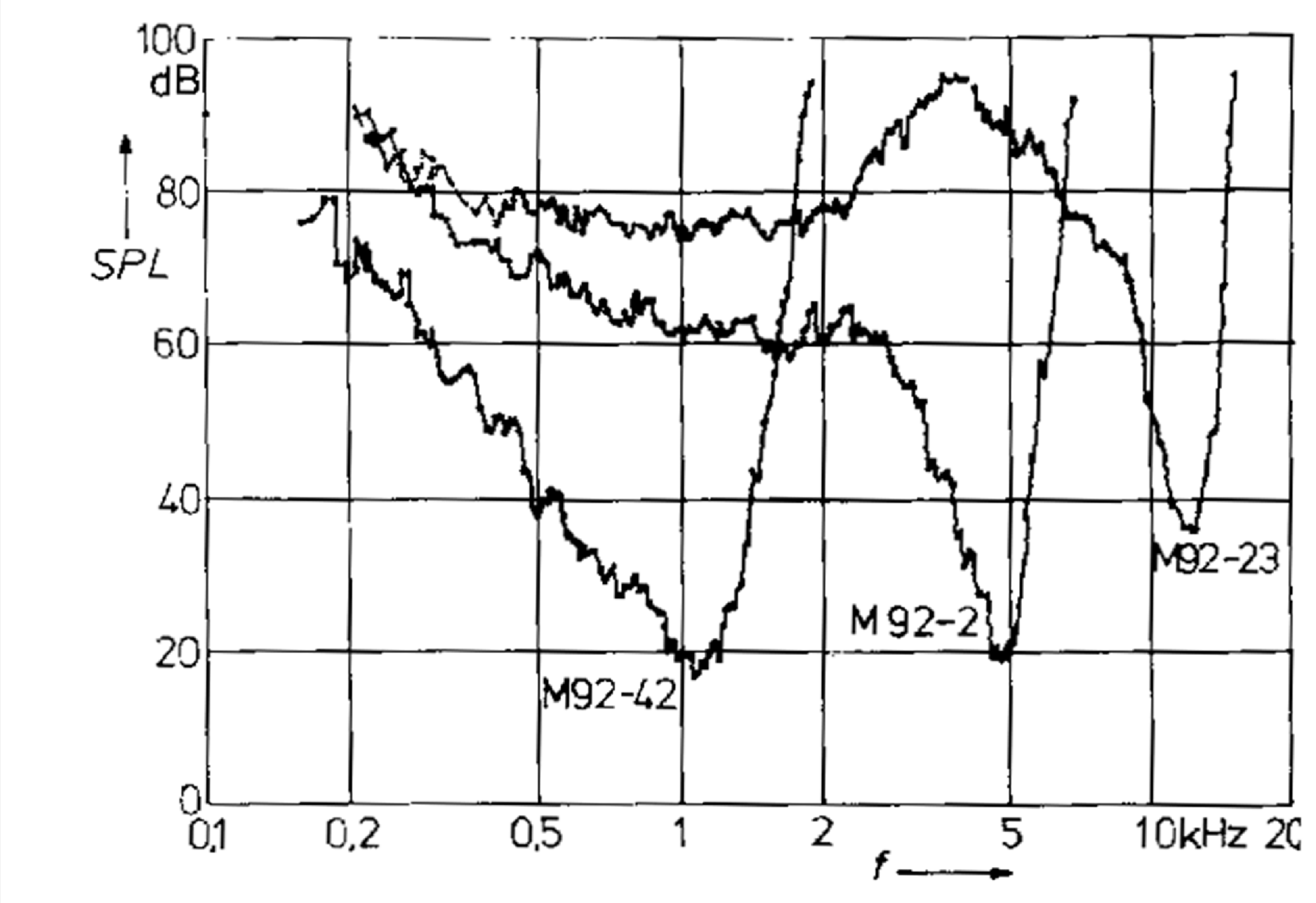


The physiological tuning curve: Auditory filtering reflected in the auditory nerve

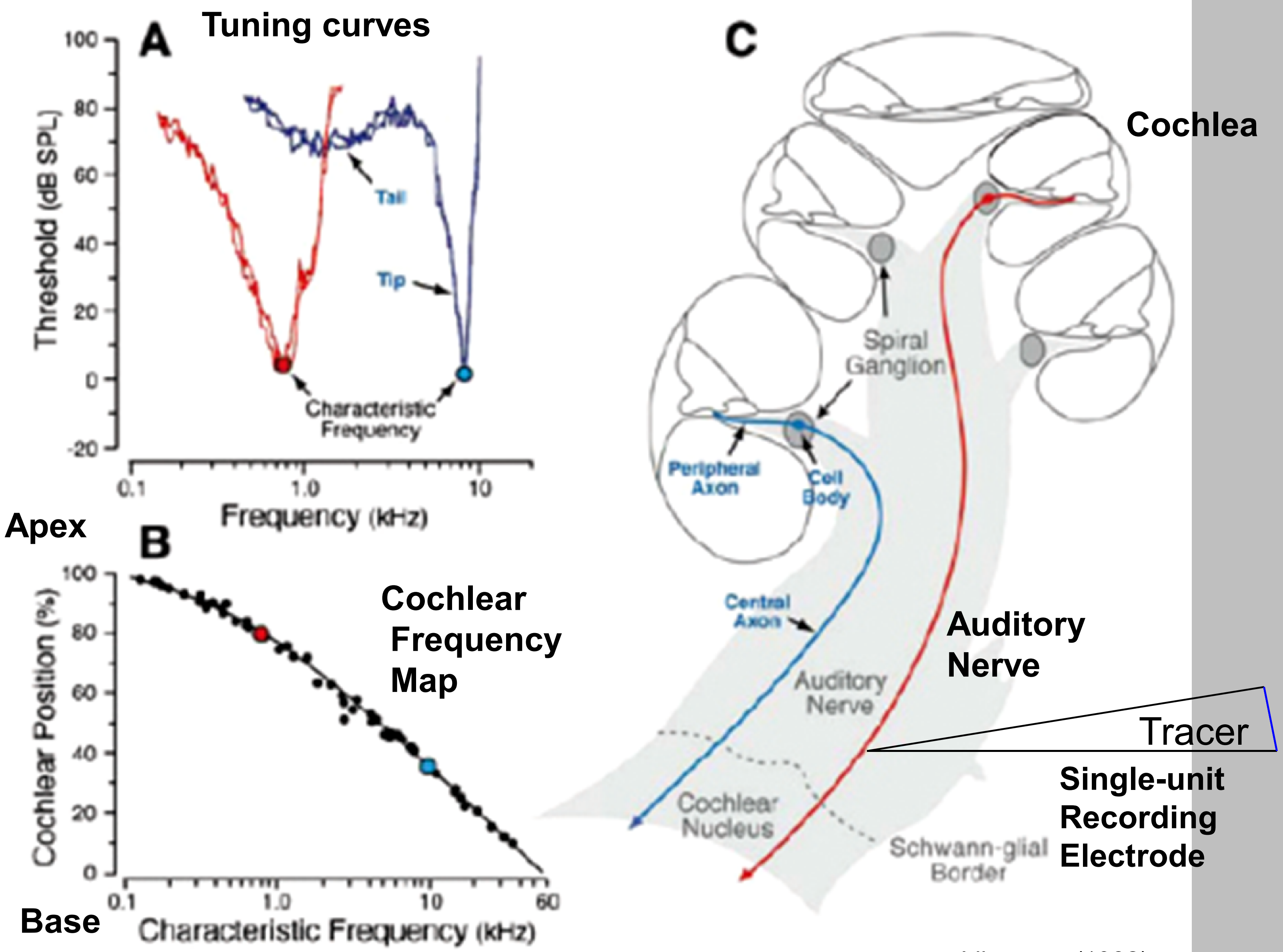


‘Audiograms’ of single auditory nerve fibres reflect BM tuning

The 'best' frequency of a particular tuning curve depends upon the BM position of the IHC to which the afferent neuron is synapsing



Auditory Nerve Structure and Function



Liberman (1982)

Masking: interactions of sounds

- Much of what we know about the properties of auditory filters has been revealed by studies of masking.

- How does the presence of one sound affect the detection of another?

  - It depends on their frequencies

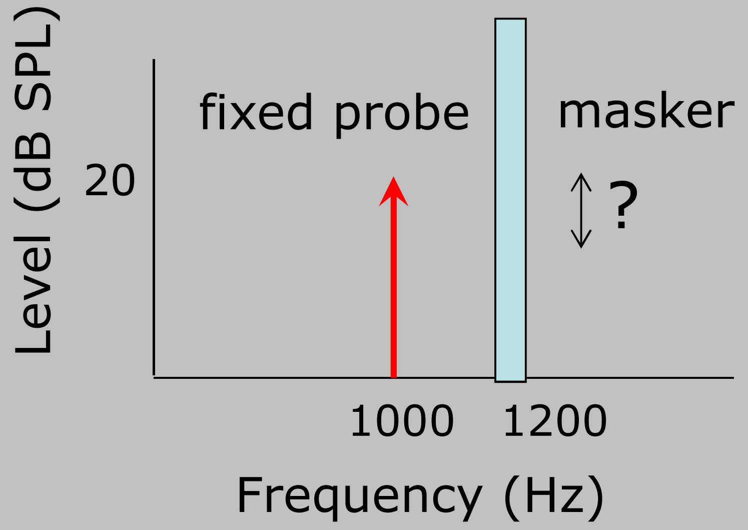
  - Maskers that are close to the signal frequency (so pass through the same auditory filter) cause much more interference

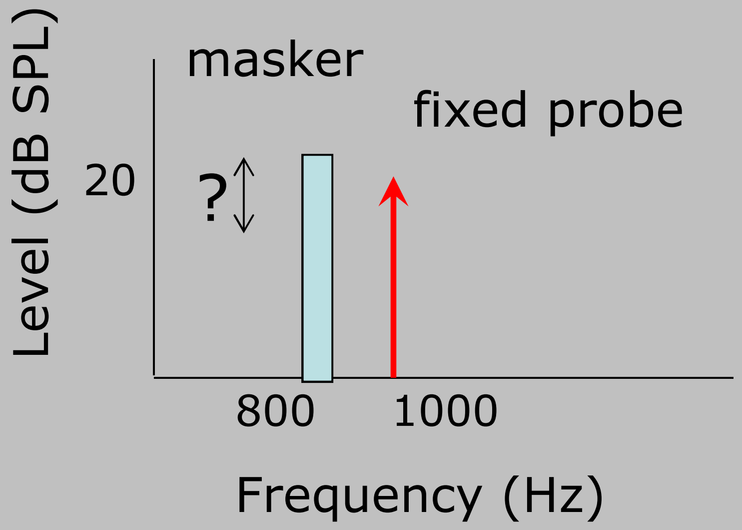
Play Demo

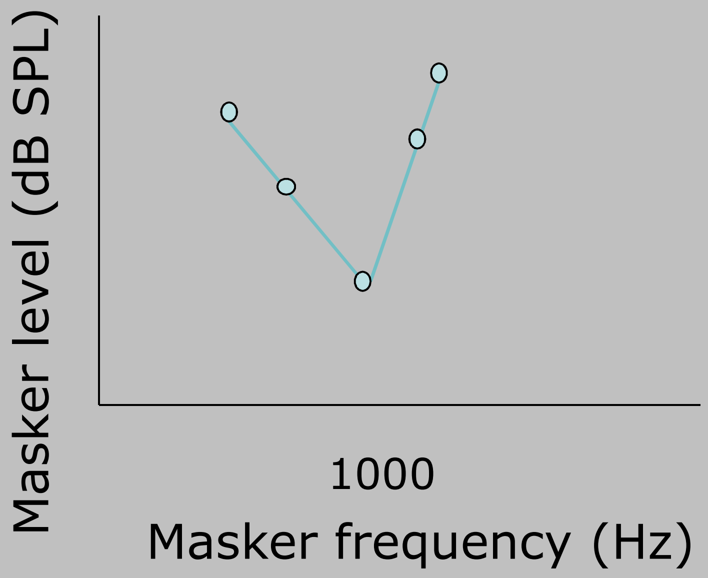
Masking experiments

- Listen for a *probe* or *target* (typically a sinusoid) in a background of a masker with a variety of spectral shapes (typically a noise).
- Assume: A listener has independent access to, and can ‘listen’ selectively to the output of an individual auditory filter – the one that will give best performance.
    - *the probe frequency controls the centre frequency of the auditory filter that is attended to*
- Assume: Only noise that passes through the same filter as the sinusoid can mask it.
- Assume: Only the ‘place’ principle applies — no temporal information.
- The *power spectrum model of masking*

Psychophysical tuning curves (PTCs)

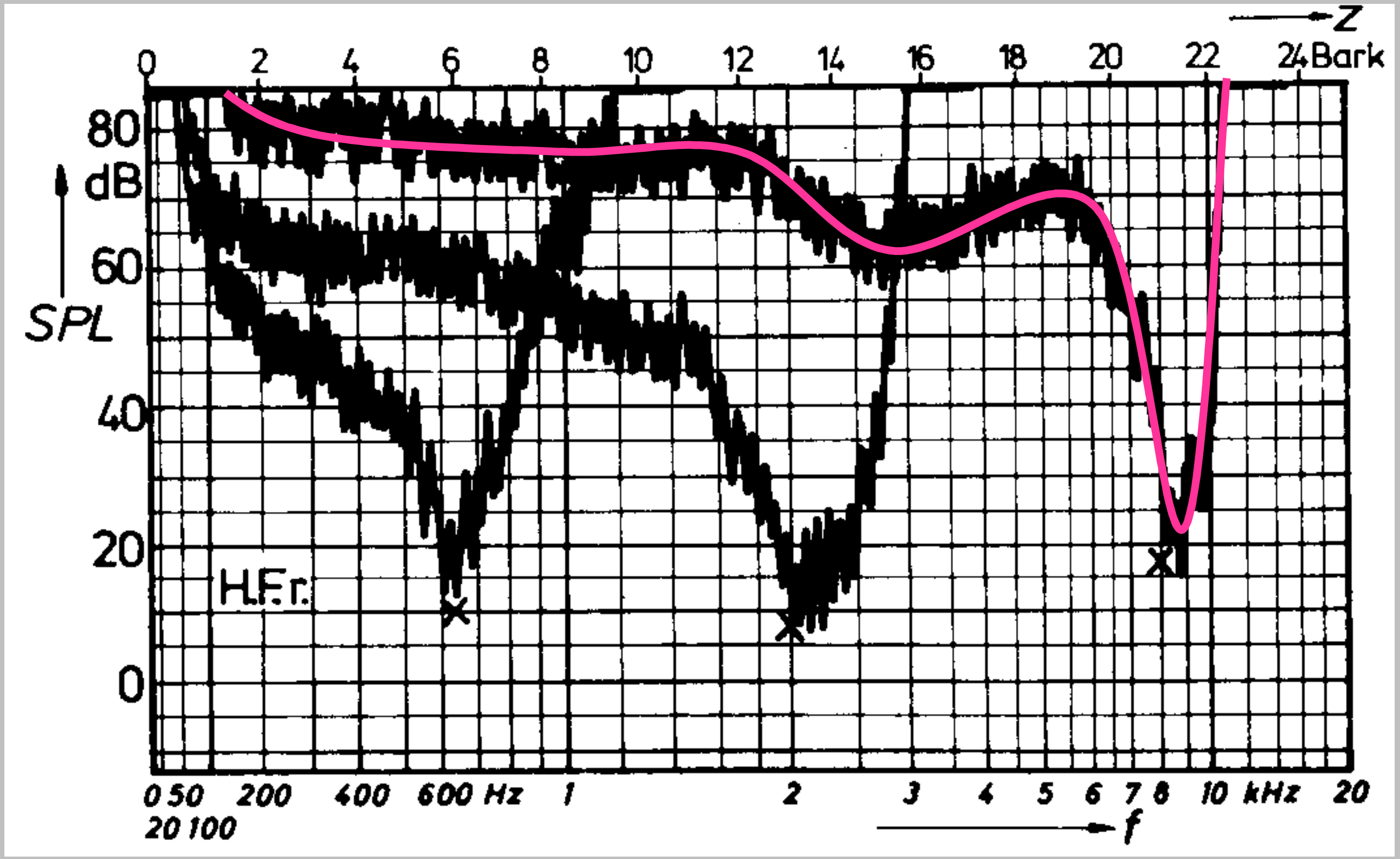




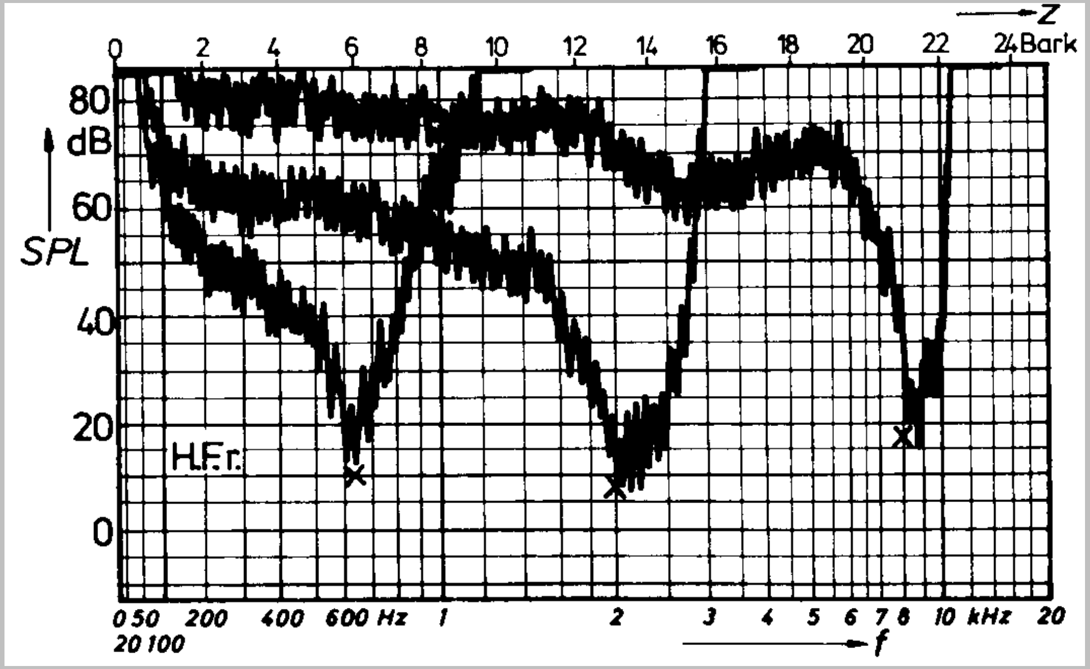


Psychophysical tuning curves (PTCs)

Determine the minimum level of a narrow-band masker at a wide variety of frequencies that will just mask a fixed ***low-level*** *sinusoidal probe.*



Psychophysical tuning curves (PTCs)



Is a psychophysical tuning curve a correlate of an excitation pattern (something like a spectrum) or a tuning curve (something like a frequency response)?

Why you can’t easily interpret PTCs at higher levels: Off-frequency/place listening

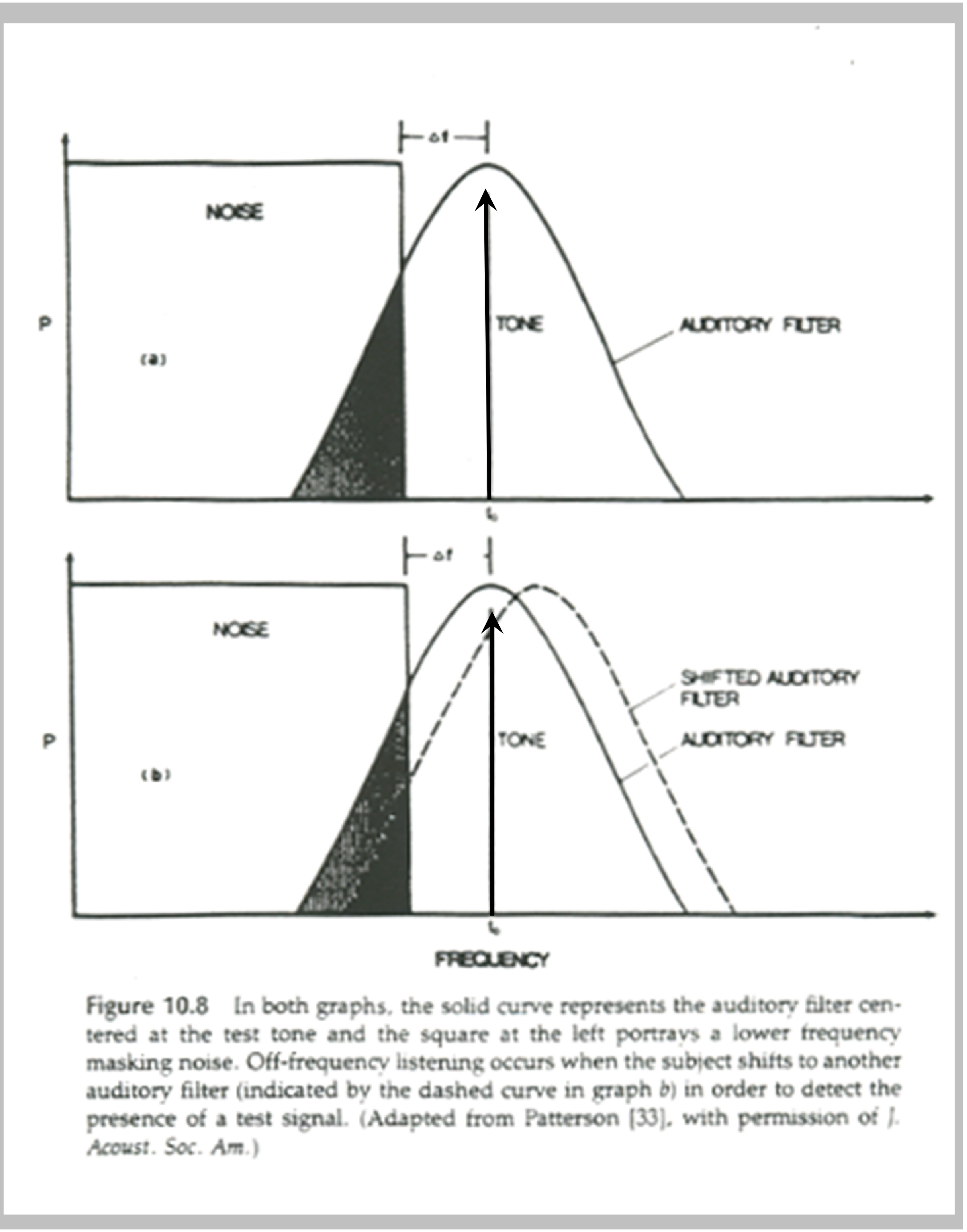


Figure 10.8 In both graphs, the solid curve represents the auditory filter centered at the test tone and the square at the left portrays a lower frequency masking noise. Off-frequency listening occurs when the subject shifts to another auditory filter (indicated by the dashed curve in graph *b*) in order to detect the presence of a test signal. (Adapted from Patterson [33], with permission of *J. Acoust. Soc. Am.*)

From Gelfand (1998)

Notch (*band stop*) noises limit off-place listening

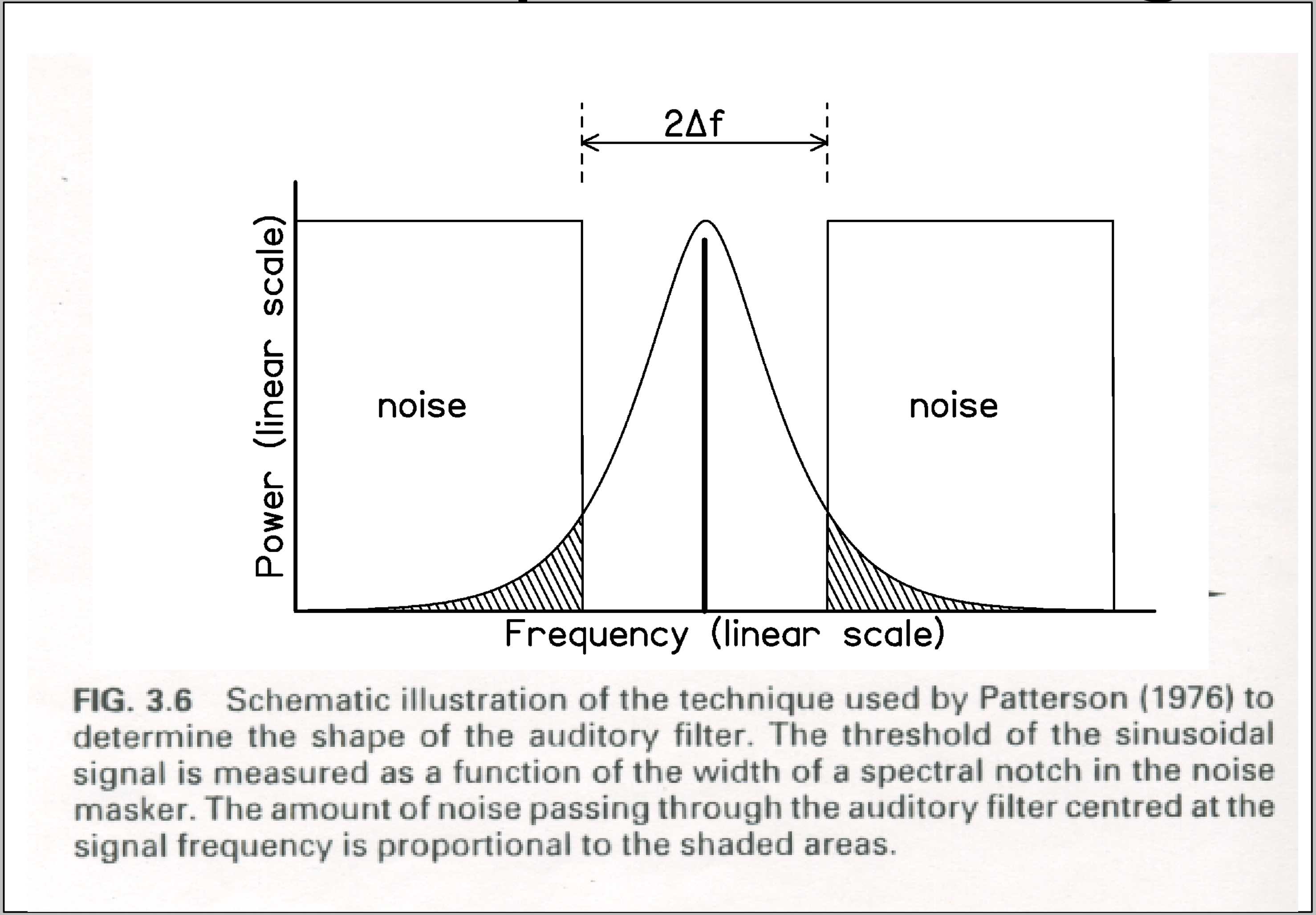
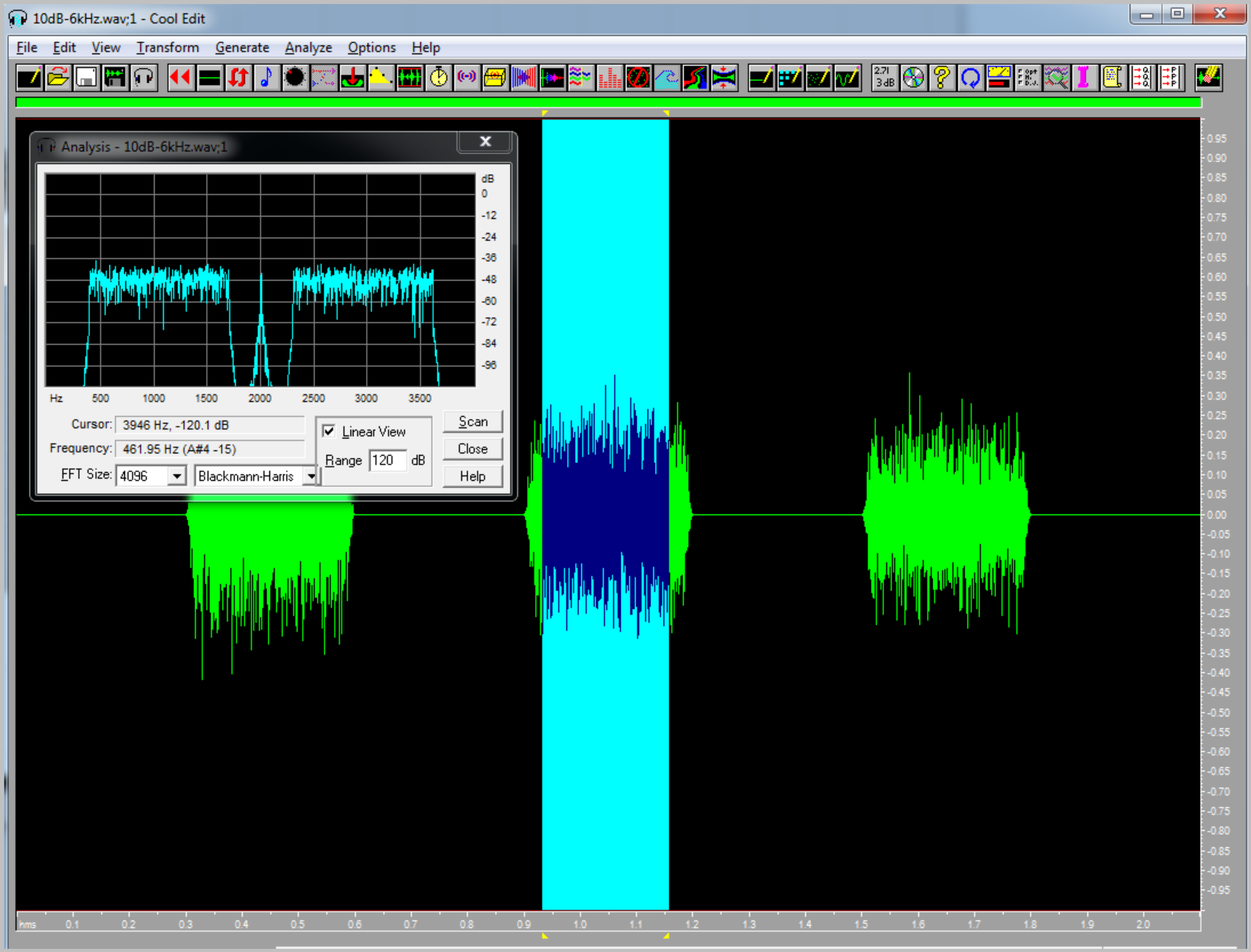


FIG. 3.6 Schematic illustration of the technique used by Patterson (1976) to determine the shape of the auditory filter. The threshold of the sinusoidal signal is measured as a function of the width of a spectral notch in the noise masker. The amount of noise passing through the auditory filter centred at the signal frequency is proportional to the shaded areas.

From Moore (1997)

Looking at spectra



10dB-6kHz.wav;1 - Cool Edit

Analysis - 10dB-6kHz.wav;1

Cursor: 3946 Hz, -120.1 dB

Frequency: 461.95 Hz (A#4 -15)

FFT Size: 4096

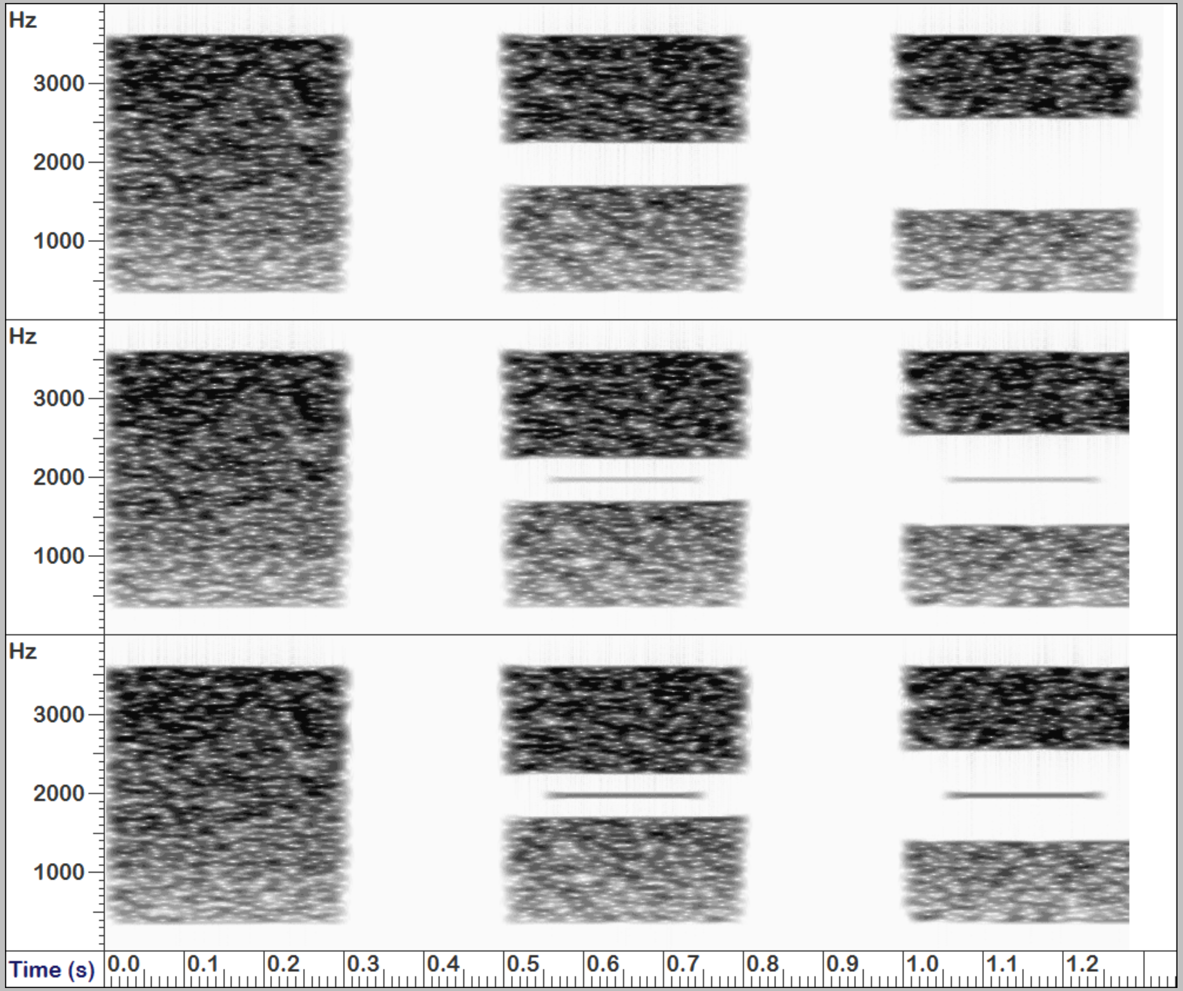
Blackmann-Harris

☑ Linear View

Range 120 dB

Notched noises

g=0.0 *g=0.15* *g=0.3*



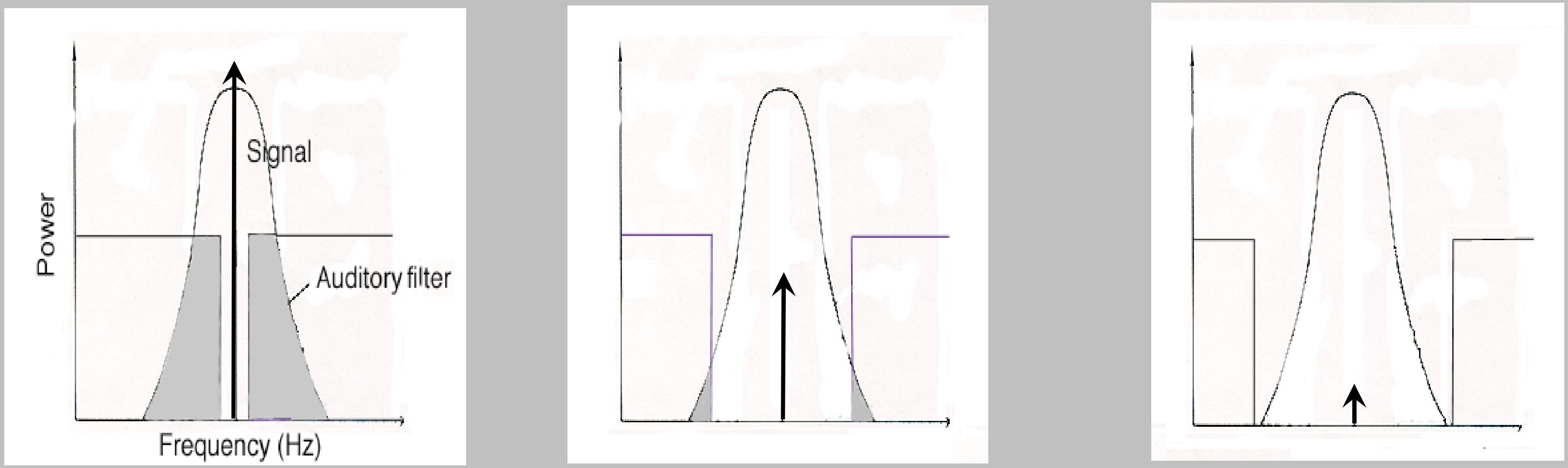
SNR= -∞ dB

SNR= 0 dB

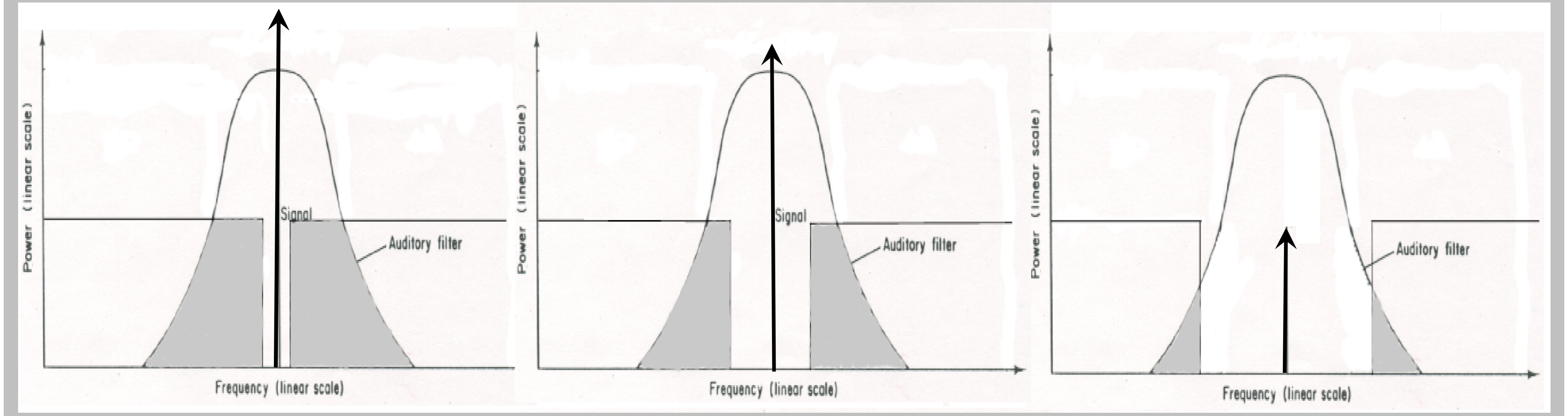
SNR= 10 dB

Narrow vs broad filters

Narrow filter

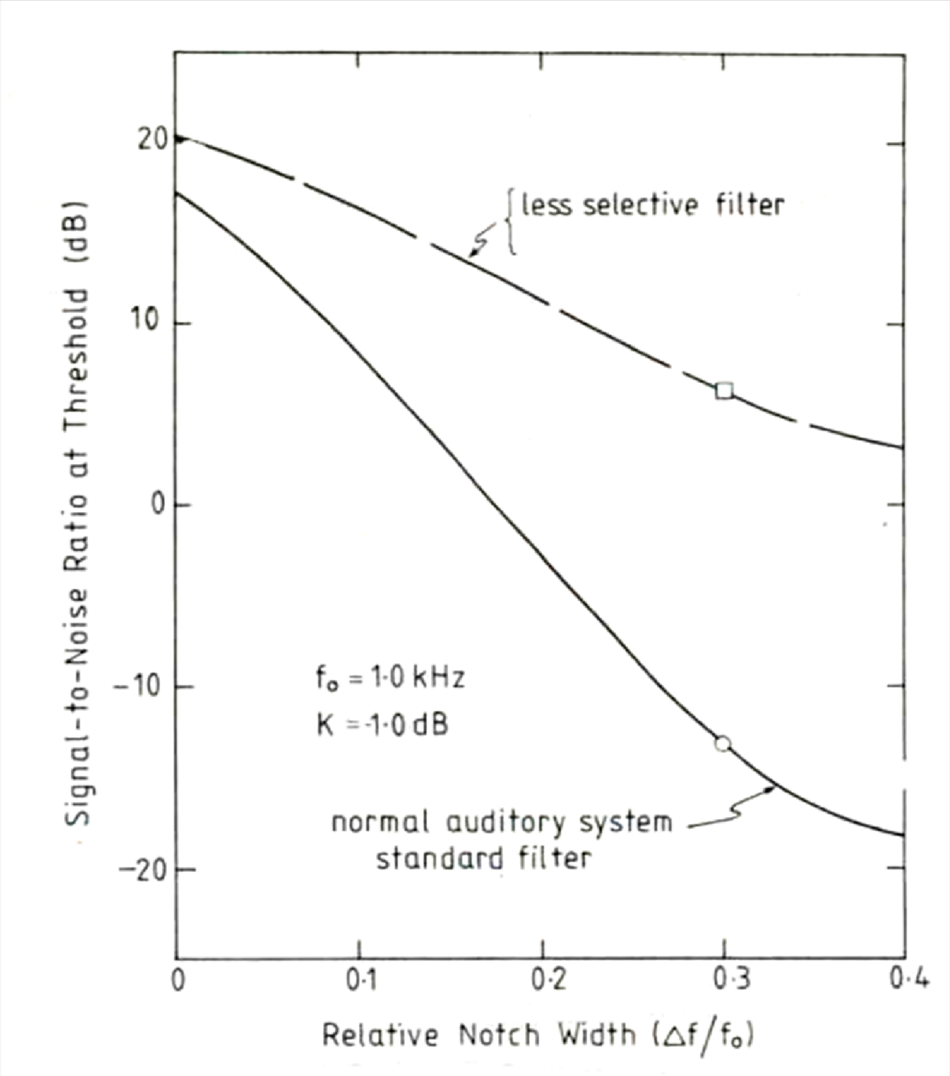


Broad filter



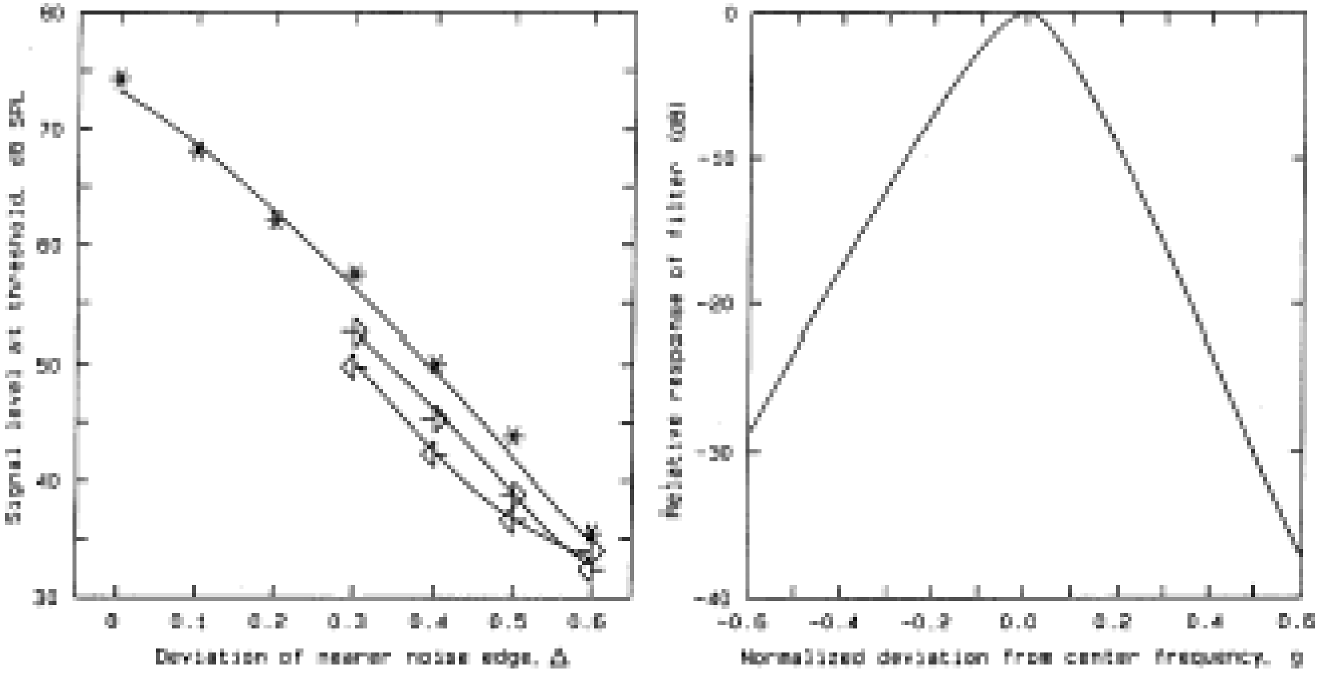
Notch gets wider ⟶

Thresholds at different notch widths

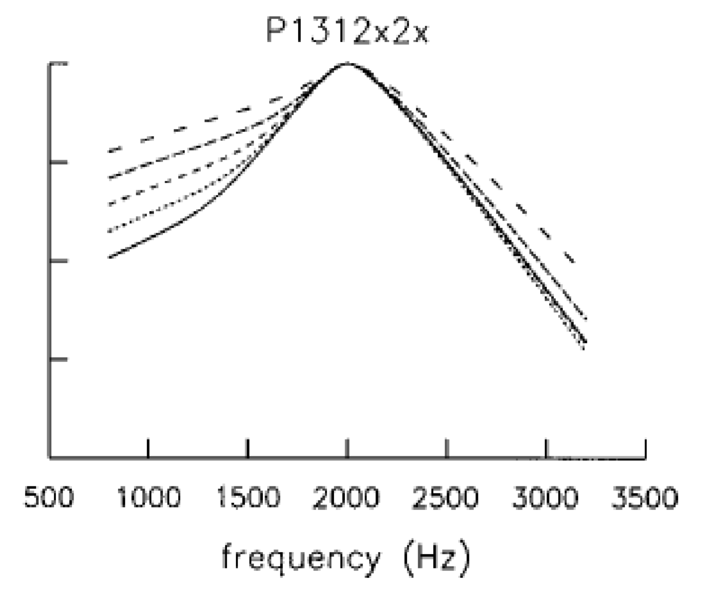


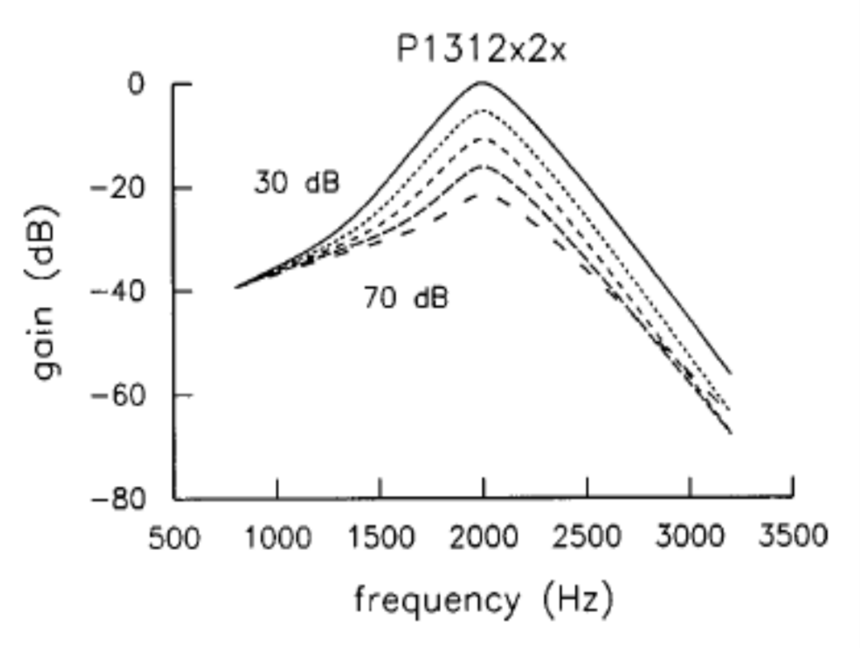
From Patterson et al. (1982)

Typical results at one level, and a fitted auditory filter shape (symmetric & asymmetric notches)

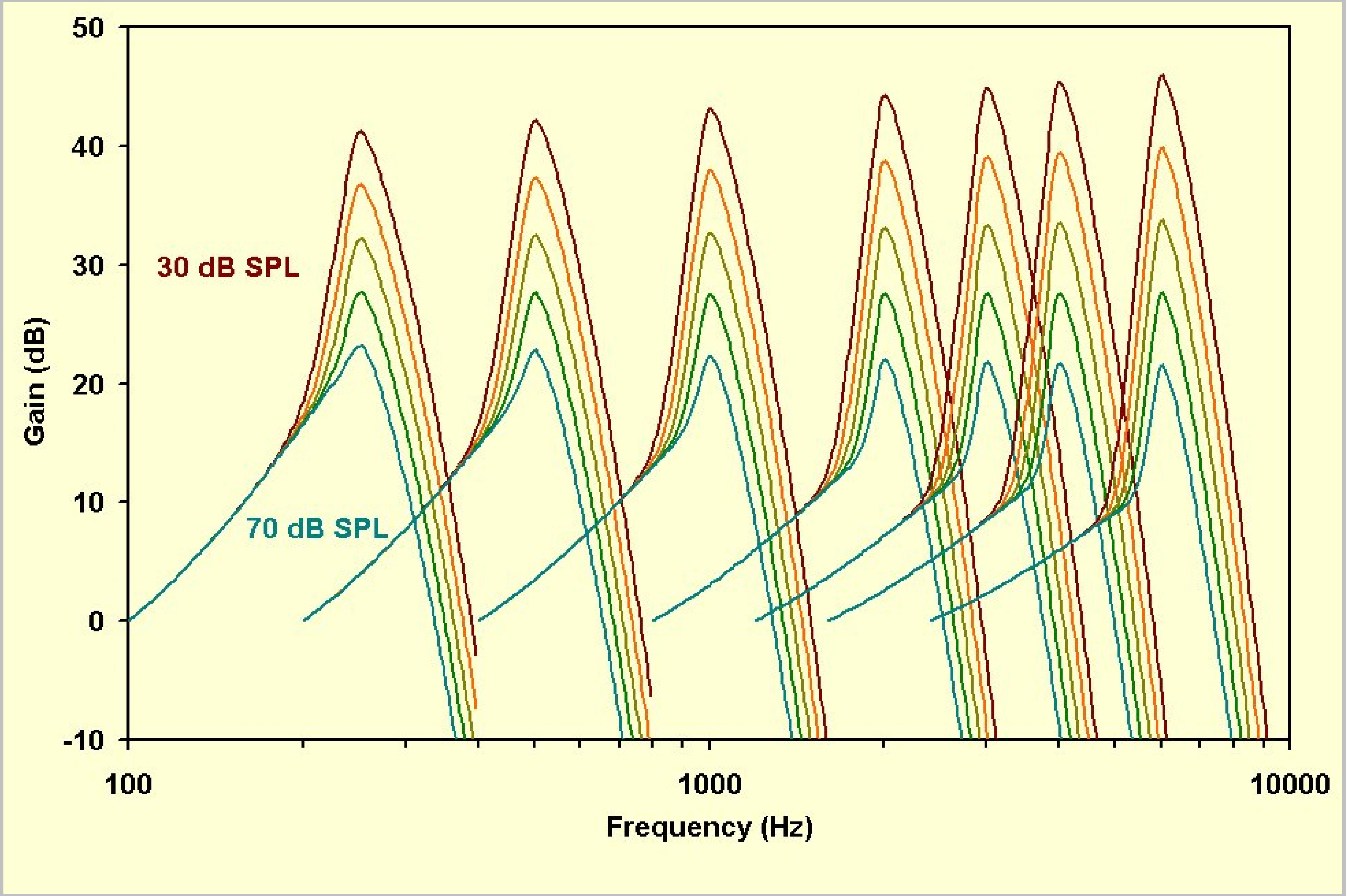


Now measure across level and assume filter linearity at frequencies substantially lower than CF

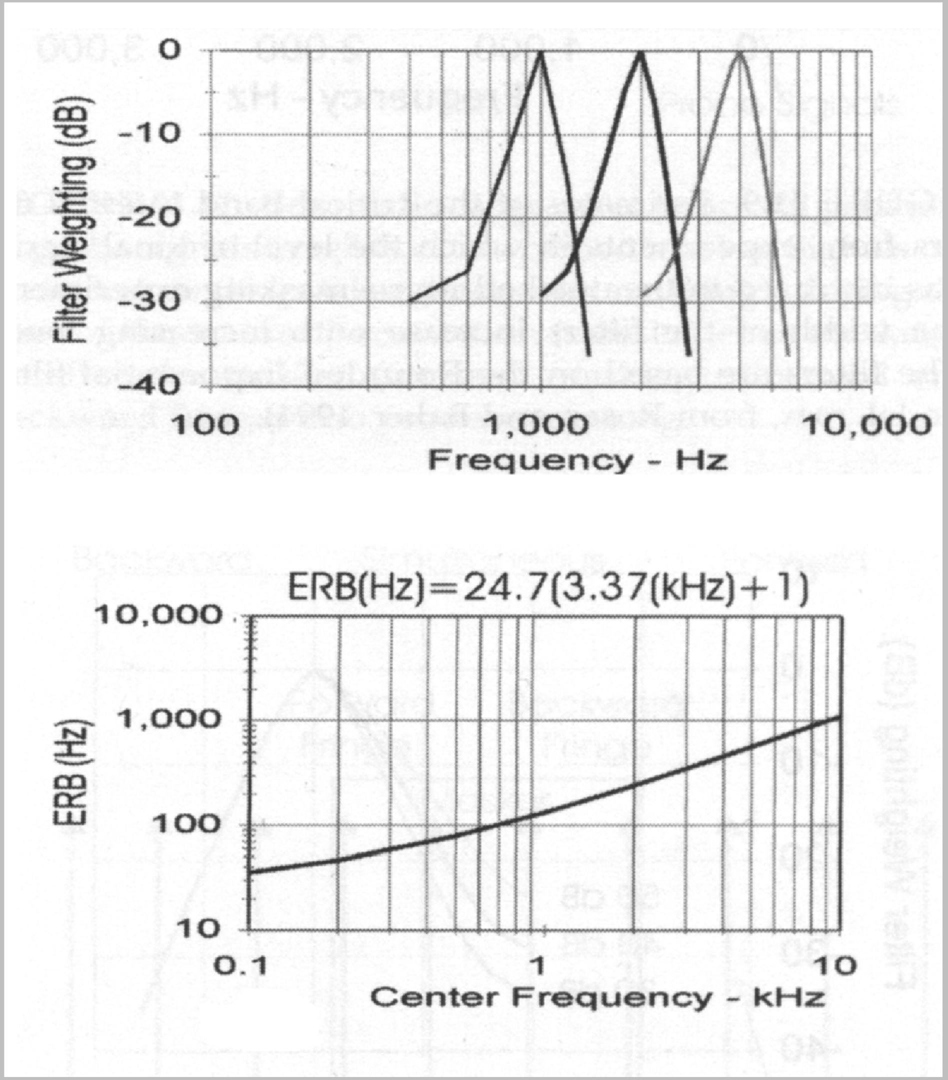




Auditory filter shapes across level & frequency: Note the asymmetry

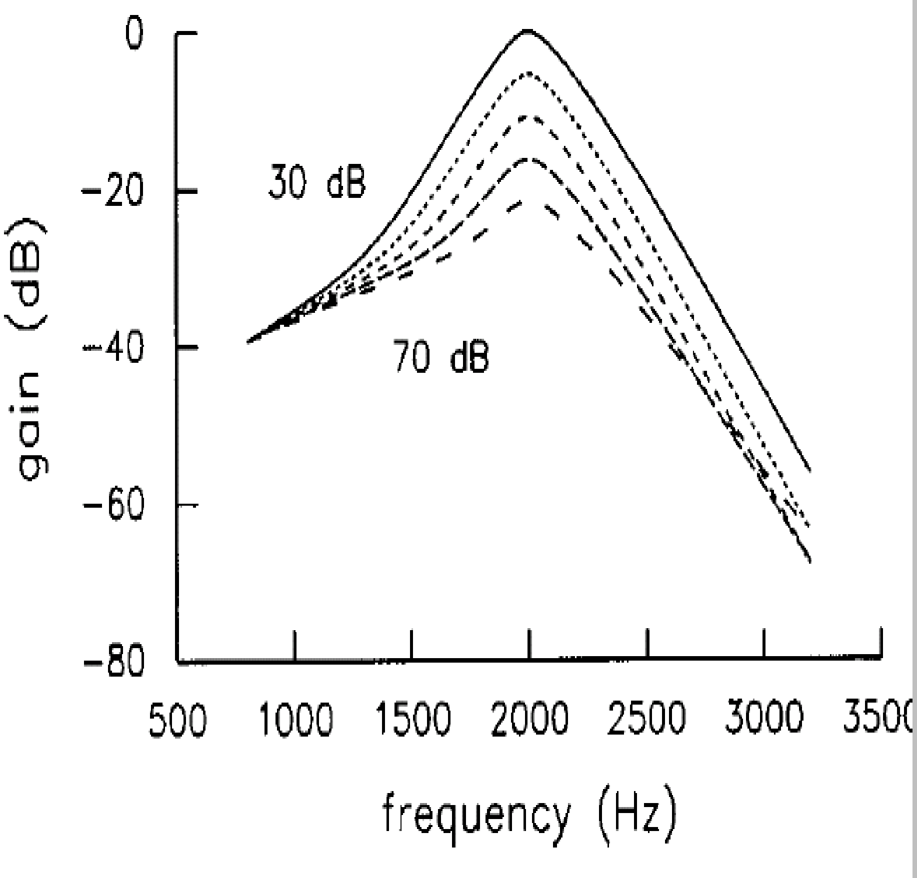


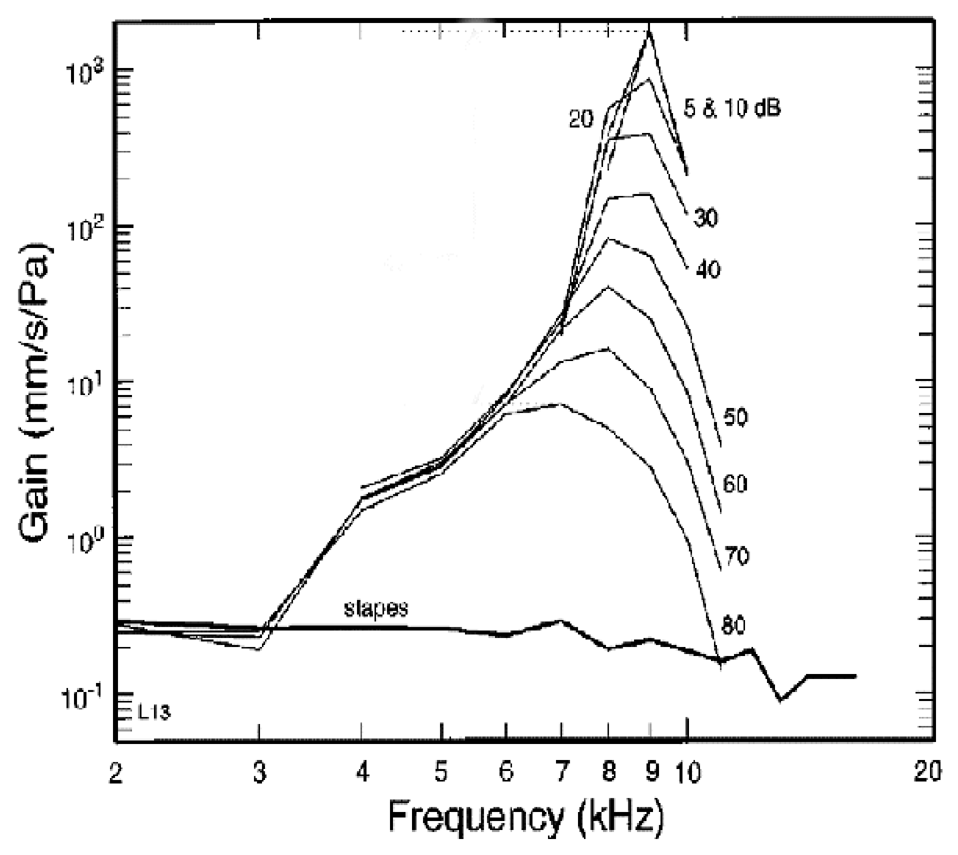
Auditory filter shape



- Filter shapes based on thresholds in notched noise (log frequency axis)
- Filter bandwidth increases with frequency (here as Equivalent Rectangular Bandwidth)
- BUT - filter shape also depends on sound level

Psychoacoustic estimates of auditory filter shape at 2 kHz as a function of pure tone SPL (left) & basilar membrane frequency response (right)





Rosen et al (1998)

Auditory filter **nonlinearity** at 2 kHz in normal hearing listeners

Ruggero et al (1990)

NB log frequency scale while left figure uses linear frequency scale.

Auditory filters

The cochlea acts like a bank of band-pass filters

- Each filter has a place (position on basilar membrane)
- At *basal* places band-pass filters have *high* frequency passband
- At *apical* places band-pass filters have *low* frequency passbands
- Filter bandwidth is larger for high frequency filters
- Filter shape also depends on sound level: higher levels lead to wider bandwidths

Importance of auditory filter shape

- Changes in auditory filter shape are a major factor in hearing loss.
- Auditory filter shape determines how spectral shape is represented in the auditory system
  - determines timbre (sound quality)
  - **critical for speech perception**

Spectral cues to speech sounds

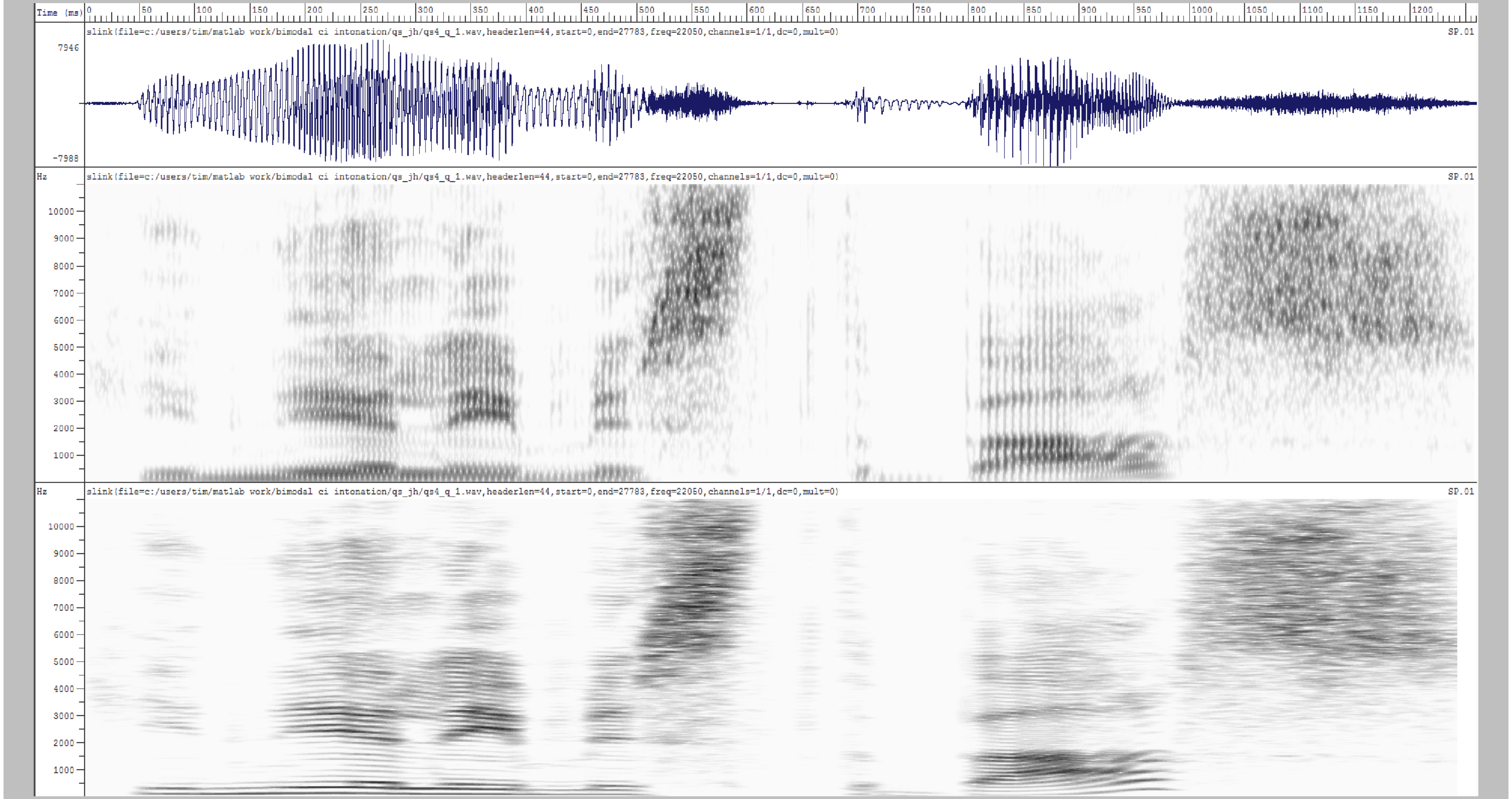
- Spectral shape (the distribution of acoustic energy over frequency) is crucial for the identification of both
  - Vowels (location of formants)

    and

  - Consonants (place of articulation)

Different ways of looking at speech

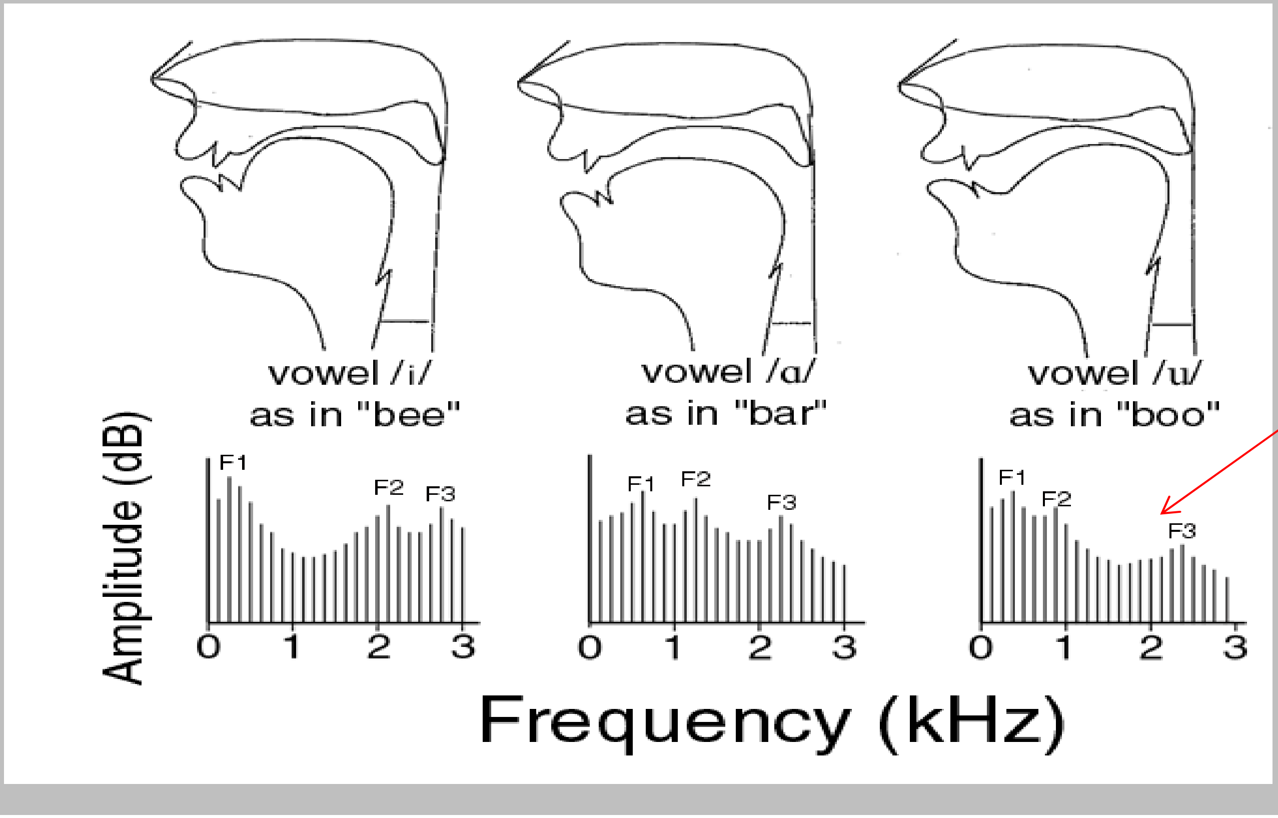
he n ear l y mi ss d the bu s



Vowels are cued by spectral peaks (formants)

Configuration of the vocal tract is reflected in oral resonances
- producing peaks in the amount of energy at particular frequency regions

F1 and F2 are main determinants of vowel identity



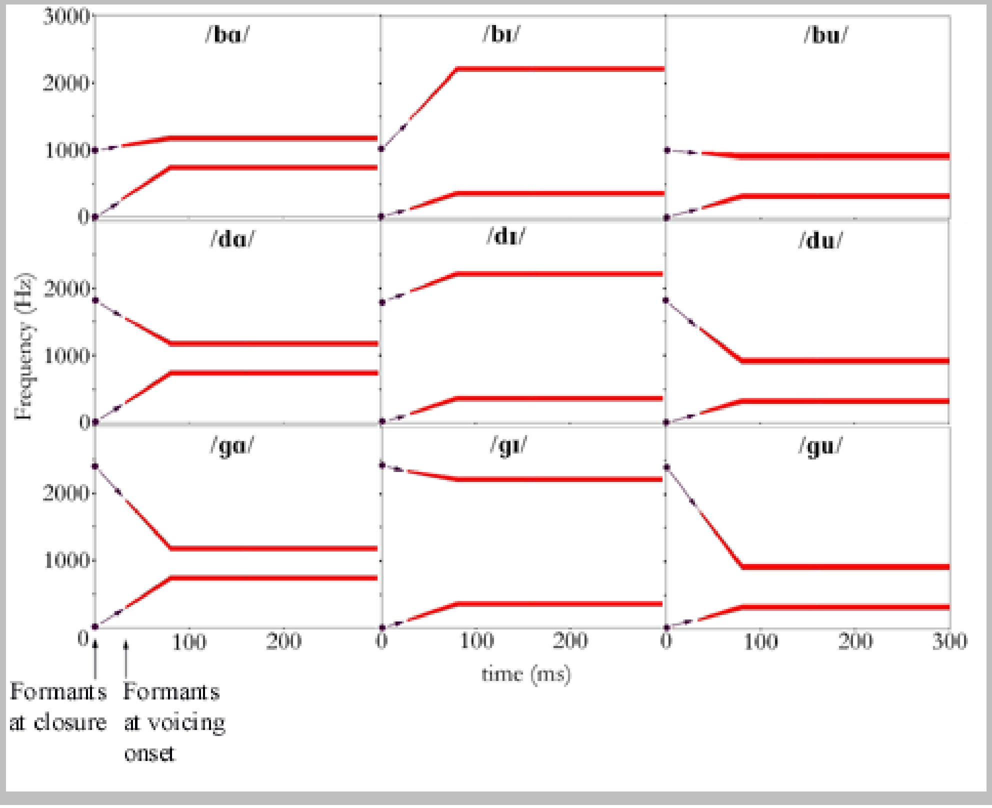
Voiced sound, source is periodic, energy only present at harmonics of F0

Cues to consonant place

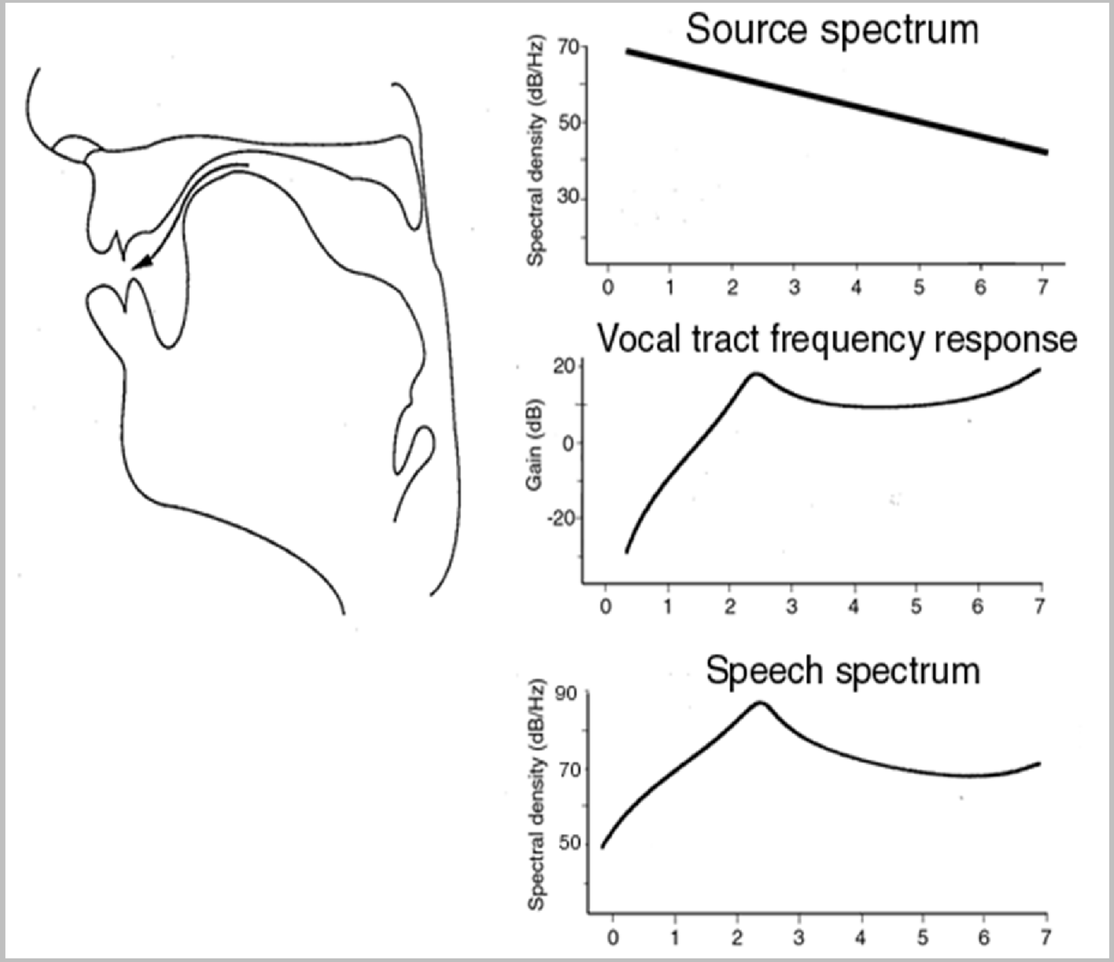
- Transitions of formants from consonant to vowel determine changing patterns of amplitude peaks in the spectrum

Transitions of formants F1 and F2 over time – from consonant opening into the following vowel.

Shown here for bilabial, alveolar, and velar voiced plosive consonants



Place of voiceless consonants



Fricative consonant ‘sh’

Location of spectral peak varies with place of articulation

e.g. ‘s’ made further forward, smaller front cavity would give higher frequency spectral peak

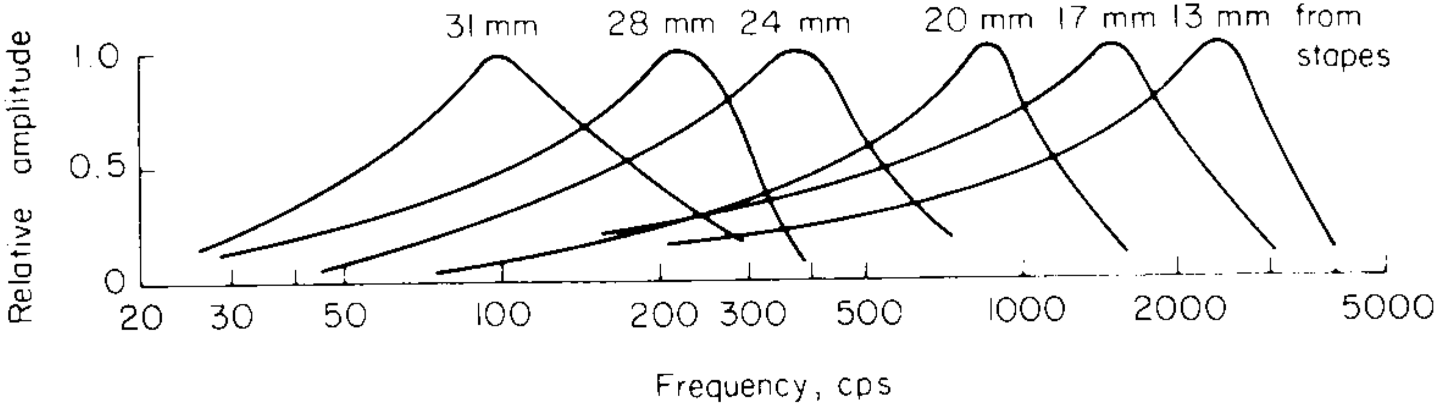
Coding of spectral shape

- Speech sounds vary in intensity over FREQUENCY and TIME
  - **Dynamic changes of acoustic spectrum**
- Auditory system acts as filter bank
- Acoustic intensity determines amount of neural activity (number of nerves firing and rate of nerve firing)
- Amount of neural activity varies over PLACE and TIME
  - **Dynamic changes of excitation pattern or auditory spectrum**

Frequency response of cochlear filters

- Frequency response — the amount of vibration shown by a particular place on the BM to sinusoids of varying frequency.
  - Input = many sine waves at different times
  - Measure at a single place on the BM for each sine wave

Figure shows frequency response of 6 different BM locations



Excitation patterns

- Excitation pattern — vibration pattern across the basilar membrane to a single sound.
  - Input = 1 sound (e.g. 300, or 200 or 100 Hz)
  - Measure at many places along the BM for each sound
- Related to a *spectrum* (amplitude by frequency).

Figure shows excitation patterns for four different tones

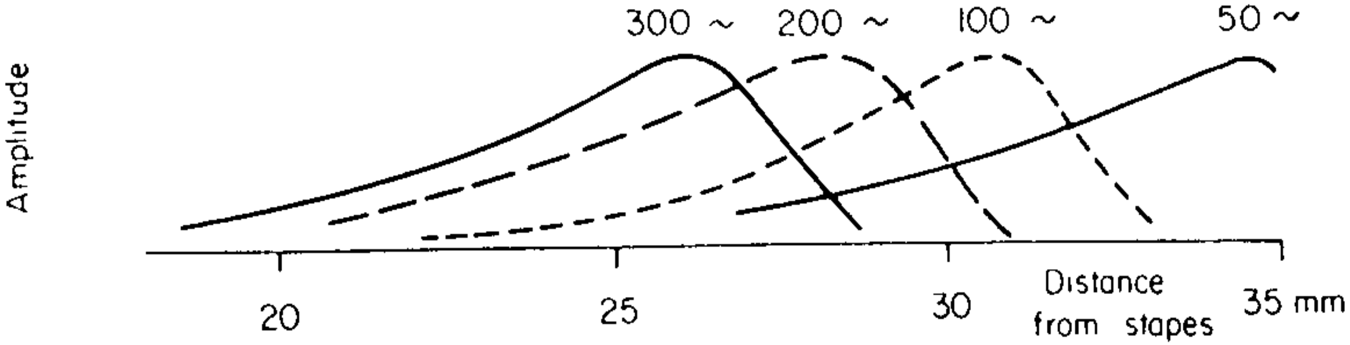
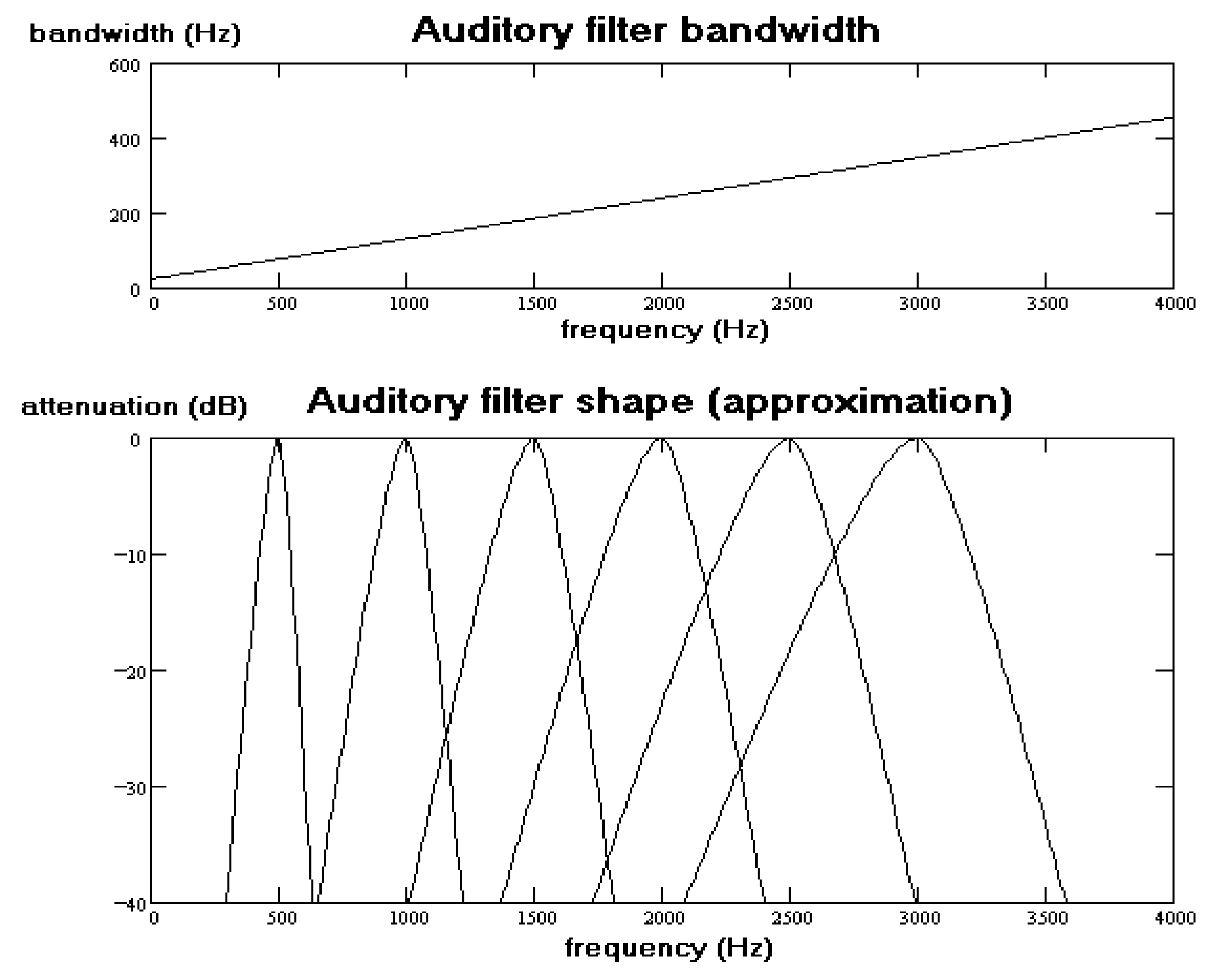
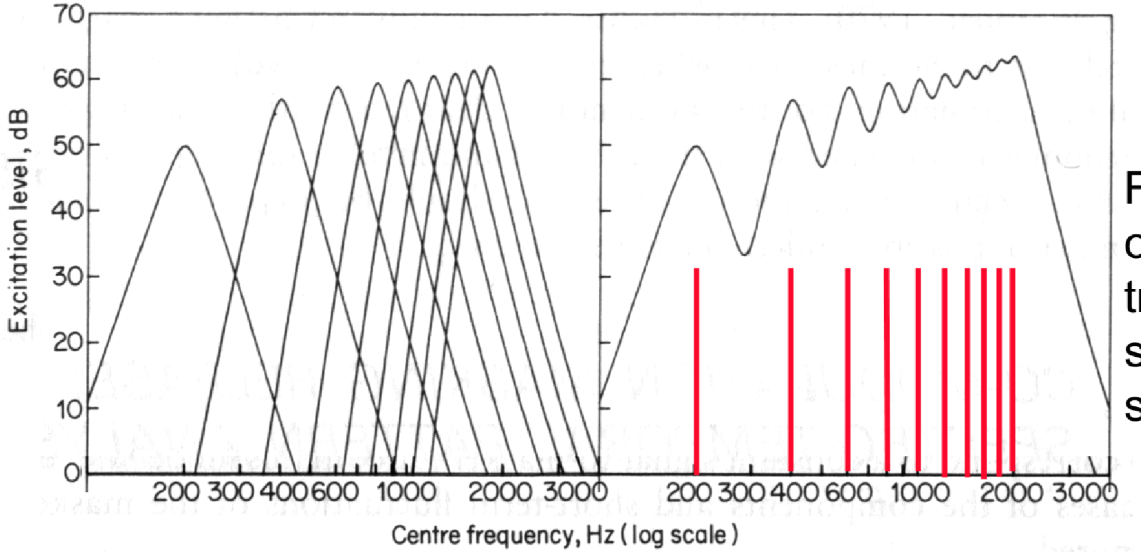


Fig. 3.8 Displacement envelopes on the cochlear partition are shown for tones of different frequency. From von Békésy (1960, Fig. 11.58).

Auditory filter bandwidth increases with frequency (while harmonics are evenly spaced)



Excitation patterns: complex sounds



First 9 harmonics of 200 Hz pulse train (red lines show acoustic spectrum)

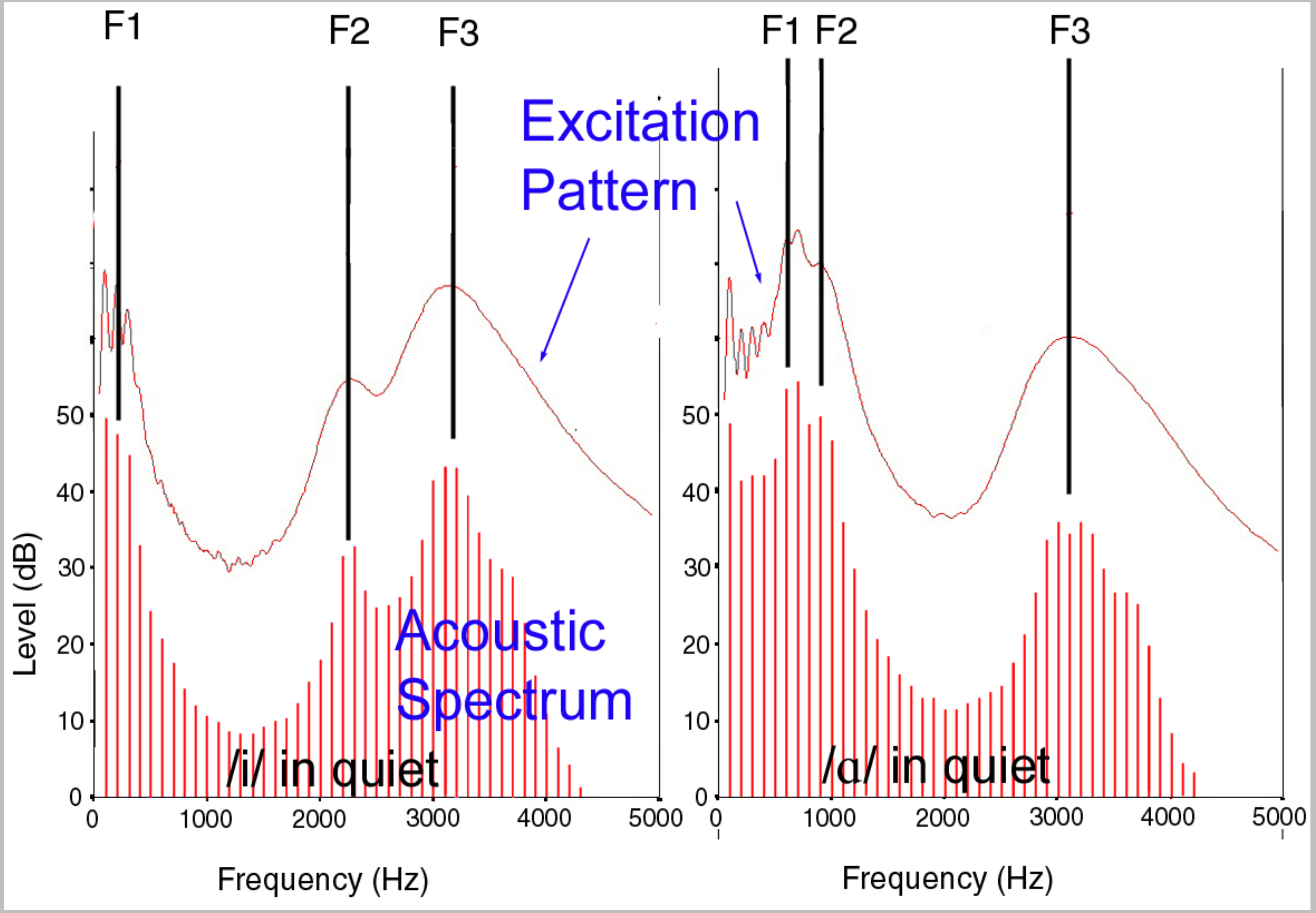
Excitation patterns of each component individually

Overall excitation pattern (harmonics all present together)

Auditory filter bandwidth increases with filter centre frequency

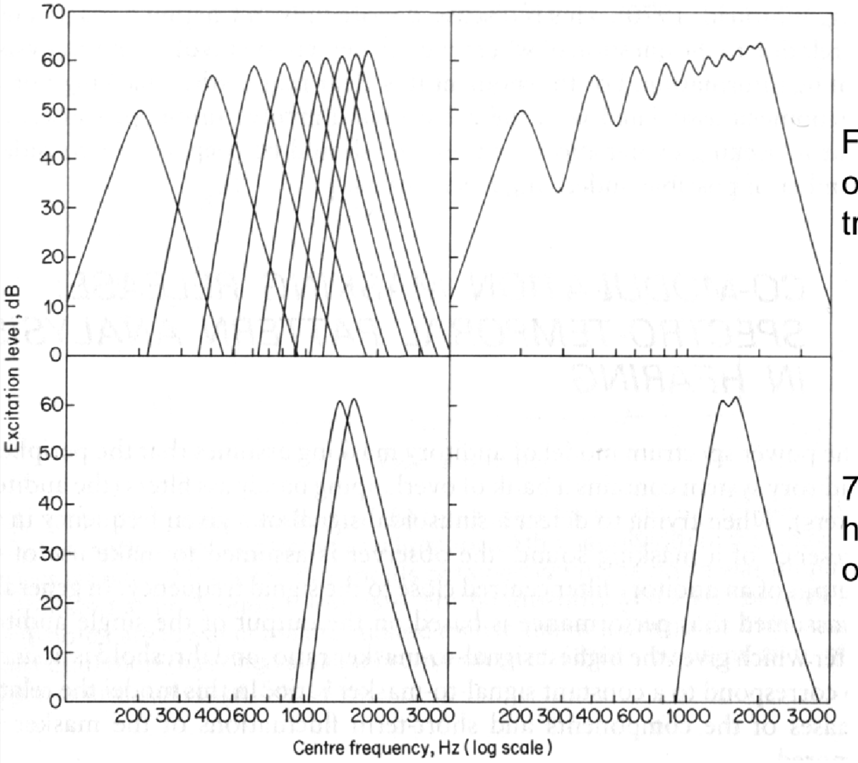
Lower harmonics are clearly resolved – above 1.6 kHz filter bandwidth is wider than 200 Hz spacing between harmonics and these higher harmonics are not resolved

Excitation patterns: vowels (100 Hz F_0)



Excitation patterns – sound spectra as activity across auditory filters

Excitation pattern of complex tone



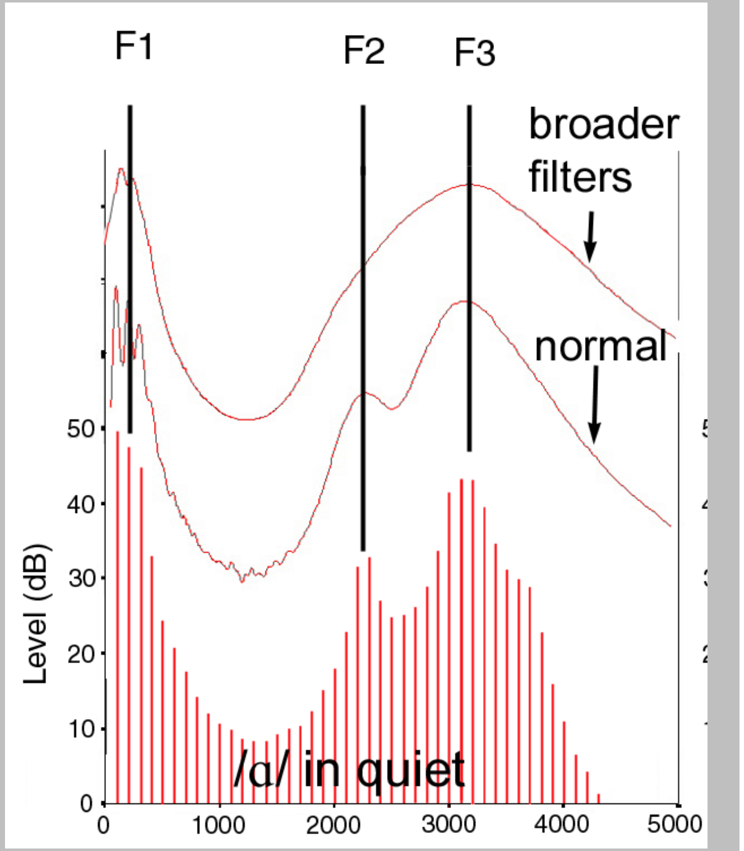
First 9 harmonics of 200 Hz pulse train

7th and 8th harmonics only

Excitation patterns of each component

Overall excitation pattern

Excitation pattern of vowel



Spectral features of voiced speech in excitation patterns

- Low order harmonics individually resolved
  - places of high excitation at harmonics near low formants cue vowel formant frequencies
- Harmonics above about the 8th are not resolved
  - formant frequencies above 8 x F_0 appear as broad peaks in excitation pattern
- Similar principles account for coding of spectral features of consonants but dynamics are much more important

Summary of auditory filtering and excitation patterns

- Masking studies reveal shape of filters
  - Filter width increases with frequency
  - Filter width increases with sound level
  - Filter width increases with hearing loss
- Auditory filtering provides a place coding of spectra as excitation patterns

Main points

- The ‘filters’ through which we listen to sounds are the filters established in the inner ear, in SNHL as well as normal hearing.
  - supported by the similarity between physiological and behavioural measurements

- The width of the auditory filter is an important determinant in how well we can hear sounds in quiet and in noise

The End