# Review of Cognitive Neuroscience: **Hearing**

Katrin Krumbholz

Contact: katrin@ihr.mrc.ac.uk, 922 3431

MRC Institute of Hearing Research (Science Road)

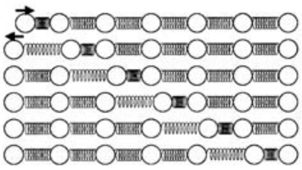
Topics:

- What is sound?
- Measures for sound pressure/intensity
- The basic building blocks of sounds: sinusoidal frequency components
- The structure of the ear (outer, middle and inner ear)
- The first step towards transforming acoustical information into neural activity: the cochlear travelling wave
- The organ of Corti, and how it converts sound information into nerve impulses
- Active feedback in the cochlear: dancing hair cells

## What is sound?

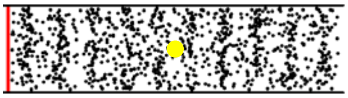
- Pressure waves

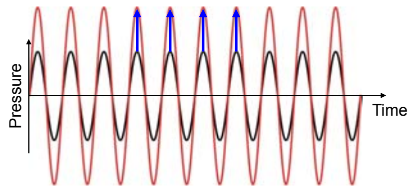
- layers of high or low pressure propagate in a similar fashion as a perturbation in a chain of marbles connected by springs



## What is sound?

- Pressure measured at a particular point in space varies over time as the layers of high and low pressure pass across the point of measurement



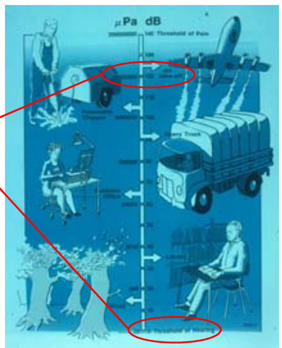



Amplitude:

➔ Pressure at the peaks of the pressure waveform

## Measures for sound pressure/intensity

- ***Intensity***: "Sound energy passing through a unit area (e.g., 1 $m^2$) per unit time (e.g., 1 s) "; unit = Watts per $m^2$ ($W/m^2$)
- The intensity of a sound ($I$) is proportional to the square of its pressure ($P$) ➔ $I = k \cdot P^2$




- ***Dynamic range*** (range of hearing):

➢ The faintest sounds humans can hear have a pressure of ~20 μPa

➢ Humans can process sounds with a pressure of ~20,000,000 μPa before feeling pain

➢ In terms of pressure, the hearing range corresponds to a factor of 1 million, and in intensity terms, 1 trillion, or 1 million millions

⇒ Thus, using the ordinary pressure or intensity units would mean having to deal with huge and unwieldy numbers

## Measures for sound pressure/intensity

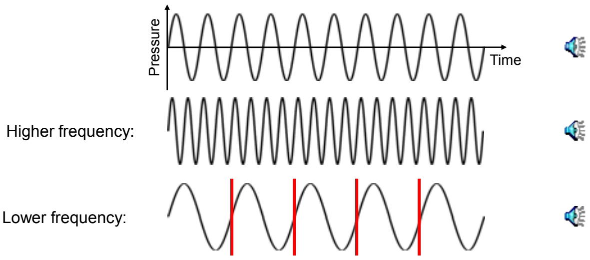
- Instead, pressure and intensity are specified in ***logarithmic units*** called ***decibels sound pressure level (dB SPL)***
- $L$ dB SPL $= 20 \cdot \log_{10}(P/20\,\mu\text{Pa})$

⇒ Faintest audible sound (20 μPa) = 0 dB SPL

⇒ Pain threshold (20,000,000 μPa) = 120 dB SPL

## Frequency decomposition

- Simplest possible sounds = pure tones (comp. to monochromatic light)
  - ➢ Pure tones have a sinusoidal pressure waveform

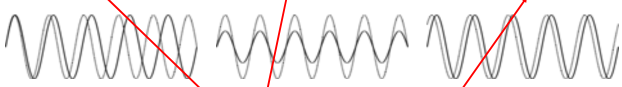



- ➢ Waveforms repeat themselves over time (temporally periodic); rate of repetition = frequency
- ➢ Frequency is measured in repetitions (cycles) per second, or ***Hertz*** (***Hz***)
- ➢ 1 repetition/s = 1 Hz; 1000 repetitions = 1000 Hz = 1 ***kHz***

## Frequency decomposition

- In order to characterise a pure tone, one needs to specify:

Frequency Amplitude (→Level) Starting phase



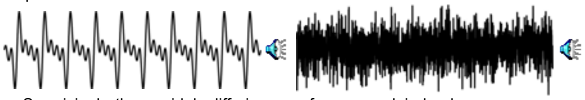
- ***Starting phase*** describes any ***temporal offset*** in the pure tone's pressure waveform
- Mathematically, the waveform of a pure tone is give by the equation

$$p(\text{time}) = A \cdot \sin(2 \cdot \pi \cdot F \cdot \text{time} + \theta),$$

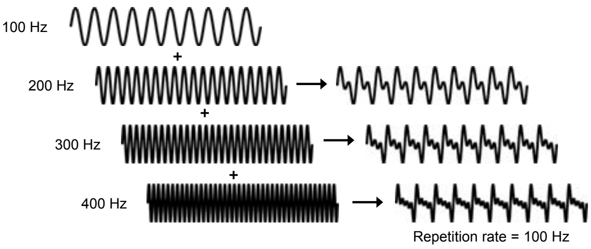
which is simply a sine function, multiplied by the amplitude, $A$; the sine function has two arguments, one that depends on the frequency, $F$, and the other, $\theta$, representing the starting phase

## Frequency decomposition

- Most sounds encountered in realistic environments look very different from pure tones:



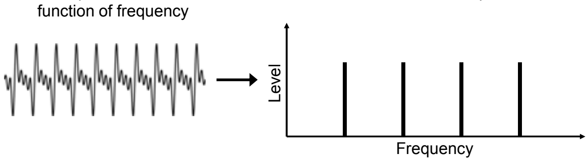
- Surprisingly, these widely differing waveforms – and, indeed, ***any waveform*** – can be produced by ***adding together*** pure tones of different amplitudes, frequencies and phases!



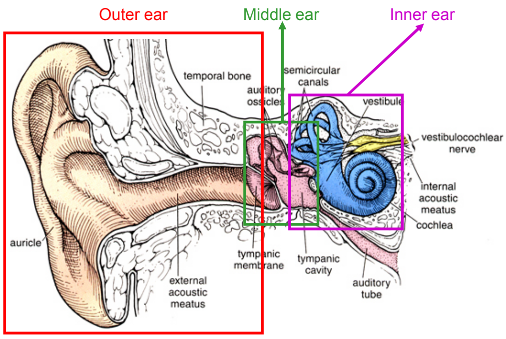

## Frequency decomposition

⇒ Pure tones/sinusoids are the basic building blocks of sounds; the whole concept of frequency is based on sinusoids! (E.g., a "low-frequency sound" is a sound that is composed of low-frequency sinusoids)

- Just as it is possible to *build up* a sound by adding together pure tones, it is also possible to *break down* a sound into a set of pure tones
- Sounds can be unambiguously described in terms of their sinusoidal frequency components
- In analogy to the spectrum of light as analysed through a prism, the description of a sound in terms of its frequency components is referred to as a ***spectrum***; the mathematical technique used to compute it is the ***Fourier analysis***
- The spectrum shows the levels of the sound's sinusoidal components as a function of frequency



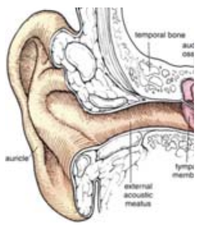

## Structure of the ear




## Structure of the ear

**Outer ear**

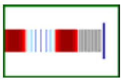
- The ***outer ear*** consists of the ***auricle*** and the ***ear canal***
- The outer ears "***filters***" the incoming sound (by letting through some frequencies better than others)
- The ***frequency transfer characteristics*** of the outer ear depend on the direction which the sound is coming from (the sound is filtered differently depending on its direction)

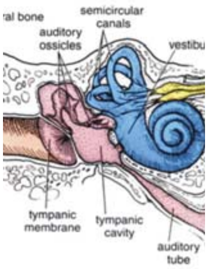


## Structure of the ear

**Middle ear**

- The outer ear is separated from the ***middle ear*** by the ***eardrum*** (***tympanic membrane***), which is pushed and pulled inwards and outwards by the sound waves



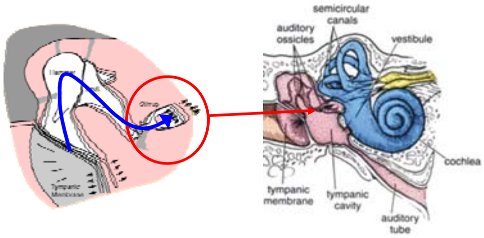



- The middle ear is connected to the back of the throat by the Eustachian (auditory) tube, which can be opened by yawning or swallowing ⇒ important for balancing the pressure inside the middle ear with the outside pressure

## Structure of the ear

**Middle ear - ossicles**

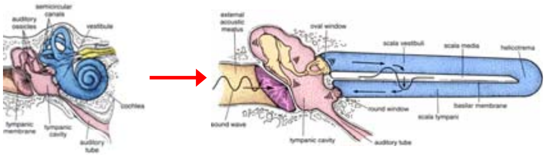
- The eardrum is connected to the ***inner ear*** by three tiny bones, the ***auditory ossicles***
- The ossicles amplify the pressure of the eardrum's vibrations like a lever and transmit them through the membrane of the ***oval window*** into the fluid-filled cavity within the bony shell of the ***cochlea*** in the inner ear




## Structure of the ear

**Inner ear**

- The ***cochlea***, in the ***inner ear***, is a thin, fluid-filled tube (3.5 cm long), which is curled up like a snail and has rigid, bony walls

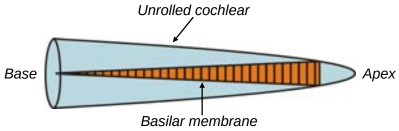


- The cochlea is divided into two compartments, the ***scala vestibuli*** and the ***scala tympani***, by a membrane called the ***basilar membrane***
- When the oval window moves inwards or outwards (due to the vibration of the ossicles) ⇒ a pressure difference between the scala vestibuli and the scala tympani is created ⇒ basilar membrane is pushed downwards or pulled upwards ⇒ similar to a water wave, this perturbation propagates from the ***base*** to the tip (***apex***) of the cochlea

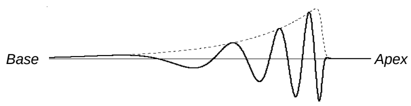
➔ This pattern of vibration is called the ***travelling wave*** [first discovered by Georg von Békésy (1960)]

## Travelling wave

- The mechanical properties of the basilar membrane change along its length: the membrane is narrow and stiff at its basal end and becomes progressively wider and more floppy towards the apex

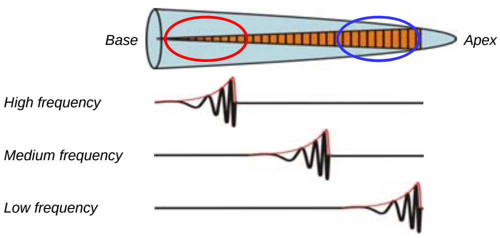



- Due to this mechanical gradient, the amplitude of the travelling wave changes as the wave progresses along the membrane, starting small and increasing until reaching its maximum at a point of resonance, beyond which the wave then dies away very abruptly




## Travelling wave

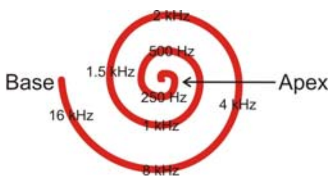
- Importantly, the point at which resonance is reached, depends on the frequency of the sound, because the narrow and stiff basal end of the basilar membrane responds better to high frequencies, whereas the wide and floppy apical end responds better to low frequencies (comp., e.g., strings of a harpsichord)




➔ Thus, different frequencies stimulate the basilar membrane at different points along its length, high frequencies at the base, medium frequencies in the middle and low frequencies at the apex

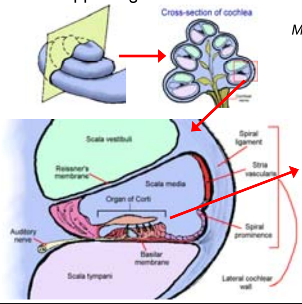
## Travelling wave

- This means that the cochlea acts as a ***frequency analyser***, decomposing sounds into their component frequencies (a bit like a Fourier analysis) and creating a ***topographic*** (spatial) map of frequency along the length of the basilar membrane. In analogy to the retinotopic map, the cochlear frequency map is referred to as the ***tonotopic*** map

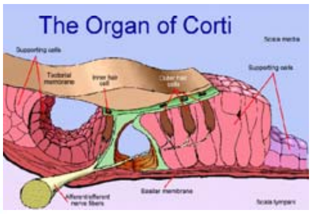


## Auditory receptor organ: Organ of Corti

- The mechanical response of the basilar membrane is converted into electrical brain activity by the ***organ of Corti***
- The organ of Corti is located on the basilar membrane; it contains several rows of ***hair cells*** (receptor cells) embedded in various supporting cells

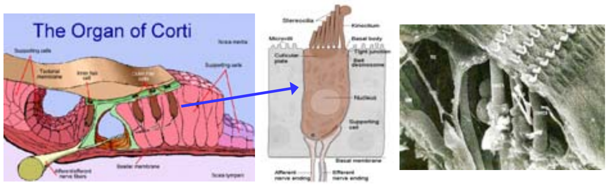



*Magnification of a part of the cochlear cross-section*



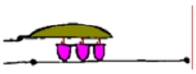

## Organ of Corti

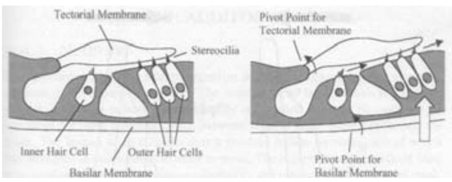
- The hair cells have little bundles of "hair" (***stereocilia***) sticking out of their tops
- Humans have 1 row of ***inner hairs cells*** (***IHCs***) and 3 rows of ***outer hair cells*** (***OHCs***)
- At the bottom, the hair cells are contacted by the endings of the ***auditory nerve fibers***
- 95% of the auditory nerve fibers are ***afferent*** fibers, which carry information from the cochlea to the brain; the remaining fibers are ***efferent***, i.e., they carry information from the brain to the cochlea
- Most of the afferent fibers (95%) innervate the IHCs, most of the efferent fibers innervate the OHCs ⇒ IHCs convey sound information to the brain




## Mechano-electrical transduction

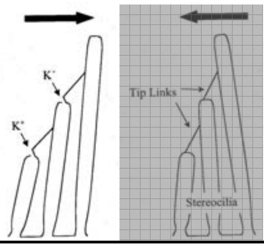
- On top of the organ of Corti lies a gelatinous structure, referred to as the ***tectorial membrane***
- Movement of the basilar membrane upwards (towards the scala vestibuli) produces a shearing force between the basilar membrane and the tectorial membrane, causing the stereocilia on the hairs cells to deflect to the right





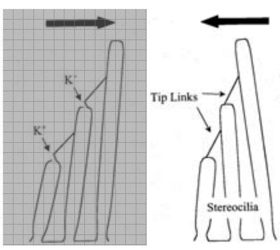

## Mechano-electrical transduction

- This stretches little protein filaments, the so-called ***tip-links***, with which neighbouring rows of stereocilia are connected
- The stretching of the tip links causes them to pull open tiny channels in the membranes of the stereocilia, causing positively charged potassium ions ($K^+$) from the surrounding fluid to flow into the hair cell and produce a ***receptor potential***, i.e., an ***increase in the electrical potential*** of the cell above its negative ***resting potential*** (***depolarisation***)
- This depolarisation causes the attached auditory nerve fibers to send ***electrical impulses*** (***spikes/action potentials***) to the brain and thus eventually elicit a perception of the sound



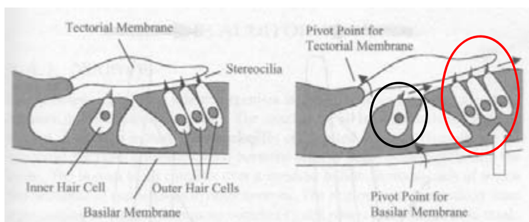

## Mechano-electrical transduction

- In contrast, the opposite movement of the basilar membrane, i.e. downwards, towards the scala tympani, deflects the stereocilia to the left and thus slackens the tip-links, and so, the ion channels stay closed



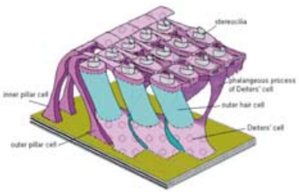

## Active feedback

- The OHCs are depolarised in the same way as the IHCs



## Active feedback

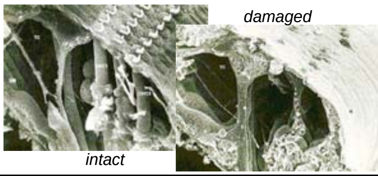
- When an OHC depolarises, the *entire cell contracts and shortens*, thereby literally pulling the basilar membrane towards the cell, because the OHCs are affixed to the basilar membrane through the supporting cells
- This phenomenon, which is known as ***electromotility***, causes the OHCs to *actively* feed *mechanical energy* back into the system!
- Electromotility it is powered by a specialized protein (prestin), lodged in the OHCs' membrane
- Movie of an OHCs, which has been isolated and whose membrane potential is being varied in the rhythm of a popular rock tune using the so-called patch clamp technique

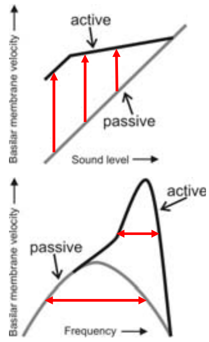



## Active feedback

- While the exact mechanism of the active feedback is not yet understood, the feedback is known to have two important consequences on hearing:
- First: to amplify the movement of the basilar membrane at low sound levels
- Second: to increase the sharpness of the frequency tuning of the basilar membrane

➢ Thus, when the OHCs are damaged or missing, there is not only a loss in sensitivity (increase in hearing threshold), but also in frequency resolution, leading to particular difficulties in hearing in noisy environments




Questions so far?

Topics:

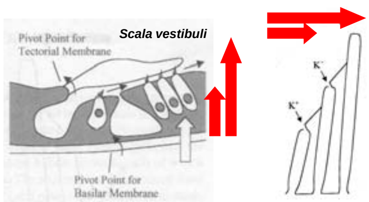
- Different aspects of information mediated by the auditory nerve (level, frequency and timing)
- How might the auditory system use this information to process complex sounds, such as speech?
- Where does it go from here? The central auditory system
- Spatial hearing

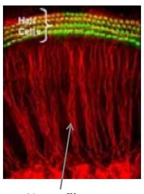
## Information in auditory nerve – (1) sound level

- Movement of the basilar membrane towards the scala vestibuli opens $K^+$ ion channels in the hair bundles of the IHCs via the tip-links ⇒ ***receptor potential*** ⇒ ***action potentials*** in attached auditory nerve fibers

➢ A ***higher sound level*** ⇒ produces a ***larger movement*** of the basilar membrane ⇒ causes tip-links to ***open more ion channels*** ⇒ creating a ***larger receptor potential*** in the hair cells and thus ***more spikes*** in the auditory nerve fibers

- Each IHC is contacted by ~20-30 auditory nerve fibers
- Most (90%) of these fibers are very sensitive ⇒ so sensitive that they produce a substantial background level of spikes (~60 spikes per s) even in the absence of any sound [***spontaneous rate*** (***SR***); ***high-SR fibers***]
- The remaining fibers are less sensitive (spontaneous rate ≤ ~10 spikes per s; ***low-SR fibers***)



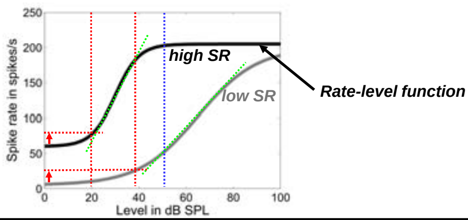




***Nerve fibers***

## Information in auditory nerve – (1) sound level

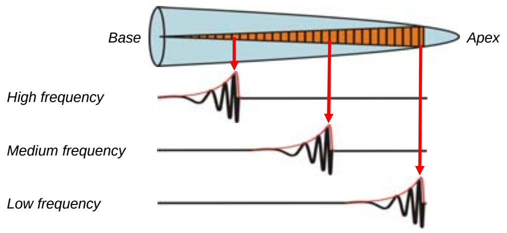
- ***Rate-level functions***: Spike rate increases with increasing sound level up to a maximum rate at which the response ***saturates***
  - ➢ High-SR fibers have lower ***firing thresholds*** than low-SR fibers (firing threshold = sound level that produces significant increase of firing rate above the spontaneous rate)
  - ➢ High-SR fibers also have steeper rate-level functions (unit increase in level produces larger increase in firing rate in the high- than in the low-SR fibers)
  - ➢ High-SR fibers saturate at higher sound levels (above about 50 dB SPL; increases in level above the ***saturation level*** do not produce any further increases in firing rate)
- Perception: Humans are very good at discriminating sound levels *even at very high levels*

⇒ Rate-level functions suggest that sound level discrimination above ~50 dB SPL must rely on the signal from the low-SR fibers!



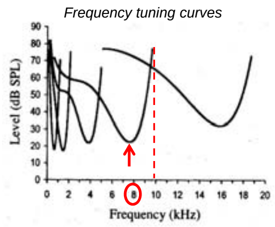
## Information in auditory nerve – (2) frequency

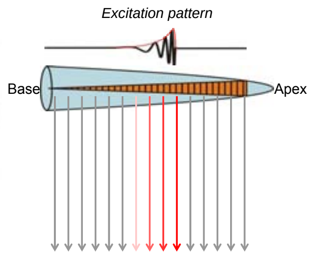
- The basilar membrane exhibits mechanical tuning for sound frequency ⇒ different places along the length of the membrane respond maximally to different frequencies




## Information in auditory nerve – (2) frequency

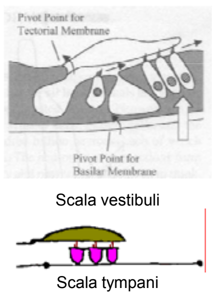
- As a consequence of this mechanical frequency tuning of the basilar membrane, the auditory nerve fibers also exhibit tuning for frequency ⇒ each nerve fiber responds only to a limited range of sound frequencies
- ➢ Thus, the profile of activation strength across the auditory nerve (***excitation pattern***) reflects the place where the basilar membrane is stimulated, which, in turn, reflects the frequency/spectral composition of the sound

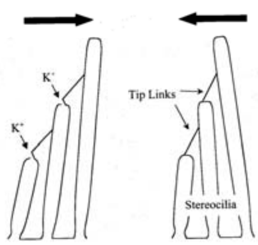



## Information in auditory nerve – (3) timing

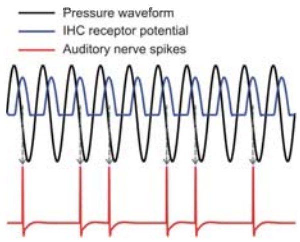
- Movement of the basilar membrane creates a shearing force, which deflects the hair bundles on the hair cells
- Only one phase of the motion of the basilar membrane (the motion towards the scala vestibuli) is effective in eliciting a receptor potential in the hair cells



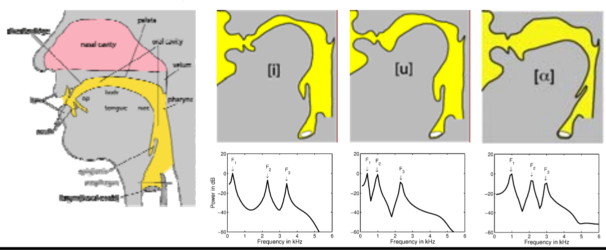

## Information in auditory nerve – (3) timing

- Receptor potential reflects only one half-cycle of basilar membrane motion
- Therefore, the receptor potential resembles a ***halfwave-rectified*** version of the original pressure waveform of the sound
- This means that the spikes in the auditory nerve are *time-locked* to a particular *phase* in the waveform of the sound
- ➢ The temporal pattern of the auditory nerve spike trains reflects the temporal structure of the sound ⇒ ***phase locking***



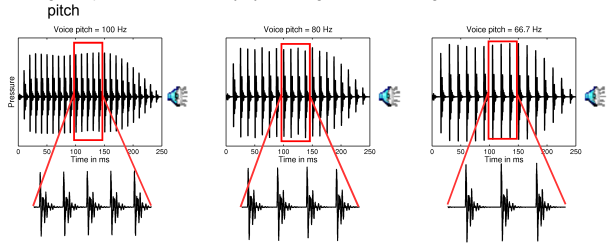

## Processing of complex sounds

- When we speak, the opening and closing of the ***vocal folds*** releases puffs of air into the oral cavity, thereby producing sound (***glottal pulses***)
- The quality of the resulting sound depends on the shape of the oral cavity (e.g., different vowels)
- Different vowels contain different frequencies
- The frequency spectra of vowels are characterised by three prominent frequencies, the ***formants*** ⇒ different vowels have different formants
- ➢ This is why the spectral information conveyed by the auditory nerve may be expected to play an essential role in the identification of speech sounds



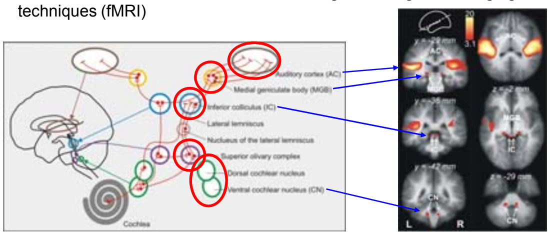

## Processing of complex sounds

- In contrast, the perception of pitch/melody in speech and music may be based on the temporal (phase-locking) information conveyed by the auditory nerve
- The pitch of the voice conveys information both about speaker identity (male/female) and about meaning (e.g., "Peter fed the dog?" versus "Peter fed the dog.")
- The pitch of the voice is determined by the rate at which the vocal folds open and close ⇒ the *faster* the rate of the ***glottal pulses***, the *higher* the pitch
- Due to phase locking, the auditory nerve tends to fire in synchrony with the glottal pulses ⇒ the auditory system might use this timing information to derive pitch



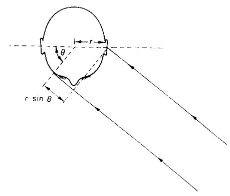
## Central auditory system

- The auditory system is exceptional in the sense that many of the auditory structures are located in the brainstem rather than the cortex as, for instance, in the visual system
- Most important auditory structures: (i) ***cochlear nucleus*** (first relay station), (ii) ***superior olivary complex*** (consists of several nuclei; this is the stage where information from the two ears is combined for the first time), (iii) ***inferior colliculus*** (relays practically all of the ascending auditory projections), (iv) ***medial geniculate body*** (auditory part of the thalamus), (v) ***auditory cortex*** (***supra-temporal plane***)
- Most of these structures can now be investigated using brain imaging techniques (fMRI)



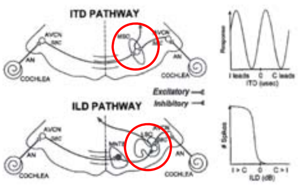
## Spatial hearing

- It is generally assumed that one of the main functions of the sub-cortical part of the auditory system is to analyse the acoustical cues for sound location
- In humans, sound localisation mainly relies on the analysis of differences in sound level and sound arrival time at the two ears [referred to as ***interaural level*** and ***interaural time differences*** (***ILDs***, ***ITDs***)]
- In a sound originating from a lateralised source, ILDs are produced by the head casting a shadow on the farther ear ⇒ ILDs more prominent in high-frequency sounds, because low-frequency sounds can "bend around" the head (***diffraction***)
- ITDs are produced by the path length differences between the sound source and the two ears ⇒ ITDs are of the order of a few tens to a few hundreds of *microseconds* (a thousands of a thousands of a second) ⇒ thus, neuronal processing of ITDs requires a phenomenal temporal accuracy! ⇒ ITDs are more important at low frequencies, because temporal processing (phase locking) breaks down at high frequencies



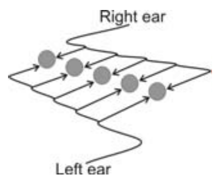
## Spatial hearing

- The initial processing of ILDs and ITDs starts in the ***superior olivary complex*** (***SOC***) in the brainstem
- Generally assumed that ILDs and ITDs are processed by different types of neurons, located in different nuclei of the SOC
- The lateral superior olive (LSO) contains neurons that receive excitatory (activating) input from one ear and inhibitory (suppressing) input from the other ear ⇒ these neurons effectively compute the difference between the signals from the two ears, which makes them sensitive to ILDs
- Medial superior olive (MSO) contains neurons that receive excitatory input from both ears ⇒ sensitive to ITDs



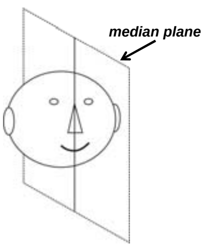

## Spatial hearing

- MSO neurons might be converting ITDs to a topographic (spatial) representation similar to the tonotopic representation of sound frequency
- Idea that ITDs might be converted to spatial code, which was first proposed by ***Jeffress*** in 1949, is still basis of most current models of spatial hearing
  - Jeffress proposed that ITDs are processed by a set of coincidence neurons receiving excitatory input from both ears and being activated only by simultaneous/coincident input
  - Input to the coincidence neurons provided by axons whose length varies systematically across the set
  - The longer the axon the longer it takes the spikes to travel to the neurons
  - Difference in the time taken for the signal from each ear to reach the coincidence neurons varies systematically across the set
  - A given ITD is represented by that neuron for which the difference in the *axonal conduction delay* from the two ears compensates the ITD



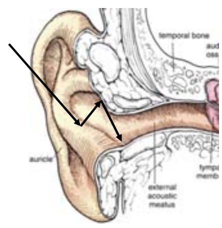

## Spatial hearing

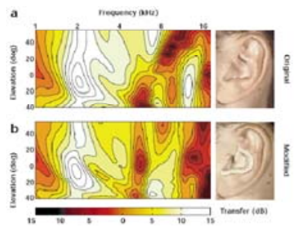
- In the ***median plane*** ⇒ sounds elicit neither ITDs nor ILDs
- Nevertheless, humans can localise the elevation of sound sources with reasonable accuracy




## Spatial hearing

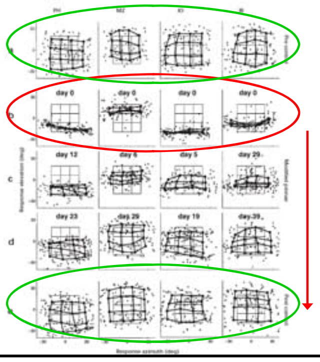
- In the median plane, sound location is conveyed by ***spectral cues***
- Sound wave impinge on head and outer ear ⇒ reflections from and deflections around surfaces enhance some frequencies and attenuate others
  - Journey towards ear drum imparts spectral "profile" to the incoming sound ⇒ the shape of the profile depends on the direction of the sound
  - Location-dependent spectral profiles are dependent on shape of outer ear ⇒ highly individual
  - Being able to use spectral profiles for sound localisation means having learned the frequency transfer characteristics of one's own ears (Hofman et al., 1998. Nat. Neurosci. 1, 417-421)





## Spatial hearing

- Ability to localise sound elevation was dramatically degraded immediately after the modification
- Over weeks of wearing the molds, subjects progressively reacquired their ability to judge sound elevation



→ Further reading (books):

1. C. J. Pack, The Sense of Hearing
2. B. C. J. Moore, An Introduction to the Psychology of Hearing
3. W. A. Yost, A. N. Popper and R. R. Fay, Human Psychophysics
4. J. O. Pickles, An Introduction to the Physiology of Hearing
5. P. Dallos, A. N. Popper and R. R. Fay, The Cochlea

→ Websites:

i. http://www.brainconnection.com/topics/?main=anat/auditory-phys
ii. http://www.cf.ac.uk/biosi/staff/jacob/teaching/sensory/ear.html
iii. http://www.biols.susx.ac.uk/home/Chris_Darwin/Perception/Lecture_Notes/Hearing_Index.html