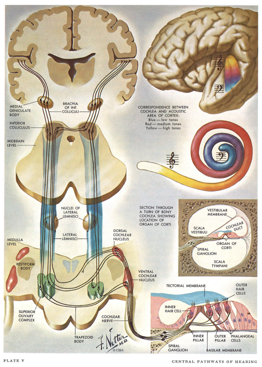
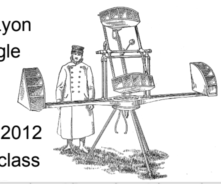
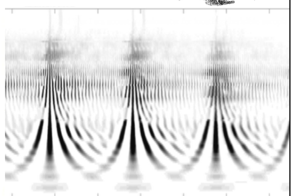
Does the brain process sound in the time domain?



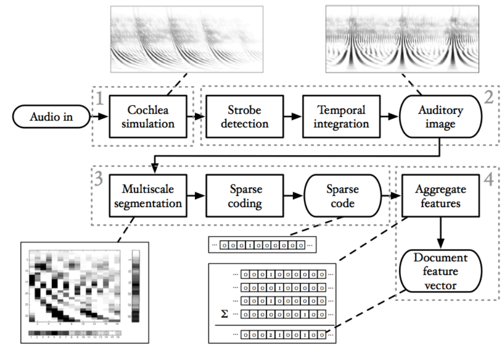
Dick Lyon
Google



May 9, 2012
cortex class



Our typical "Machine Hearing" system makes good use of fine time structure



Respect the Auditory Nerve

Ed E. David, Jr. 1958
"Artificial Auditory Recognition in Telephony"

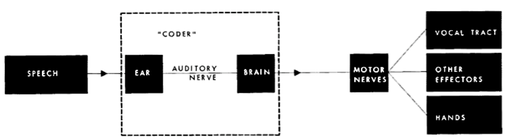
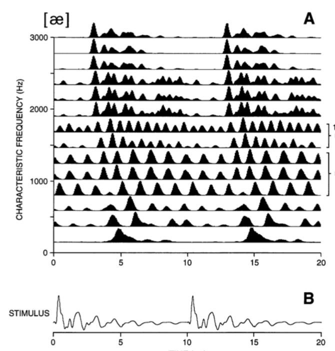


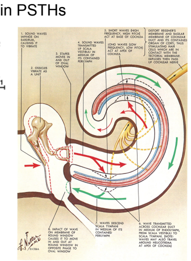
Figure 5 **Representation of efficient speech transmission.**

- What's on the auditory nerve?
- What happens between auditory nerve and cortex?

Auditory nerve carries fine time structure

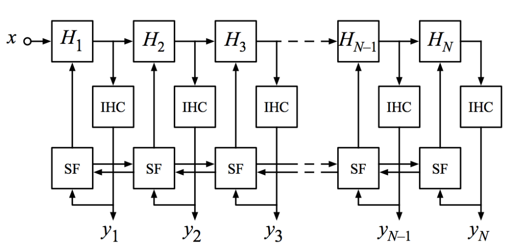


Delgutte 1997
shows pitch-synchronized response at all CFs in PSTHs



Cochlear model: CAR-FAC

(cascade of asymmetric resonators with fast-acting compression)



outputs along cochlear place, or frequency, axis (high to low)
with several kHz bandwidth, plenty of fine time structure

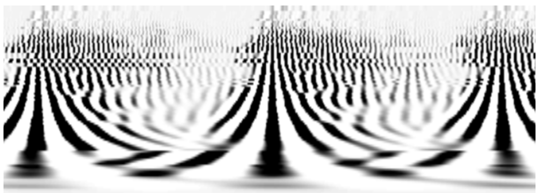
Tonotopic projection to cortex, but…

What is the other dimension in auditory cortex?

What corresponds to image retinotopy in vision?

Auditory images…

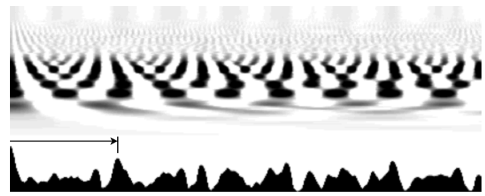
Correlograms are movies; this one is based on Roy Patterson's "Stabilized Auditory Image" (SAI)



(speech + complex tone)

Timing explains pitch of aperiodic sounds

- "Strike note of a chime" demo from ASA *Auditory Demonstrations* CD.
- Correlogram (SAI) shows near alignment at interval corresponding to the musical pitch:



Temporal coincidence detection theories: Jeffress 1948 binaural, Licklider 1951 pitch

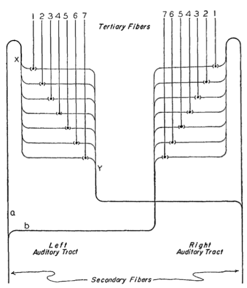


FIGURE 1. HYPOTHETICAL MID-BRAIN MECHANISM FOR THE LOCALIZATION OF LOW FREQUENCY TONES

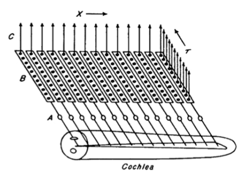
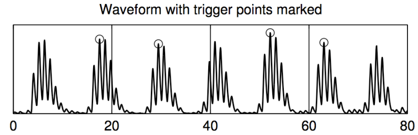


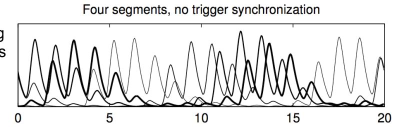
Fig. 2. – *Schematic diagram of overall analyzer.* At the bottom is the uncoiled cochlea. Its lengthwise dimension and the corresponding dimension in the neural tissue above it is designated the x-dimension. The cochlea performs a crude frequency analysis of the stimulus time function, distributing different frequency bands to different x-positions. In the process of exciting the neurons of the auditory nerve, the outputs of the cochlear filters are rectified and smoothed. The resulting signals are carried by the groups of neurons A to the autocorrelators B, whose delay- or τ-dimension is orthogonal to x. The outputs of the autocorrelators are fed to higher centers over the matrix of channels C, a cross-section through which is called the (x, τ)-plane. (Output arrows arise from all the dots; some are omitted in the diagram to avoid confusion.) The time-varying distribution of activity in the (x, τ)-plane provides a progressive analysis of the acoustic stimulus, first in frequency and then in periodicity.

A trigger method for Patterson's *triggered temporal integration* approach to making stabilized auditory images:

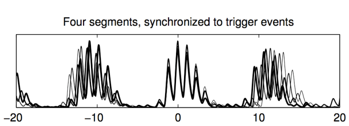
One trigger per segment, at max; channel 60:

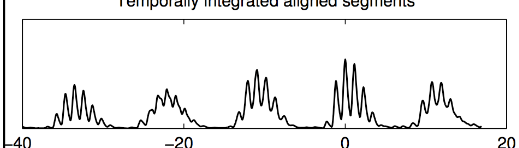


Without trigger alignment, integrating (averaging) segments doesn't make sense:

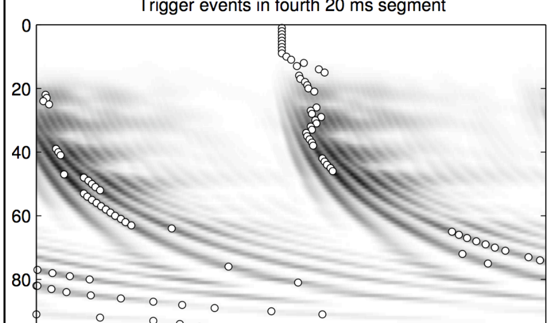


Align to triggers, and temorally integrate

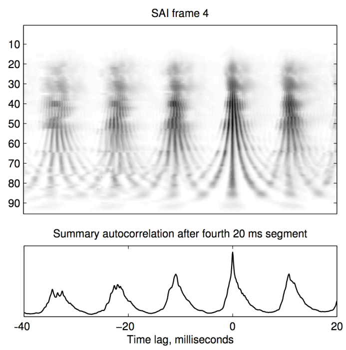




Trigger points in 20 ms cochleagram segment



SAI and Summary SAI



Licklider's autocorrelator neural circuit

$$g(t,\tau) = (f(t)f(t-\tau)) * w(t)$$

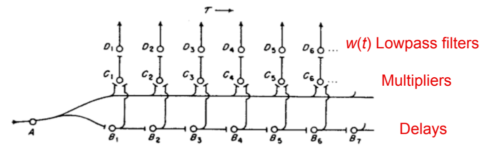
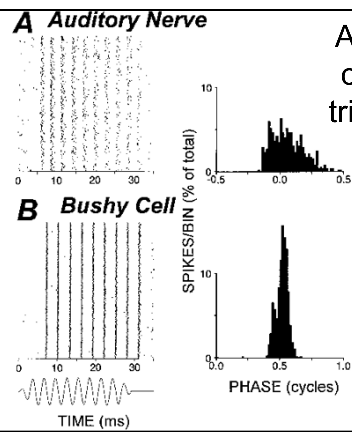


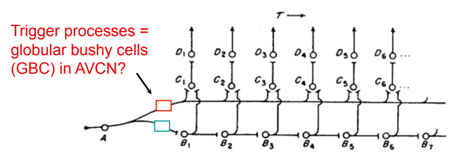
Fig. 1. – *Basic schema of neuronal autocorrelator.* A is the input neuron, B_1, B_2, B_3, ... is a delay chain. The original signal and the delayed signal are multiplied when A and B_k feed C_k, and a running integral of the product is obtained at the synapse between C_k and D_k, where excitation accumulates whenever C_k discharges and dissipates itself at a rate proportional to the amount accumulated. Since these operations correspond to the definition of running autocorrelation, the excitatory states at D_1, D_2, D_3, ... provide a display of the running autocorrelation function of the input time function, the temporal course of the discharges of A.



AVCN bushy cells extract trigger points?

"Primary-like" bushy cells sync well to stimulus period

Cross-corr versions



$$g_1(t,\tau) = (\hat{f}(t)f(t-\tau)) * w(t)$$

$$g_2(t,\tau) = (f(t)\hat{f}(t-\tau)) * w(t)$$

Demo break

real-time
AIM-C movies
of music and stuff

then binaural…

Lord Rayleigh's 1907 "duplex theory" of binaural lateralization: interaural intensity and phase differences

Phase is ambiguous above about 650 Hz.

Thus, although there might be right and left sensations from sources obliquely situated, these sensations would fail when most needed, that is when the source is really in the line of the ears. In this case a perception of phase-differences would seem to do more harm than good. At a pitch a little higher, ambiguities of a misleading and dangerous kind would necessarily enter. For example, the same sensations might arise from a sound a little on the left and from another fully on the right.

On the whole it appears that the sensation of lateralness due to phase-difference disappears in the region of pitch where there would be danger of its becoming a misleading guide. ... It is fortunate that when difference of phase fails, difference of intensity comes to our aid.

But sounds other than sinusoids are easier…

Jens Blauert *Spatial Hearing* 1997

2.4 Evaluating Nonidentical Ear Input Signals 149

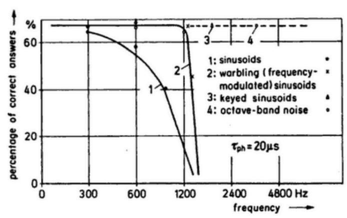
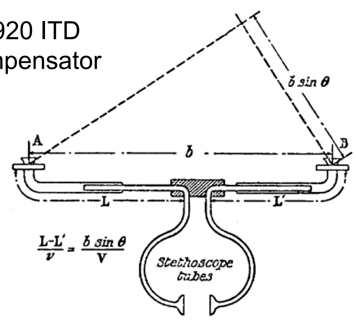


Figure 2.73
Ability to notice an interaural phase delay, shown as a function of frequency for various types of signals (after Scherer 1959; 1 subject).

Mallock 1908 observation of ITD cue

A sound which is caused by the detached waves, such as those which accompany a bullet, can scarcely be said to have a pitch, but the wave-length is certainly small compared with the distance between the ears, and is indeed comparable with the dimensions of the bullet itself. It would seem, therefore, that the ears can determine the direction of a sound, not only by difference of phase, but by the actual difference in the times at which a single pulse reaches them.

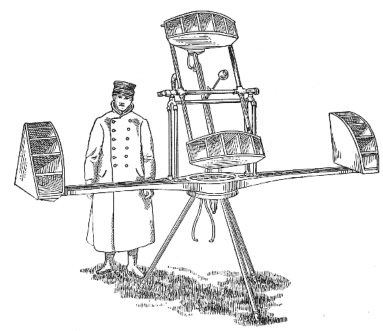
1920 ITD compensator



Binaural method with rectilinear compensator.

World War I era directional listener using a linear compensator to compe
ave arrivals (Drysdale, 1920). The dashed line on the left is a wavefront

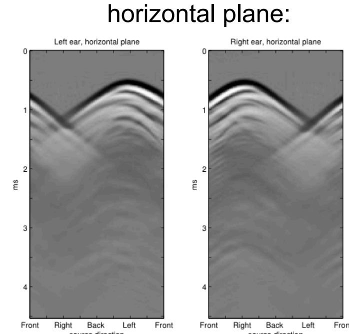
Rotating ITD eliminator measures angles directly



World War I era acoustic goniometer for locating "invisible aeroplanes" (Ferry, 1921)

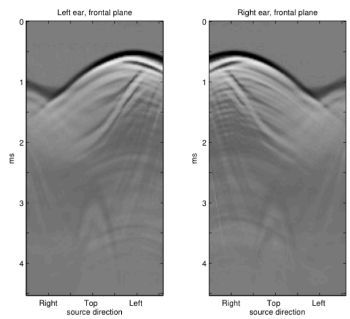
Daniel Tollin at U. Colo. Denver

HRTFs and HRIRs "encode" direction; horizontal plane:



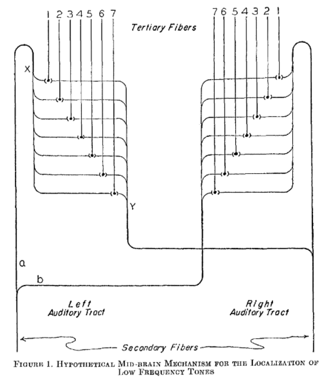
Head-related impulse responses of a dummy head, for sounds from various directions

HRIRs for “frontal plane” (over top of head)



Head-related impulse responses of a dummy head, for sounds from various directions

Neural circuit in MSO



Jeffress’s cross-correlator circuit as a model of coincidence detection in the two MSOs.

Barn Owl’s ITD system in N.L.

M. Konishi / Comparative Biochemistry and Physiology, Part A 126 (2000) 459–469 463

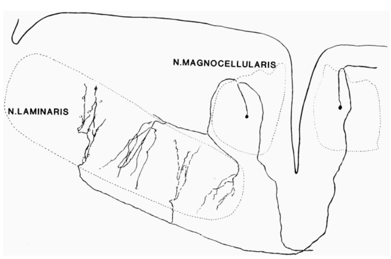


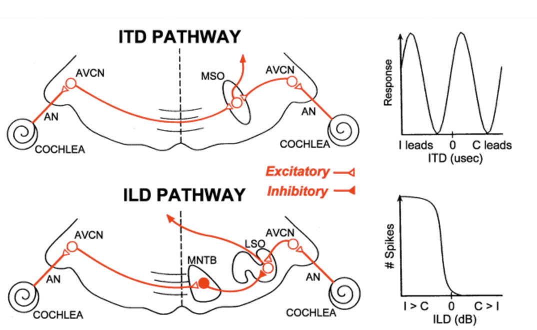
Fig. 6. Axonal delay lines and coincidence detection. In the owl's auditory system, axons from nucleus magnocellularis serve as delay lines and nucleus laminaris cells work as coincidence detectors. This figure shows a reconstruction of an axon and its branches from each side. The lines are axons and the circles are laminaris cell bodies. Nucleus laminaris is tonotopically organized and the detection of ITDs occurs separately in different frequency bands (from Konishi et al., 1988).

Lloyd Jeffress’s model…
Bowlker 1908

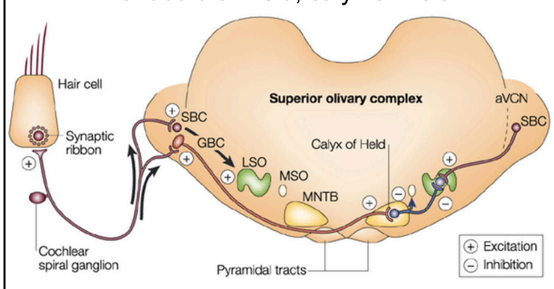
In order to explain the existence of a movable image of the sound within this zone, we may suppose that the transmission of the sound impulse through some specialized part of the auditory apparatus or brain takes a definite time from each ear, and that the point where the impulses meet is the focus that gives rise to the sensation of a sound-image.

This was regarded as an “unwelcome hypothesis”.

Binaural circuits
fig. from Daniel J. Tollin 2003



Specialized fast cells and synapses for triggering and correlation: endbulb of Held, calyx of Held



Henrique von Gersdorff & J. Gerard G. Borst 2002

Nature Reviews | Neuroscience

About this book | Preview this book | **Physiology of the ear** By Anthony F. Jahn, Joseph Santos-Sacchi

Page 620

620 PHYSIOLOGY OF THE EAR

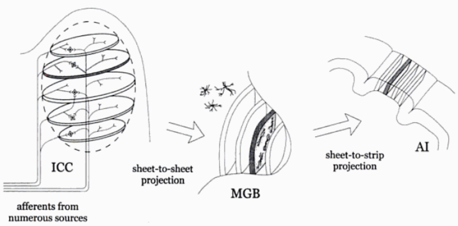
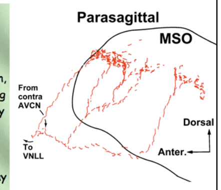


Figure 26-3. Schematic depiction of the connectivity between the auditory midbrain and the auditory cortex. Note that the inferior colliculus receives input from many sources and that these inputs are neither completely overlapping in their spatial distribution nor completely segregated. Note that in the laminated, ventral division of the MGB, the principal cells have tufted dendrites and are oriented in the isofrequency plane; in the nonlaminated, medial division of the MGB, the cells are often larger and may have stellate dendritic arrangements (after Winer[84]). There is a sheet-to-sheet connectivity between the tonotopic arrays of the ICC and the ventral division of the MGB, but a convergent, sheet-to-strip connectivity between the thalamus and the cortex.

Tom C. Yin's lab: confirming Jeffress-style delay lines in MSO

Spherical bushy cells

Spherical bushy cells are associated with the *primary-like* response to pure tones at CF. Their axons exit the AVCN in ventral stria (or trapezoid body) and travel in the medial division, projecting bilaterally to the MSO. The contralaterally projecting axon has a distinct **delay-line** (right) form, like that proposed by the Jeffress model, but the ipsilaterally projecting axon does not. All of the contralaterally projecting axons that we injected had a more caudal than rostral projection while maintaining the tonotopic organization in the MSO. This caudally projecting delay line is in agreement with the topographic map of ITDs found in the MSO of the cat.

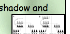


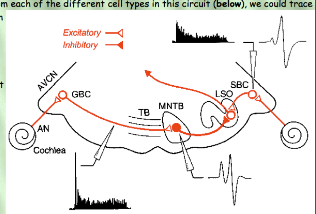
Tom C. Yin's lab: ITD in LSO!

Studies of the LSO

- **Responses to ITDs of amplitude modulated tones**

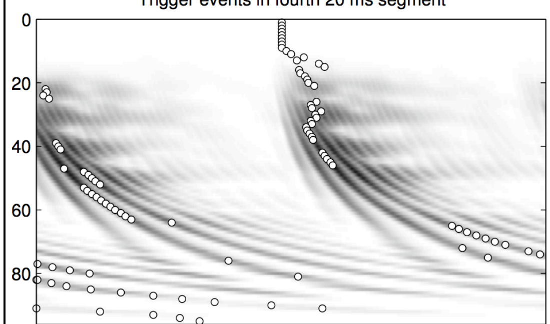
Classically the LSO is thought to encode ILDs: it receives excitation from spherical bushy cells from the ipsilateral AVCN and inhibition from the contralateral side relayed by MNTB cells of the same side. This interaction is referred to as IE (contra inhibitory and ipsi excitatory). ILDs are effective cues at high frequencies where the head acts as a shadow and phase locking to tones disappears. However, human psychophysical studies have shown that ITDs can also be detected for high frequency amplitude modulated signals, and **physiological studies** have shown that high frequency auditory nerve fibers can phase lock to AM signals. Therefore, we studied responses of LSO cells to ITDs of **high frequency AM tones.** As expected, the responses were consistent with the ILD sensitivity of the cells, with a minimal response when the AM signals were in phase (**above**). **Characteristic delays of LSO cells** were studied by varying the modulation frequency. As expected the characteristic phases of LSO cells were clustered around 0.5, reflecting the IE binaural interaction. By recording AM responses from each of the different cell types in this circuit (**below**), we could trace the gradual decline of phase synchronization at each synaptic relay. We found that the **restricted modulation frequency range** over which LSO cells show ITD-sensitivity does not result from loss of envelope information along the afferent pathway but is due to convergence or postsynaptic effects at the level of the LSO. Measurements of time delay through the two sides of the circuit to the LSO showed that on average the contralateral inhibitory input arrived only 200 microsecs later than the ipsilateral excitatory input, even though the contralateral pathway is longer and has an extra synapse. This is in accord with characteristic delay measurements and estimates derived from single cells.





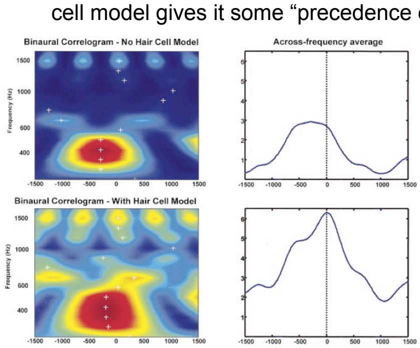
Highest peaks are not best arrival-time triggers

Trigger events in fourth 20 ms segment

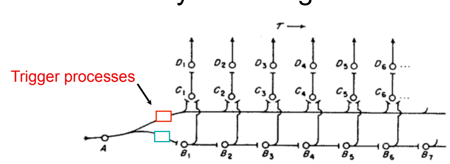


Binaural Auditory Image

Hartung & Trahiotus 2001: an appropriate hair cell model gives it some "precedence effect"



Monaural cross-corr versions can easily be changed to binaural



$$g_1(t,\tau) = (\hat{f}(t) f(t-\tau)) * w(t)$$

$$g_2(t,\tau) = (f(t) \hat{f}(t-\tau)) * w(t)$$

Binaural test: speech and ping; Lyon 1983

inputs:

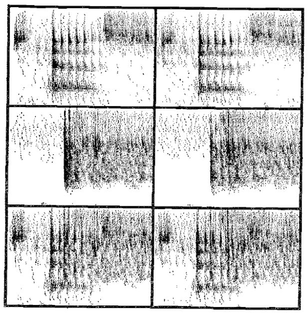


Figure 2. Cochleagrams of test signals. Top: left and right channels of speech sound. Middle: left and right channels of interfering *ping* sound, with reverberation. Bottom: left and right composite sounds, the inputs to the binaural separation test.

Binaural localization/separation outputs

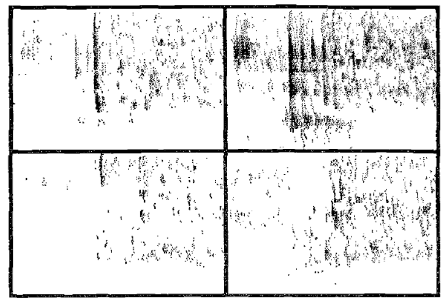


Figure 3. Separation results. Top: left and right separated sound streams. Bottom: left and right echos, or reverberation.