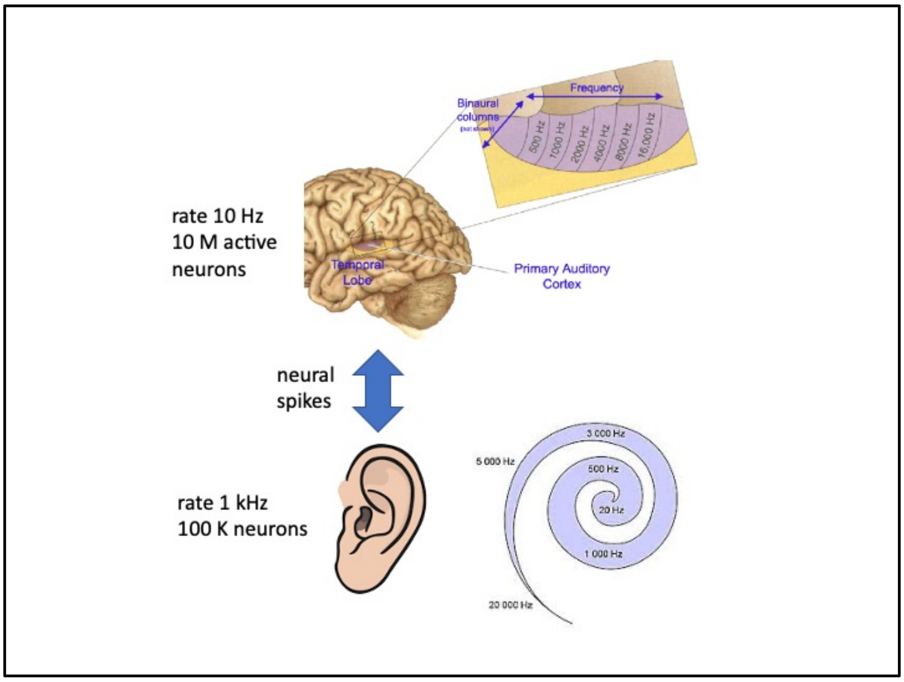
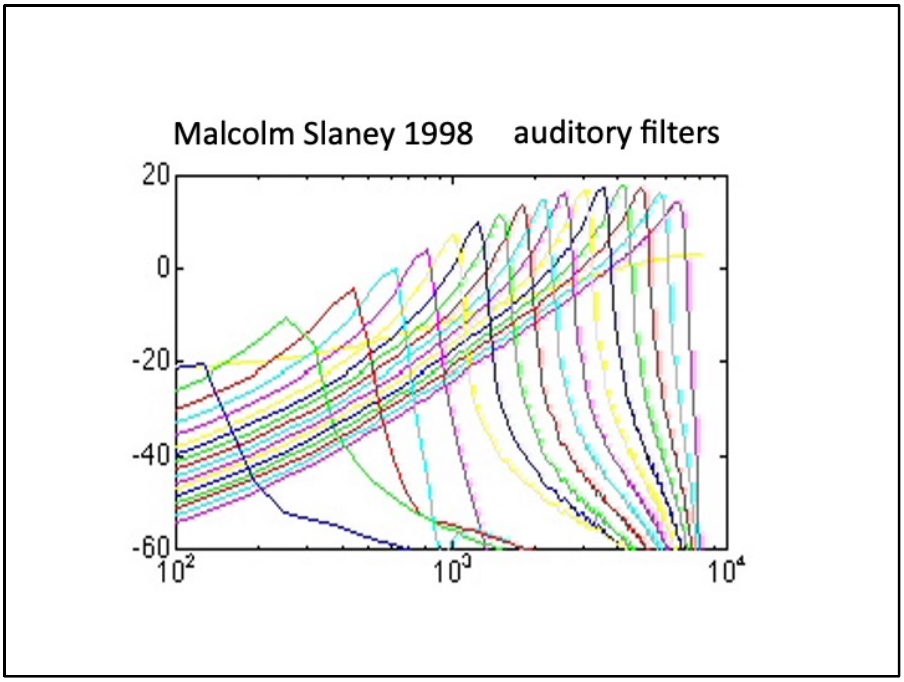
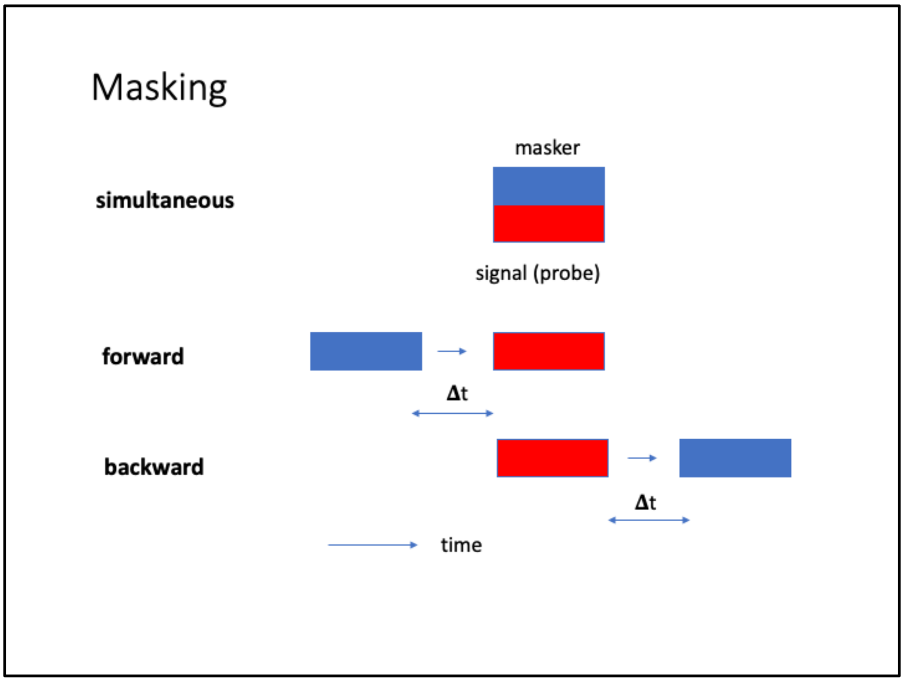
A BRIEF (AND SUBJECTIVE) SUMMARY OF SOME IMPORTANT PHENOMENA IN HUMAN HEARING



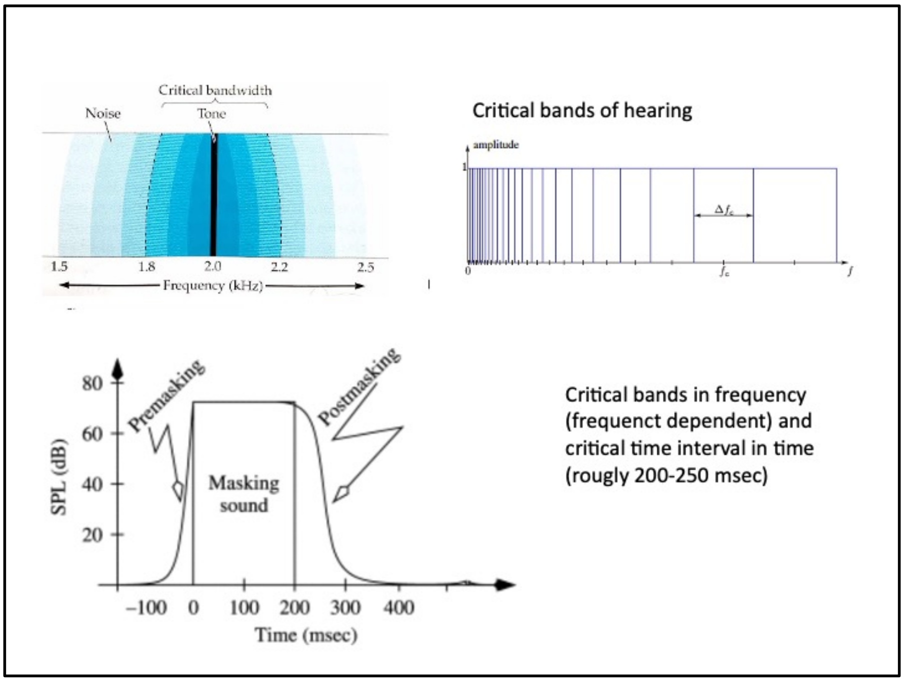
Hearing periphery (ear) divides signal into different frequency chanels. FOring rates on the audotory r=nerve are of the order of 1 kHz, there is less that 100K figers in the nerve,
Info from the periphery goes throu several processing stages and end in the auditory cortex. Frequency channels formed in the periphery are maintained all the way to the cortex. FIring rates in the cortex are of the order of 10 Hz. Number of active neurons at any given time there is extimates to less that 10 milliom (only about 5% of the total corical neurons).



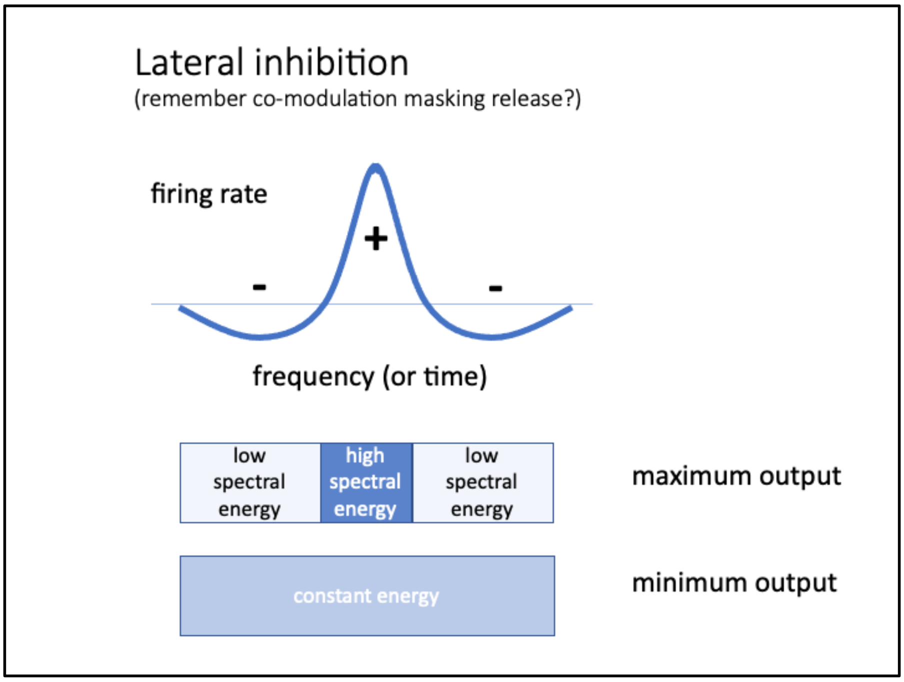
The most important function of the periphery is to bandpass filter the signal..The implied cortical filters have similar shapes on log frequency scale (i.e., they are getting broader towrds higher freuencies. They are asymmetric, slopes towerds lower frequncies are much shallower to account for the upward sperad opf masking (low frequencies influence higher frequencies and not vive versa).



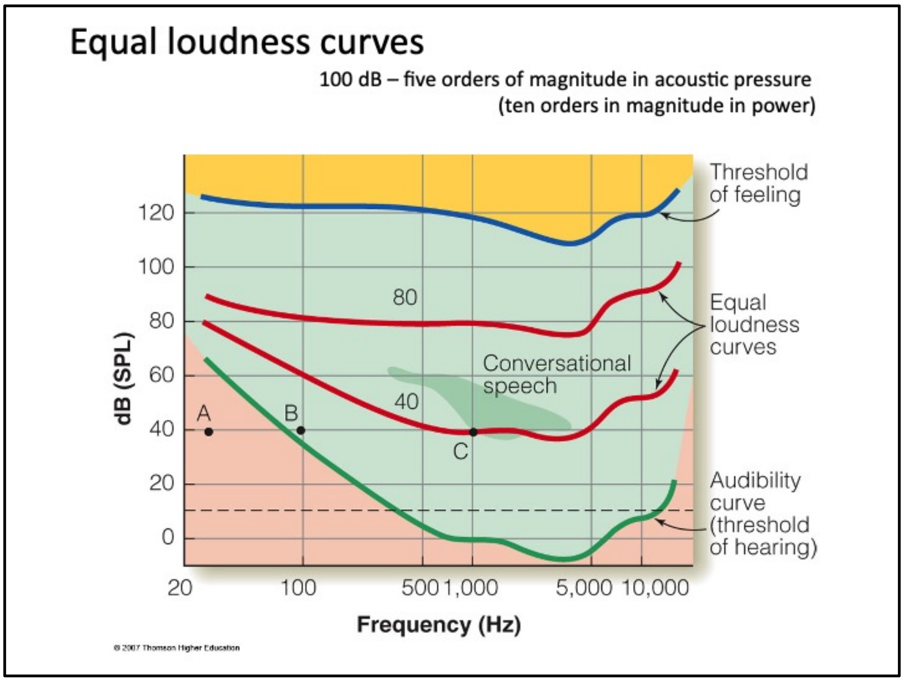
Masking os anither important phenomenon. When ooccuring at the same time as the signal (probe) has obviously larges effect on the probe threshold (simultaneous masking). However, when the masker is presented befor the probe, the masking effect is still there, as long as the interval between the msaker and the probe is not too large (forward masking). Even the masker presented after the probe has some effect on the probe threshold (backward masking). The forward masking last longer (about 200-300 msec.) The significant backward masking interval is shorter (not more than 40 msec).



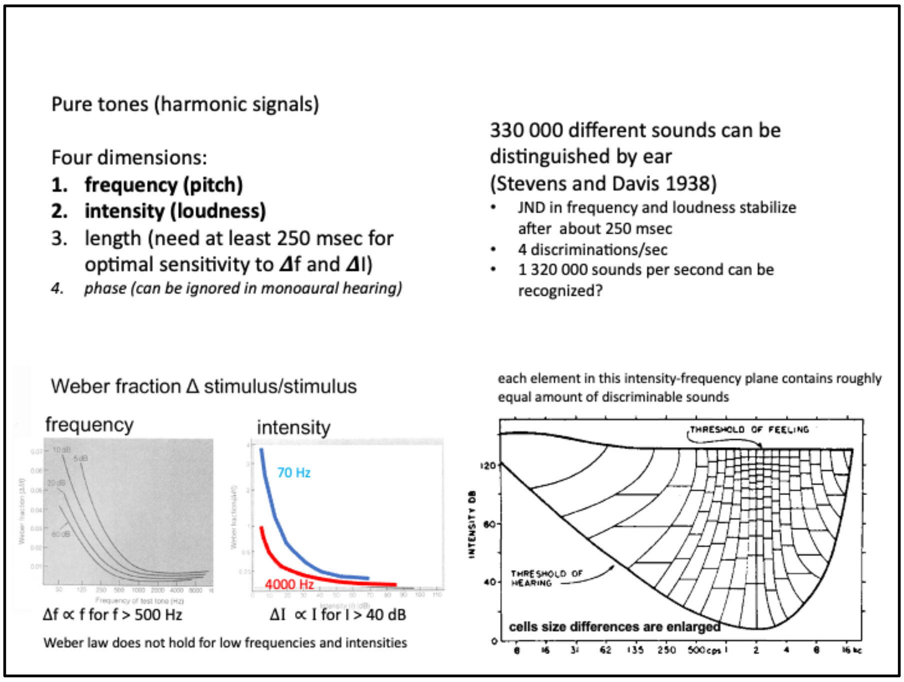
The simultaneous masking happens over critical bands (which are frequency-dependent) and over the critical time interval of about 200-250 msec. Masking implies than only the components within these critical time and frequency intervals interfere with a detection od signal within these itervals, The component outside these intervals are heard but do not interfere with the detection of the signal within the interval.



Important prenomenon we talked about (when talking about the two-tone suppression) is the lateral inhibition, which show that responses of a system are less sensitive to constant stimuli and enhances stimulti with a particular energy distrubution. We talked mostly about lateral inhibition in frequency but it also exists in time. It is one of the most important mechanism seen in perceptual systems.



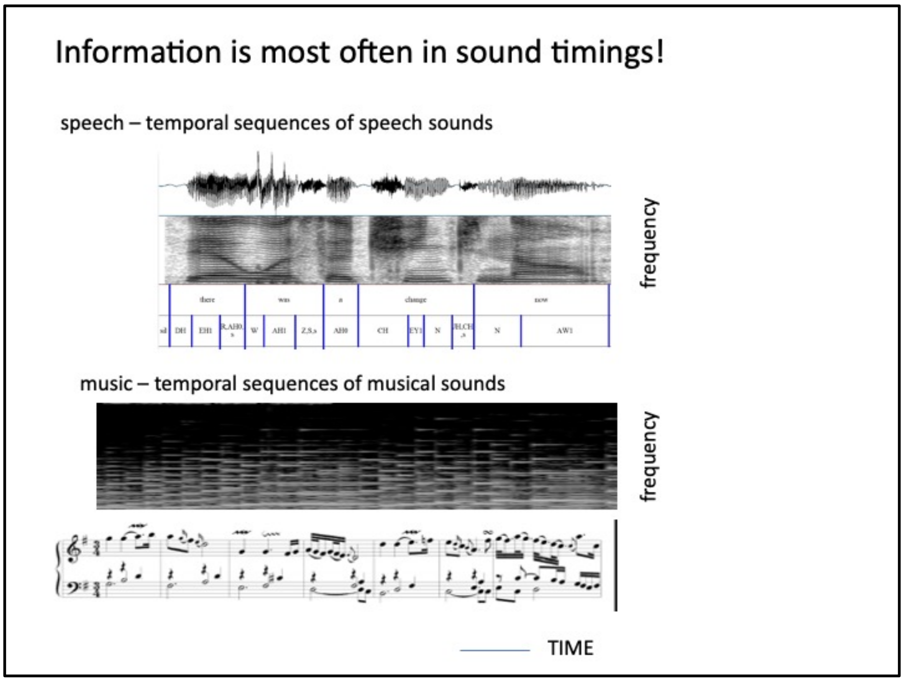
Equel loudness curves show how the intensity of the signal needs to change with frequency to maintain the constant loudness of the sound. It is seen that hearing is less sensitive below 600-700 Hz and above 5 kHz. Intensity range of human hearing is around 100 dB (which implies 10 orders of magnitude in acoustic power).



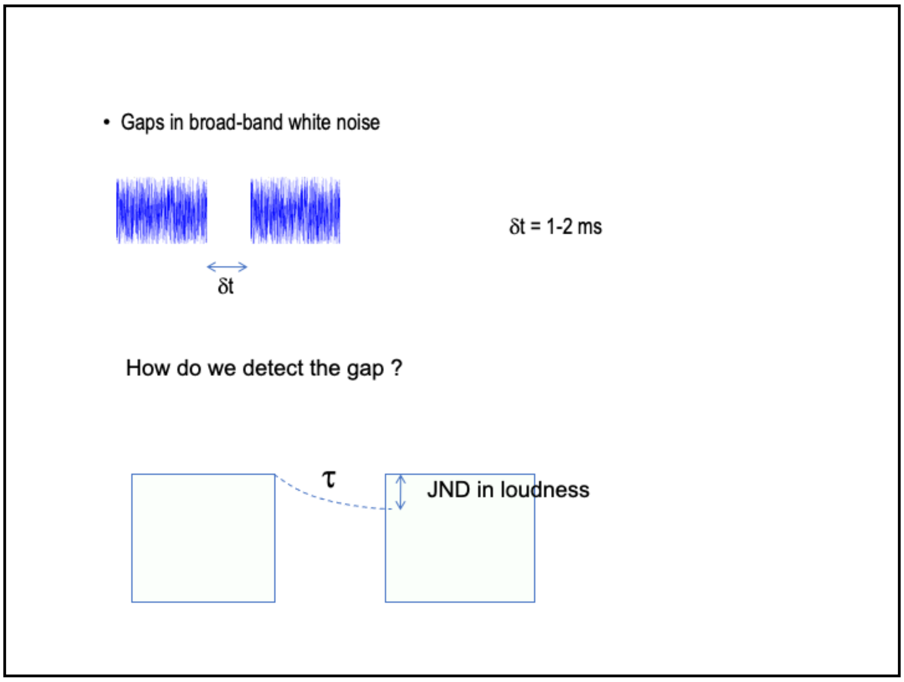
For sounds longer that 250 ms and more intense than 40 db (but less that 100 dB) the JNG in loudness is roughly 1 dB and JND in frequency is rougly 0.2 % of the frequency. Knowing the JNDs for both the intensity and the frequency allows for estimation of a total number of distinguishable sounds – this number is similar for both the audition and for the vision (this number would represents about 18 bits of information every 200 ms).

Weber fractions for frequency Δf and for sound intensity ΔI and proportional to frequency of the sound for frequencies higher than about 500 Hz and for intensities higher that about 40 dB, both for sounds longer than about 250 ms. For lower frequencies and intensities and for shorter sounds the fractions increase considerably (near miss of Weber's law and the effect of sound critical interval). When creating areas of equal discriminable sounds, the whole hearing area is most crowded in the center of the range of hearing frequencies and towards higher intensities. In principle, more than 1.3 million sounds should be discriminable within 1 second.` However, one has to be careful with such statements. The **discrimination** is different from the **recognition**. The fact that hearing can tell difference between two sounds does not mean that these two sounds can be independently recognized. For the discriminations, two sounds need to be heard next to each other. For recognition, one needs to recognize the sound which is heard – and human listeners are notoriously bad in that !

TIME

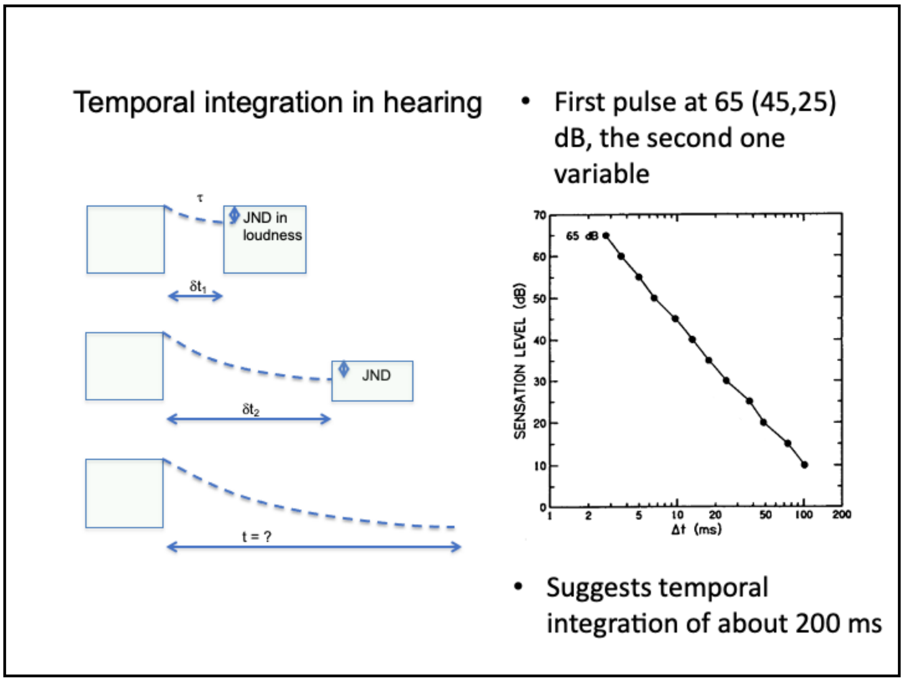


Information in both speech and music is it timing of speech or musical sounds

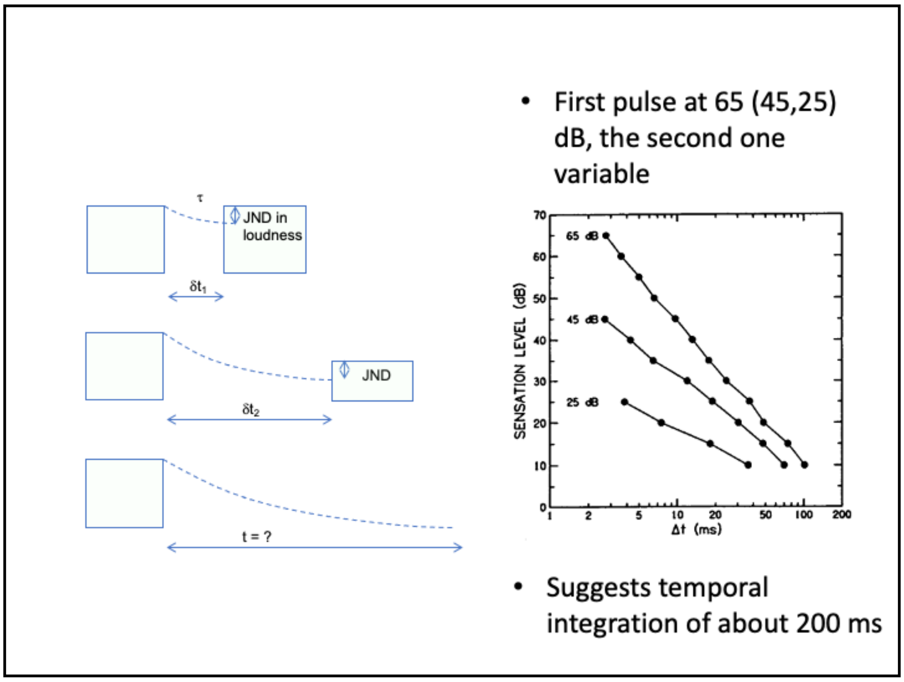


Hearing is more sensitive to gaps in broad-band noise. It may be that in this case, all frequency channels of hearing are used for the detection (since the signal has all frequency components and all are interrupted by the gap).

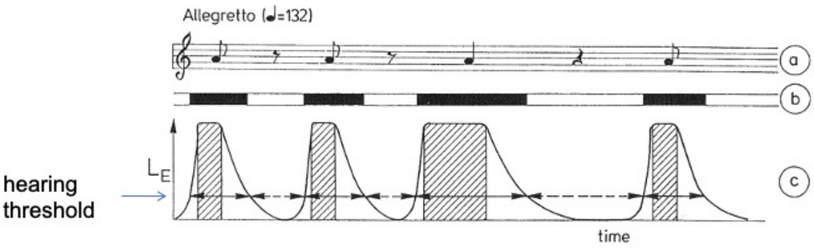
How is the gap in the signal perceived? Here is one possibility: After the signal ceases, there is a gradual decline in sensation caused by the preceding signal. When the signal is re-introduced, the sensation increases again. If the gap is long enough and the sensation from the signal before the gap decreased significantly (more that JND in loudness between the declined sensation and the sensation from the re-introduced signal, the gap is perceived.



The gap is perceived when the decline in the intensity of the first sound (shown here as dashed curve) is such that there is perceived discontinuity when the second sound comes. This happens when the difference between the fading of the first sound and the amplitude of the second sound represents the just noticeable difference (JND) in intensities. Weaker the second sound is, longer it takes for the effect of the first sound to decay. The effect of the first sound dies away completely after some time, which can be extrapolated from the perceived gaps for different levels of the second sound. Think about similarities with the effect of forward masking which we discussed earlier.

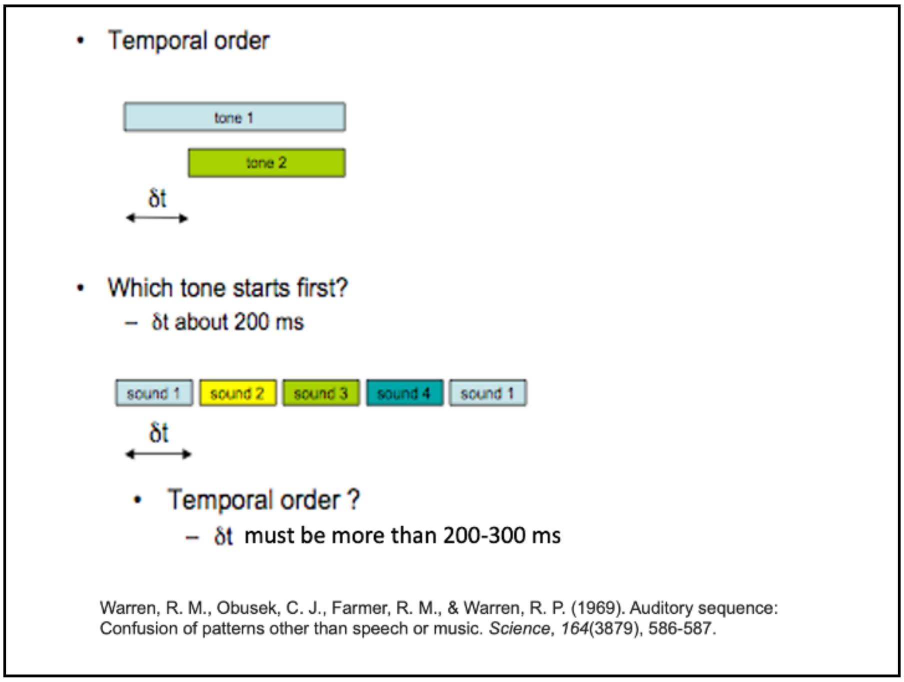


The gap is perceived when the decline in the intensity of the first sound (shown here as dashed curve) is such that there is perceived discontinuity when the second sound comes. This happens when the difference between the fading of the first sound and the amplitude of the second sound represents the just noticeable difference (JND) in intensities. Weaker the second sound is, longer it takes for the effect of the first sound to decay. The effect of the first sound dies away completely after some time, which can be extrapolated from the perceived gaps for different levels of the second sound. Think about similarities with the effect of forward masking which we discussed earlier.



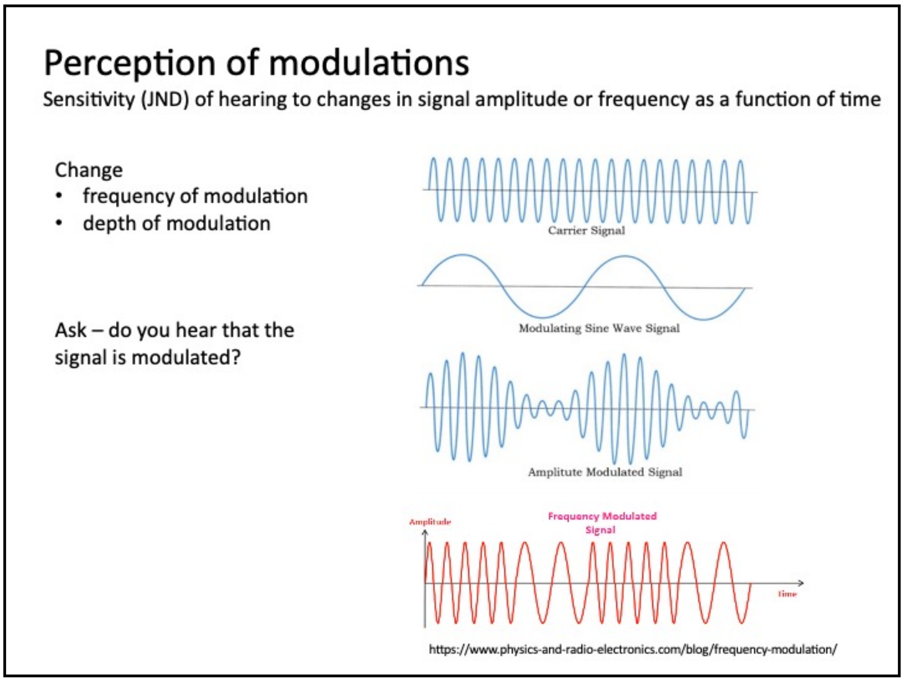
b) perceived

c) physical length of tones

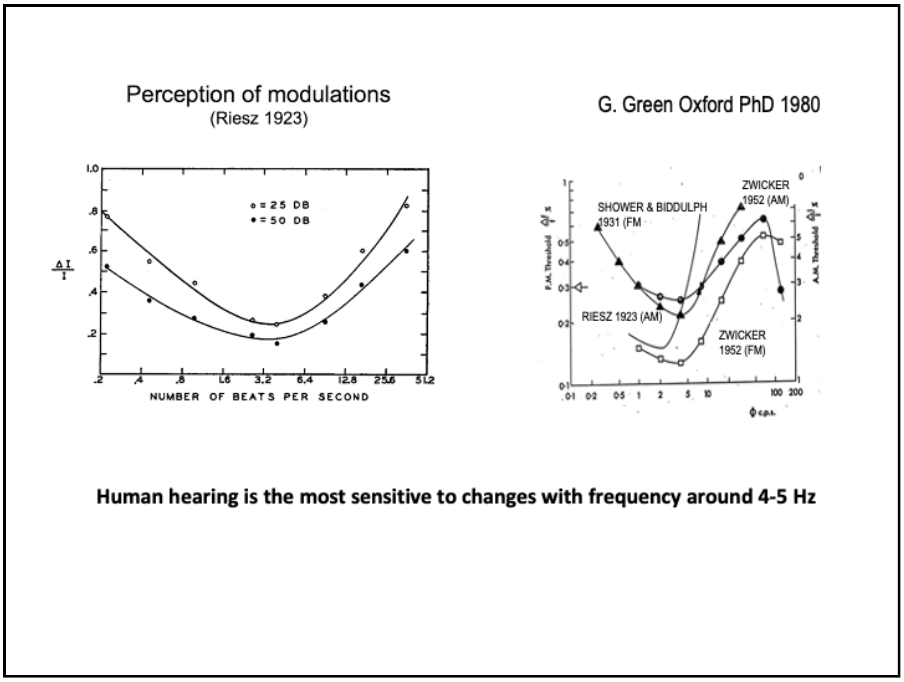


When the question is how long the time offset between two signal needs to be for the subject to be able to decide which tone started first, we her very similar number – about 200 ms.

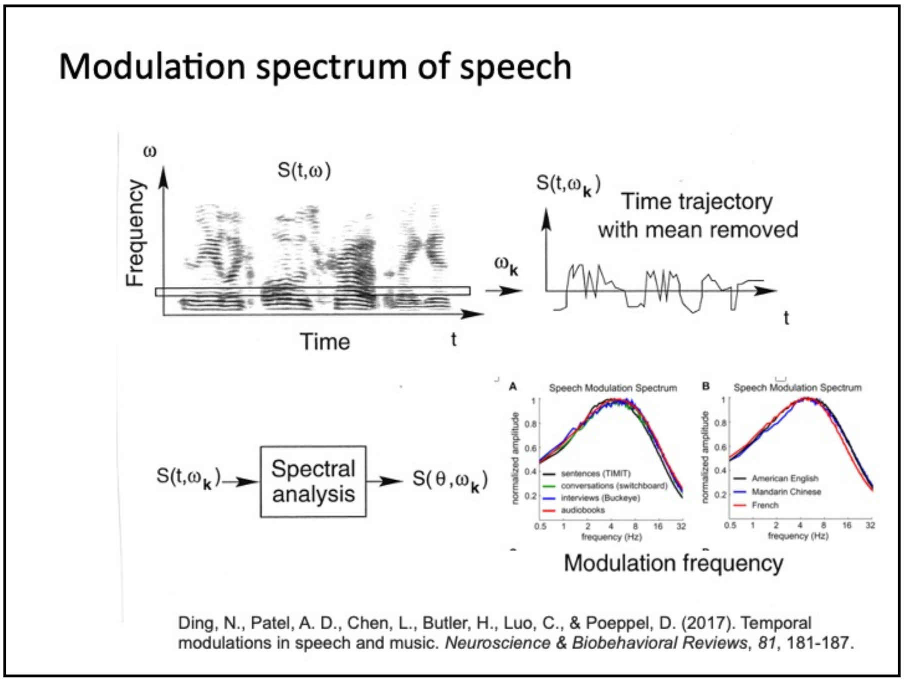
Similarly, it is hard to determine temporal order of several different (simple) stimuli, unless the stimuli are longer that about 200 ms.



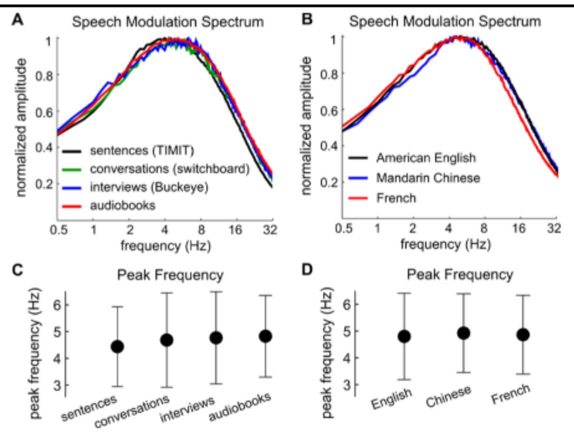
Weber's law for signal intensity approximately holds only for "normal" signal intensities above about 40 dB (where it is about 1 dB). For weaker sounds, the Weber fraction is much larger (as much as 40 times larger!)



Human hearing is most sensitive to modulations around 4-5 Hz (period 200-250 ms).

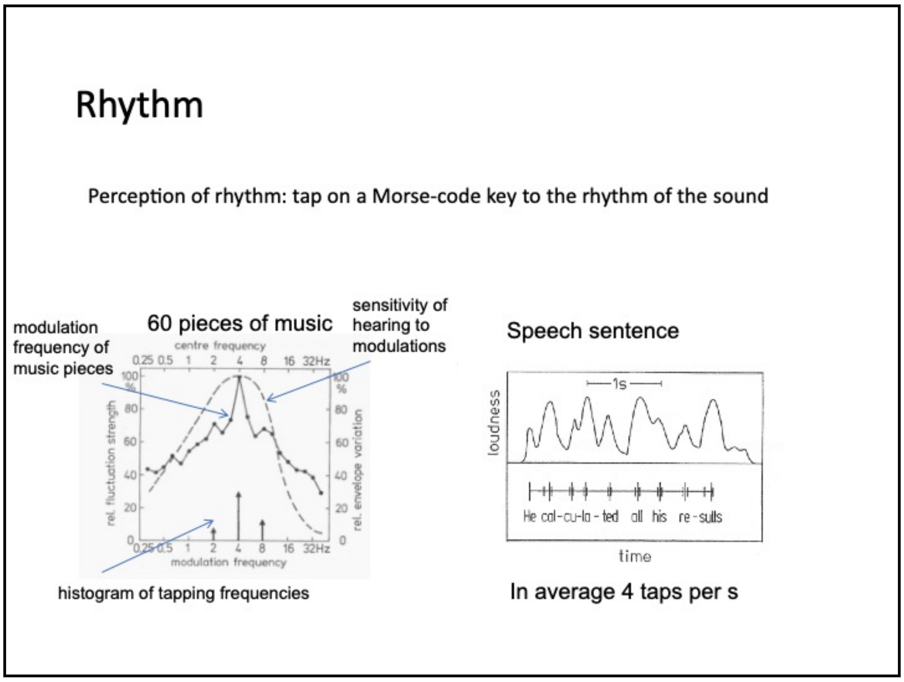


Modulation spectrum of speech can be computed by computing spectrum of temporal evolution of spectral envelope at a given frequency. The modulation spectrum of speech peaks at 4 Hz (where the sensitivity of human hearing to modulations is highest). Dominant modulations in speech were measured for several languages and for several different speech materials. Amazingly, amlitudes if modulationa were amazingly similar, quite independent of the language used or the speech material in which the modulation amplitudes were evalated.

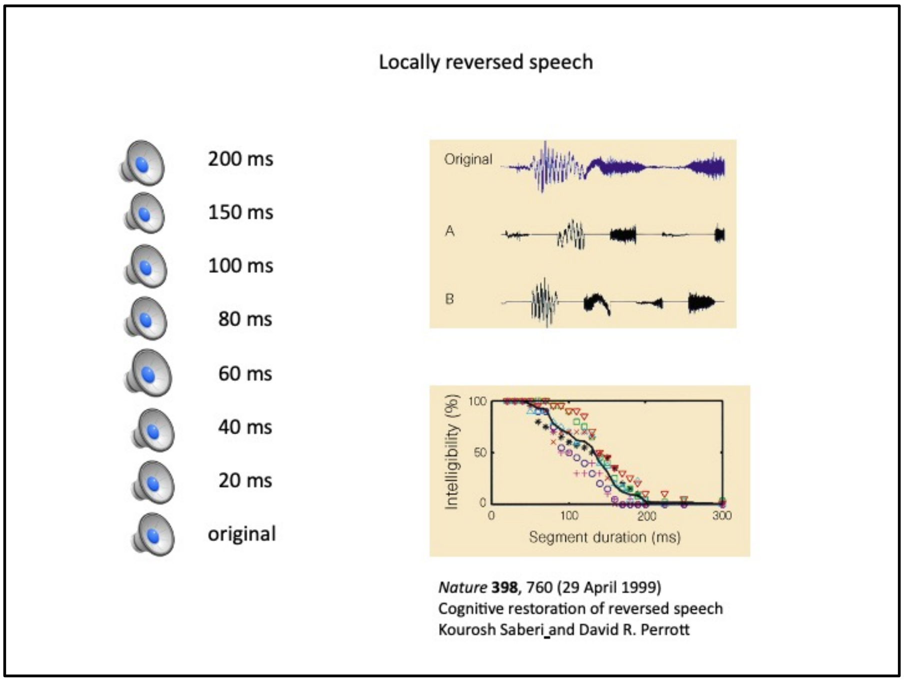


Ding, N., Patel, A. D., Chen, L., Butler, H., Luo, C., & Poeppel, D. (2017). Temporal modulations in speech and music. *Neuroscience & Biobehavioral Reviews*, *81*, 181-187.

Dominant modulations in speech were measured for several languages and for several different speech materials. Amazingly, amlitudes if modulationa were amazingly similar, quite independent of the language used or the speech material in which the modulation amplitudes were evalated.



The 4 Hz modulations dominate in both music and speech.

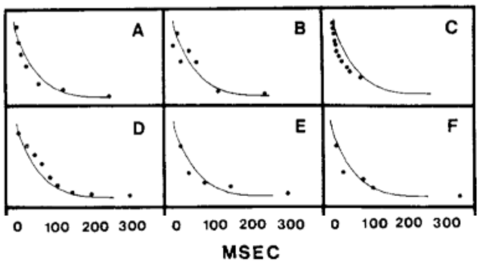


When subjects listen repeatedly to locally time-reversed sentences with moderately long windows (100 ms), they report that previously unintelligible words become clear. For longer time reversal towards 200 msec, the speech becomes unintelligible.

Coincidences?

f = 1/T => **4-5 Hz**
T = 200-250 ms

temporal integration in threshold of intensity, perception of loudness, threshold of frequency, perception of pitch, temporal order, perception of gaps, temporal masking.....



- A-forward masking
- B-backward masking
- C-gap detection
- D-overestimation of short burst duration
- E-loudness decrement
- F-JND in frequency

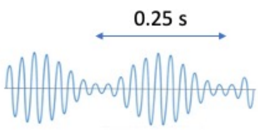
N. Cowan, On Short and Long Auditory Stores, Psychological Bulletin 1984

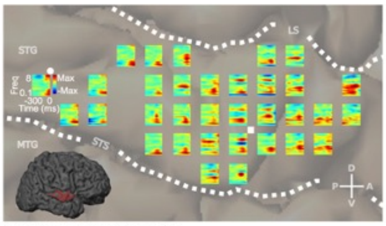
COllecting data from different perceptual experiments show similarities of time constants in different experiments. All hese works indicate that perception stabilizes only after more that 200 ms if acosutic signal is seen.

Length of a temporal buffer in hearing?

4-5 Hz implies 200 - 250 ms period of change

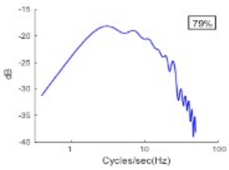
4 Hz modulation

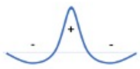




Patrick W. Hullett et al. J. Neurosci. 2016;36:2014-2026

1st principal component
(79% of variance explained)
of temoral component of
2619 spectro-temporal
receptive fields of ferrets.
Mahajan et al, JASA 2019





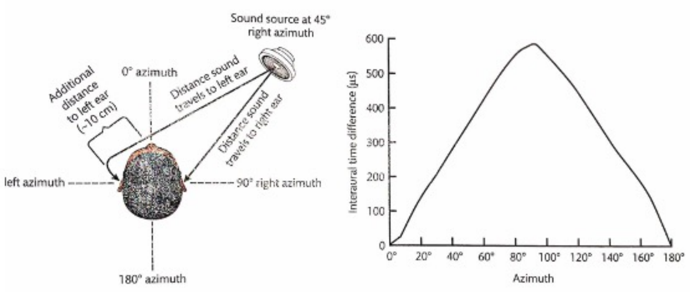
Lateral inhibition in
auditory cortex
- matched filter for 4 Hz ?

Perception of space

Interaural time differences

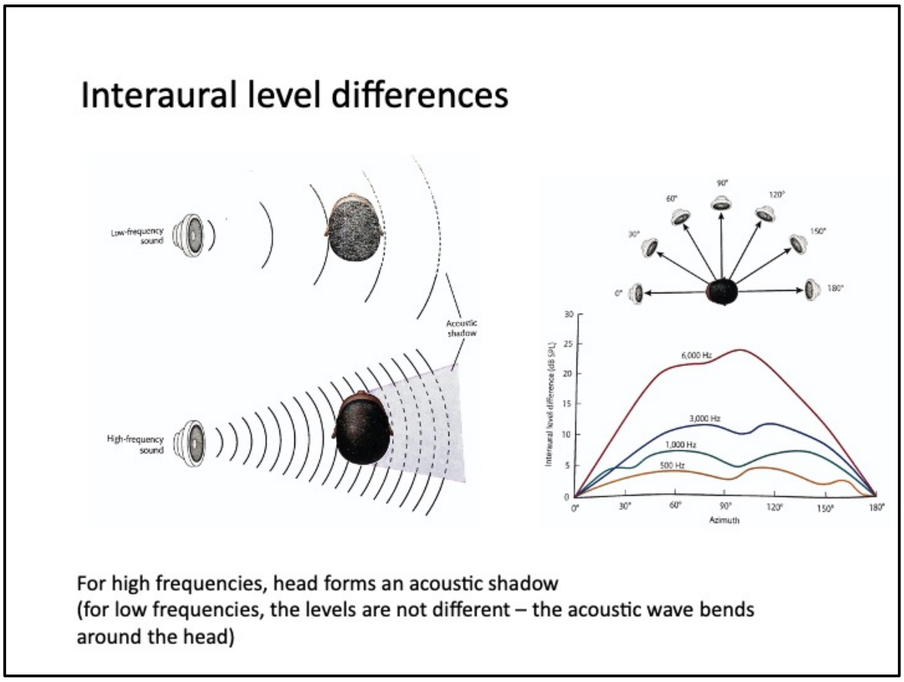
speed of sound 340 m/s
acoustic wavelength at 500 Hz $\lambda = 68$ cm

at 17 cm (distance between ears) the 500 Hz signal 90^0 out of phase



At low frequencies, the phase of the wave in the left and right ears differs, unless the source is exactly in front ior in back) of the head

When the listener is facing the sound (or the sound comes exactly from the back of the head), the signals arrive et the same time, i.e., the signals are in phase. At around 500 Hz, the acoustic wavelength is about 68 cm, that is the sound which comes perpendicullary to the direction the listener (with ears spaced at 17 cm) is facing with phase shift of 90^0. The phase differences for lower frequencies are smaller but still easily perceivable by binaural hearing. They become ambiguous at higher frequencies.



At higher frequencies, on the other hand, the sound is more severley attenuated bu the listener head and the sound level differences become very noticeable and can be used for estimating of the sond direction arival.