HST.725 Music Perception and Cognition, Spring 2009
Harvard-MIT Division of Health Sciences and Technology
Course Director: Dr. Peter Cariani

Neural mechanisms of musical pitch

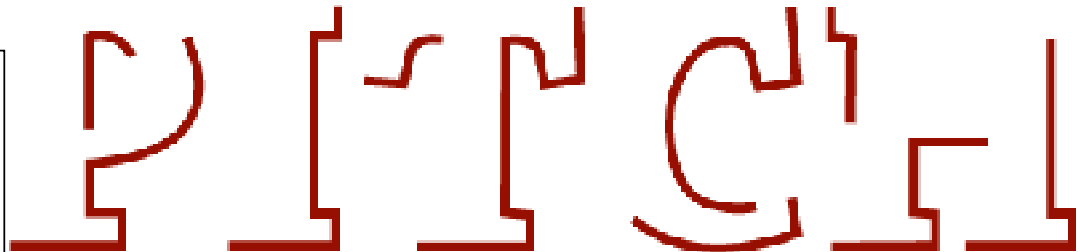


Figure by MIT OpenCourseWare.

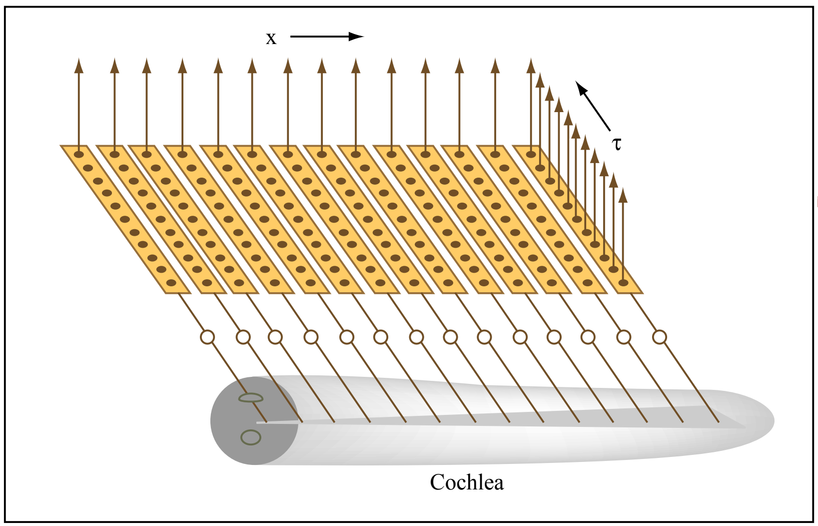


Figure by MIT OpenCourseWare.

www.cariani.com

Friday, March 13, 2009

Big questions - why music?

- **What does music do for us?**
- **Why is music effective at this?**
- **How is music structured to make it effective?**
- **What are the neural codes & computations?**
- **Why is music the way it is?** (e.g. why scales?)
- **How/why did music arise?**
- **How can I become a rock star?**

evolutionary adaptation?

hijack internal rewards?

self-control of states?

Roadmap

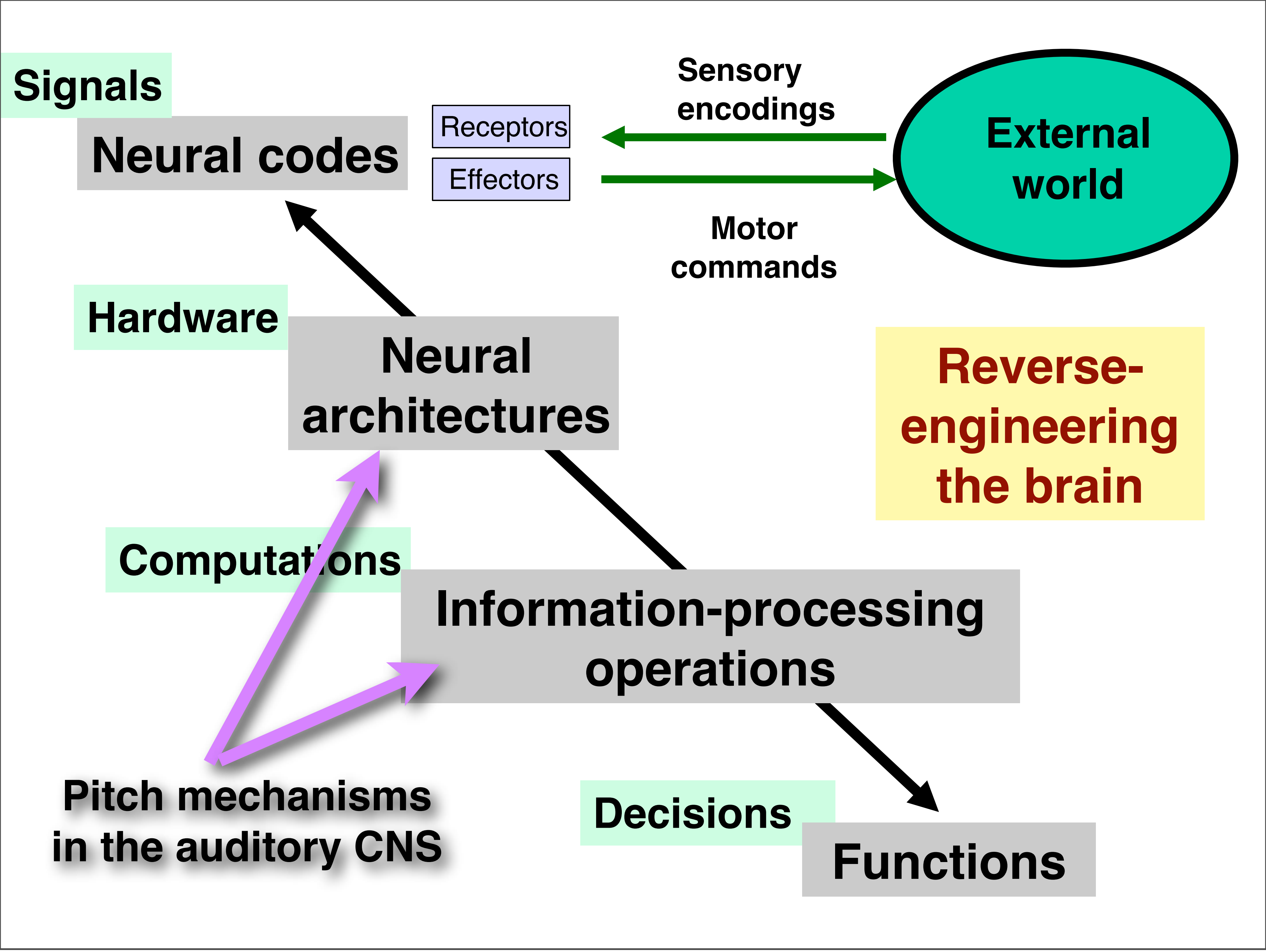
- overview
- functions of music
- sound, ear
- loudness & pitch
- timbre

  basic qualities of notes
- consonance, scales & tuning — interactions between notes
- melody & harmony — patterns of pitches
- time, rhythm, and motion — patterns of events
- grouping, expectation, meaning — interpretations
- music & language

Roadmap

- music therapy
- neurology of music
- developmental & comparative psychology
- origins: evolutionary psychology
- how/why music fulfils its diverse functions

my own belief is that music
speaks the language of the brain,
a temporal pattern code
and that this is why music can affect us
in so many different ways


Signals
Neural codes
Receptors
Effectors
Sensory
encodings
Motor
commands
External
world
Hardware
Neural
architectures
Reverse-
engineering
the brain
Computations
Information-processing
operations
Pitch mechanisms
in the auditory CNS
Decisions
Functions

Basic analysis strategies

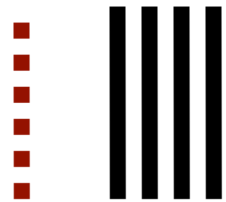
- Frequency-domain
  - Place-codes form central spectra
  - In some models, interspike intervals form central spectra (Goldstein)
  - Patterns of partials are analyzed to infer F0
  - Architectures: feature-detectors and connectionist networks
  - Output: pitch detectors
- Time-domain theories
  - Temporal patterns of spikes form autocorrelation-like representations
  - Dominant interval patterns correspond to F0-pitch
  - Architectures1:
    - Time-to-place conversion (Jeffress, Licklider. time-delay neural networks)
    - Output: Pitch detectors
  - Architectures2:
    - Time-to-time conversion (neural timing nets)
    - Output: Temporal patterns of spikes; pattern similarity detection

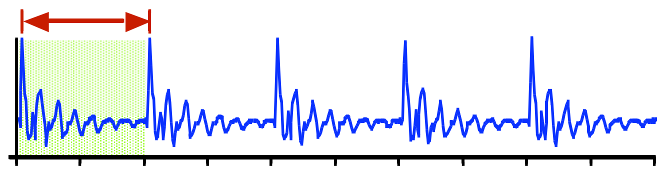
- Evidence in the auditory pathway
- Neural timing models
  - Pitch matching, similarity, and F0-based separation

Search for the missing fundamental: theories & models of musical pitch

- Distortion theories (nonlinear processes produce F0 in the cochlea)
- Spectral pattern theories
  - Pattern-recognition/pattern-completion
  - Fletcher: frequency separation
  - The need for harmonic templates (Goldstein)

    Terhardt's Virtual pitch: adding up the subharmonics
  - Musical pitch equivalence classes
  - Pitch classes and neural nets: Cohen & Grossberg
  - Learning pitch classes with connectionist nets: Bharucha





- Temporal theories
  - Residues: Beatings of unresolved harmonics (Schouten, 1940's)
  - Problems with residues and envelopes
  - Temporal autocorrelation models (Licklider, 1951)
  - Interspike interval models (Moore, 1980)
  - Correlogram demonstration (Slaney & Lyon, Apple demo video)

Basic aspects of pitch to be explained

- **Pure tone pitches (50-20,000 Hz)**
- **Complex tone pitches (periodic sounds F0s 30-1000 Hz)**
- **Pitch equivalence classes (pure & complex tones w. diff spectra)**
- **Precision and robustness of pitch discrimination**
- **Pitch salience** (why some pitches are strong or weak)
- **Pitch similarities** (octave relations)
- **Musical interval recognition/transposition/pitch relativity**
- **Role of common periodicity in auditory grouping**
  - **How multiple notes are simultaneously represented.**
- **Pitch memory (for relative & absolute pitch)**

Harmonic series

A harmonic series conists of integer multiples of a fundamental frequency, e.g. if the fundamental is 100 Hz, then the harmonic series is: 100, 200, 300, 400, 500, 600 Hz, ... etc.

The 100 Hz fundamental is the *first harmonic*, 200 Hz is the *second* harmonic. The fundamental is often denoted by F0.

The fundamental frequency is therefore the greatest common divisor of all the frequencies of the partials.

Harmonics above the fundamental constitute the ***overtone series***.

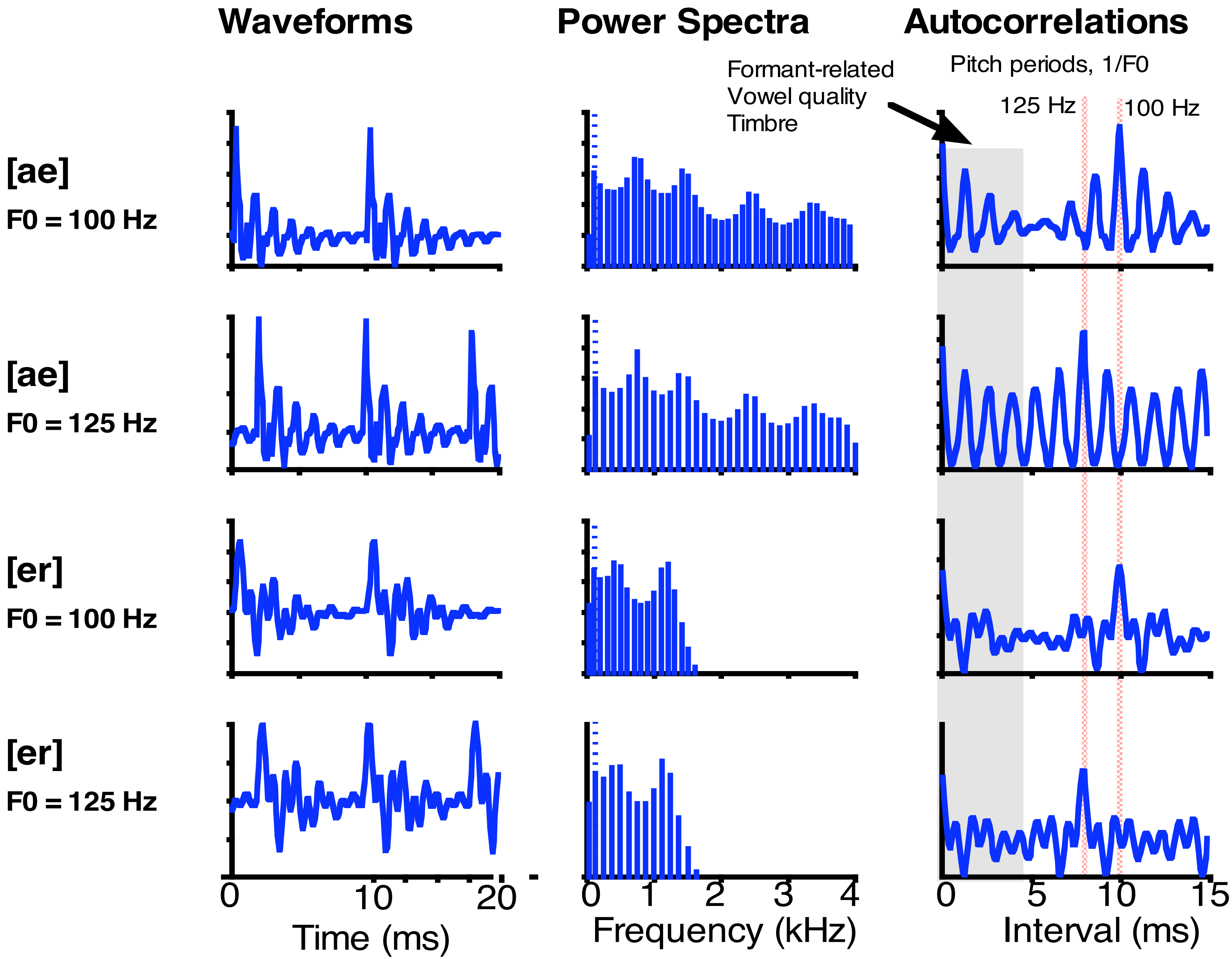
Subharmonics are integer divisions of the fundamental:

e.g. for F0= 100 Hz, subharmonics are at 50, 33, 25, 20, 16.6 Hz etc.

Subharmonics are also called ***undertones***.

The fundamental period is 1/F0, e.g. for F0=100 Hz, it is 1/100 sec or 10

Periodic sounds: time and frequency domains



Pitch : basic properties to be explained

- **Highly precise percepts**
  - **Musical half step: 6% change F0**
  - **Minimum JND's: 0.2%** at 1 kHz (20 usec time difference, comparable to ITD jnd)
- **Highly robust percepts**
  - **Robust quality** Salience is maintained at high stimulus intensities
  - **Level invariant** (pitch shifts < few % over 40 dB range)
  - **Phase invariant** (largely independent of phase spectrum, f < 2 kHz)
- **Strong perceptual equivalence classes**
  - **Octave similarities** are universally shared
  - **Musical tonality** (octaves, intervals, melodies) 30 Hz - 4 kHz
- **Perceptual organization ("scene analysis")**
  - **Fusion:** Common F0 is a powerful factor for grouping of frequency components
- **Two mechanisms? Temporal (interval-based) & place (rate-based)**
  - **Temporal:** predominates for periodicities < 4 kH (level-independent, tonal)
  - **Place:** predominates for frequencies > 4 kHz(level-dependent, atonal)

Periodic sounds produce distinct pitches

Many different sounds produce the same pitches

Strong
- Pure tones
- Harmonic complexes
- Iterated noise

Weaker
- High harmonics
- Narrowband noise

Very weak
- AM noise
- Repeated noise

Strong pitches

Weaker low pitches

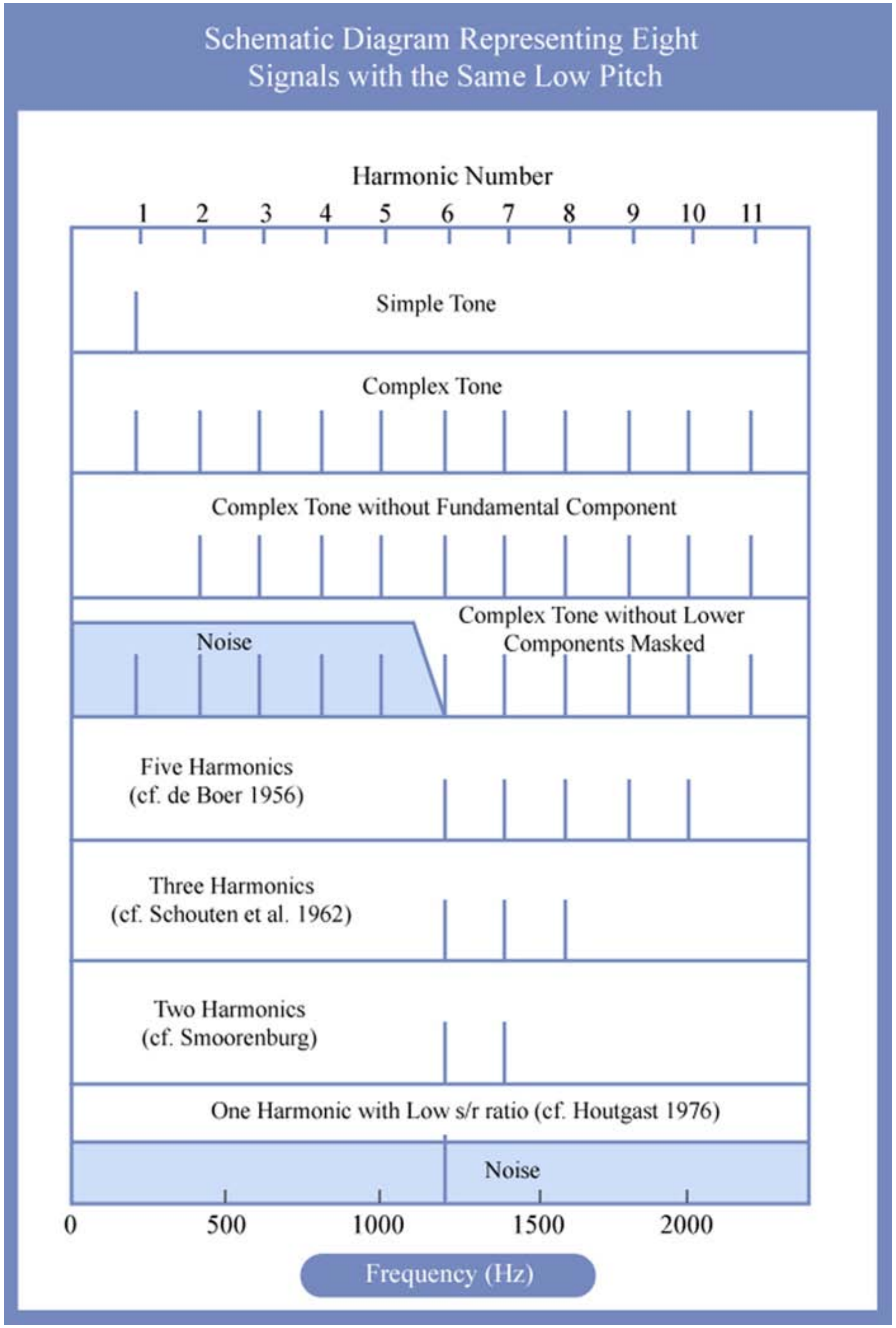
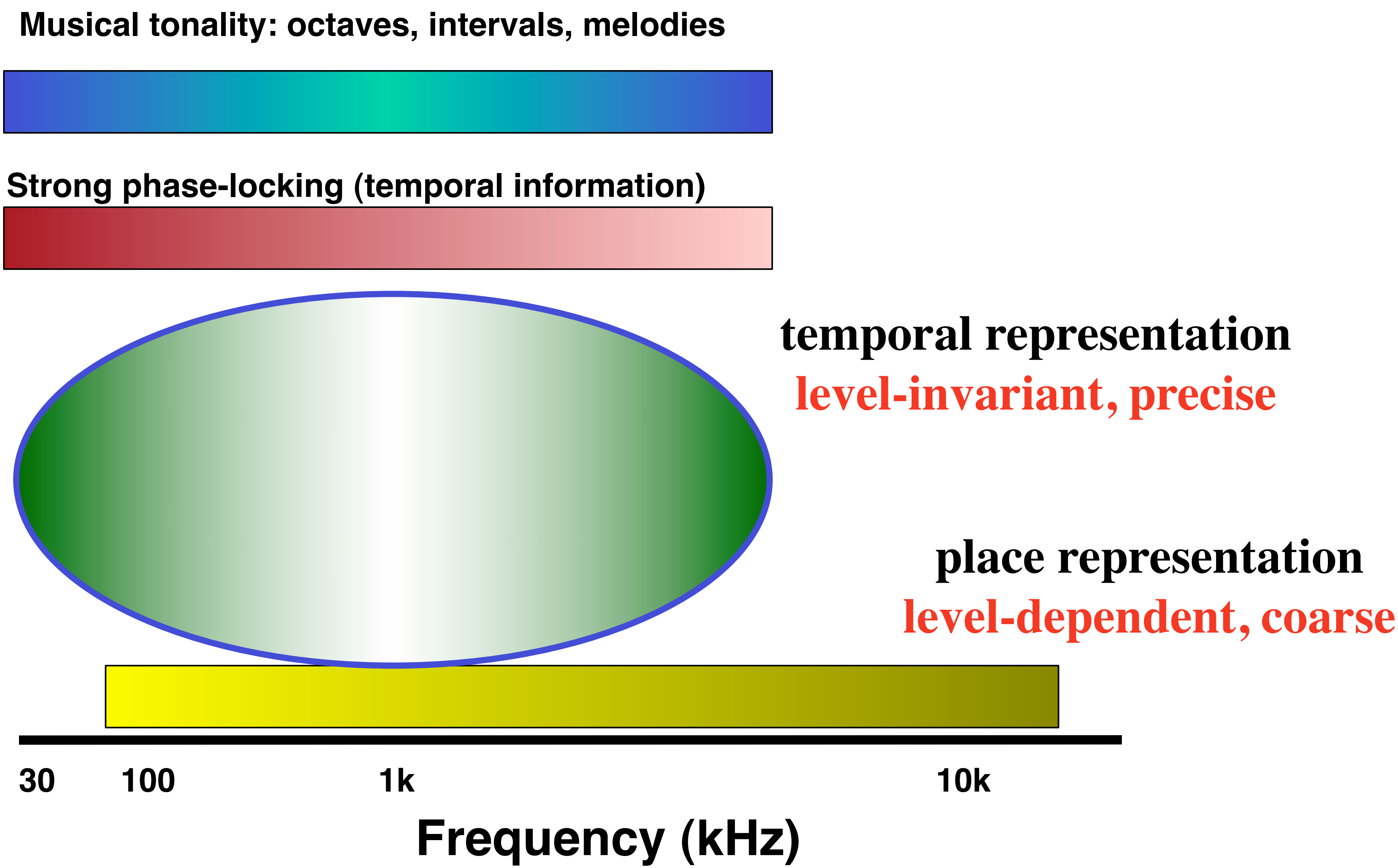
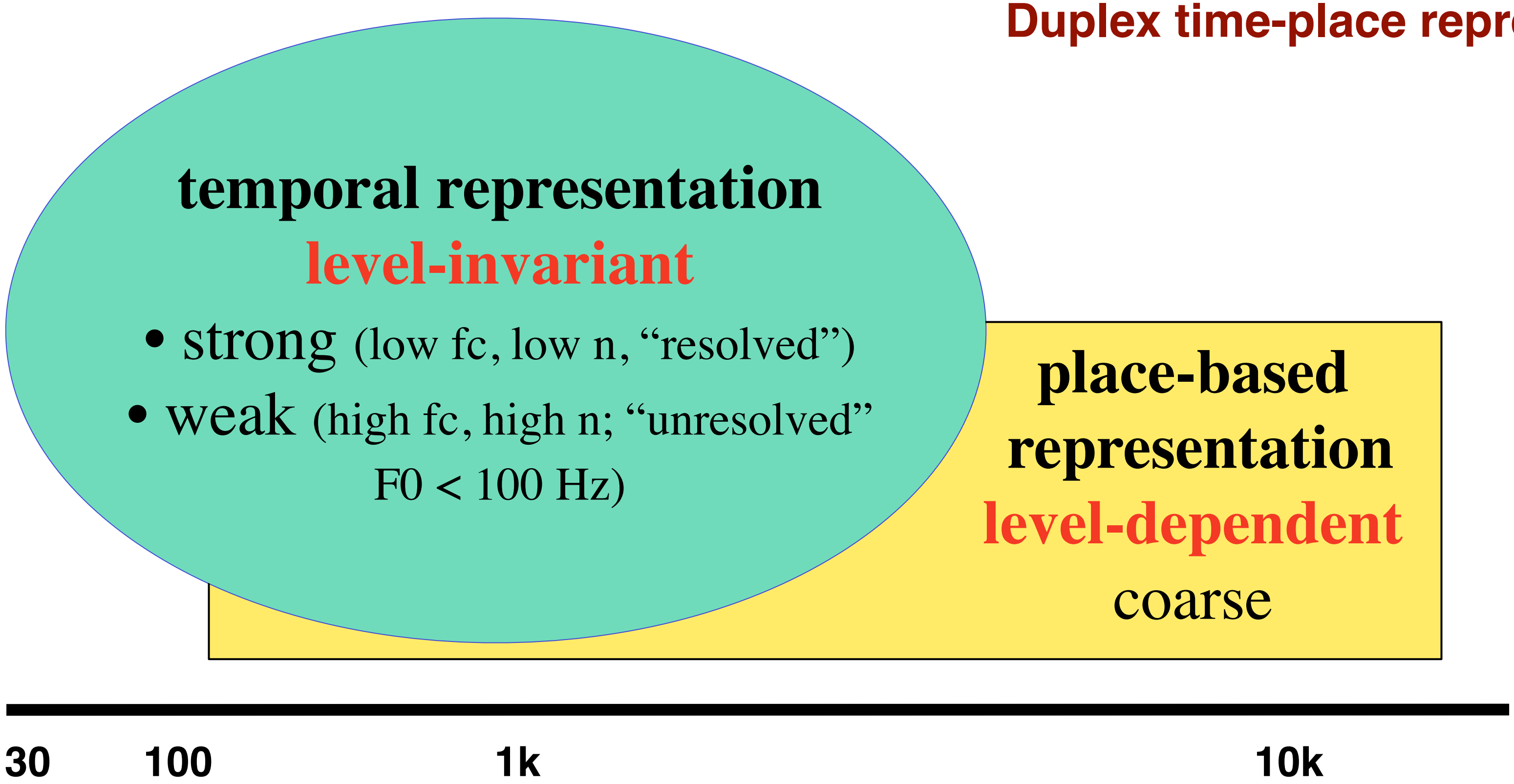


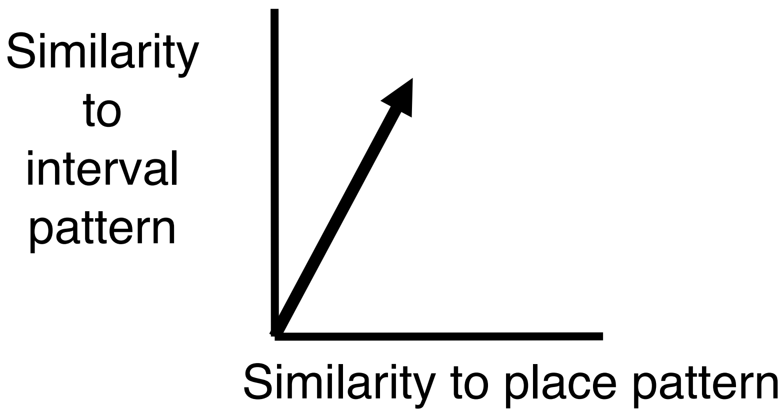
Figure by MIT OpenCourseWare.

Duplex time-place representations
"Pitch is not simply frequency"



Duplex time-place representations

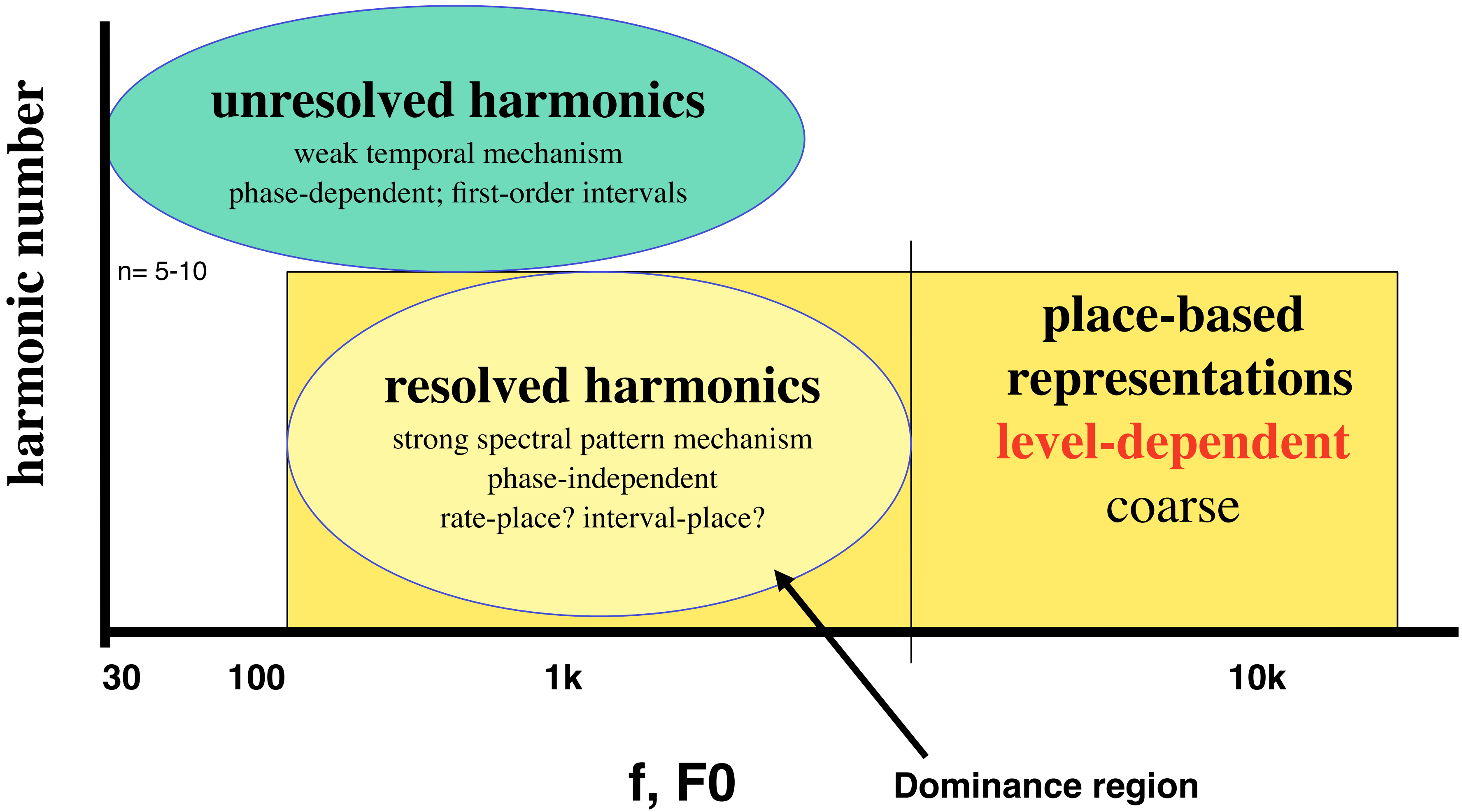




cf. Terhardt's
spectral and virtual pitch

A "two-mechanism" perspective

(popular with some psychophysicists, compatible with spectral pattern models of F0 pitch)



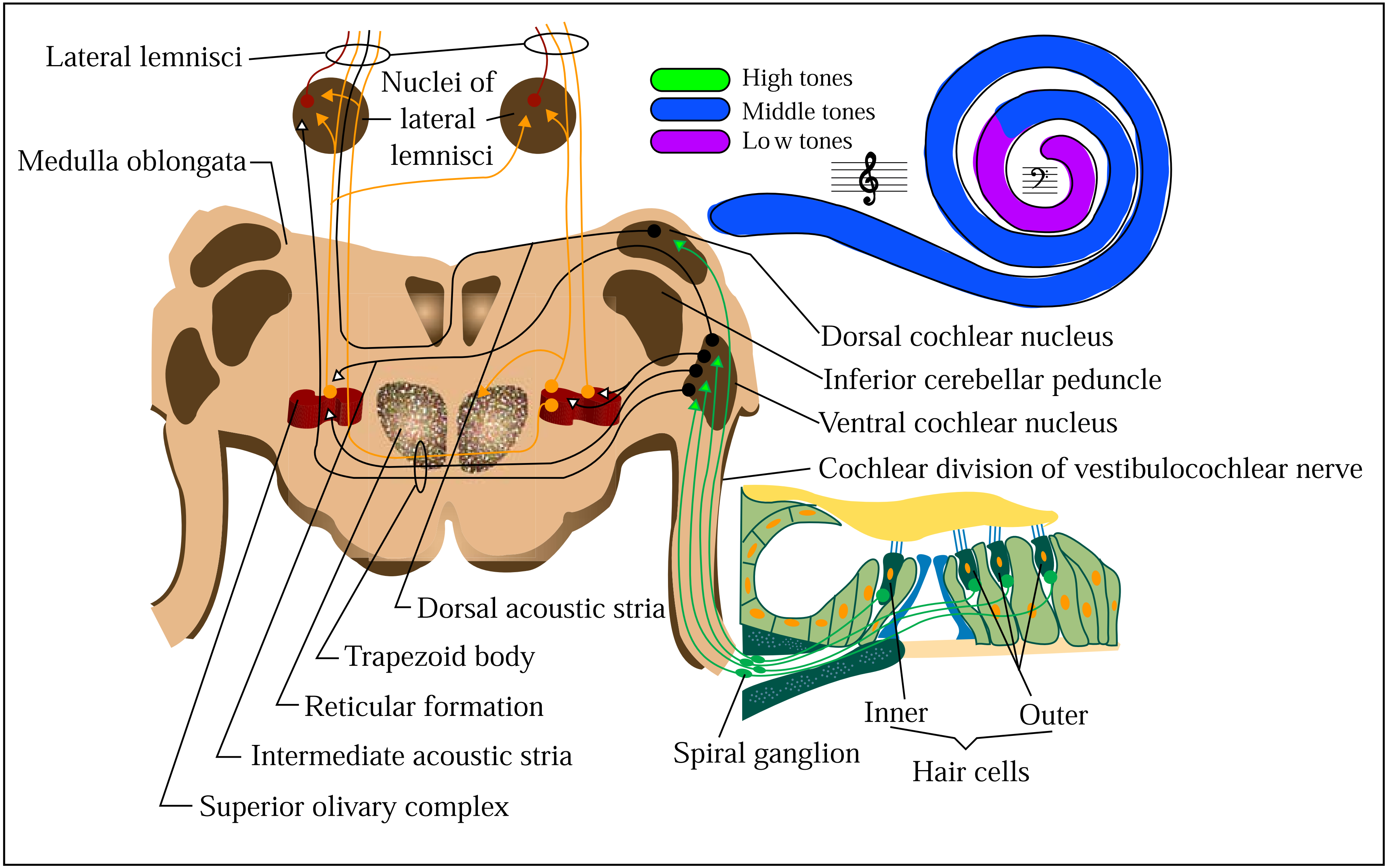
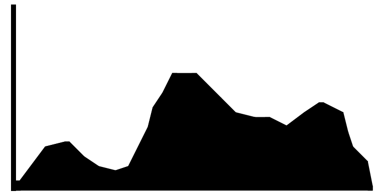


Figure by MIT OpenCourseWare.

Some possible auditory representations

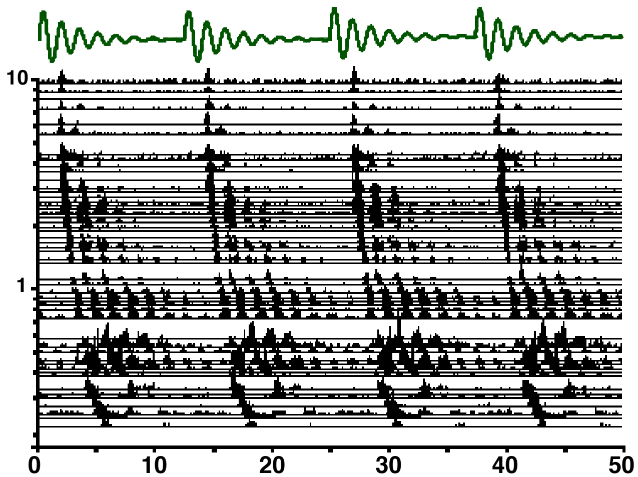
Local

Central spectrum



CF

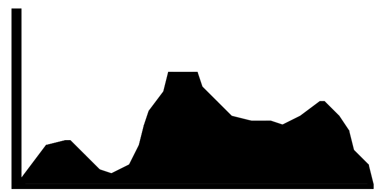
Masking phenomena
Loudness



Synchrony-place

Phase-place

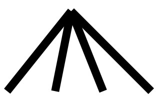
Interval-place

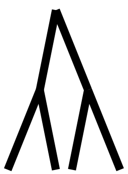


Pure tone pitch JNDs: Goldstein

Stages of integration

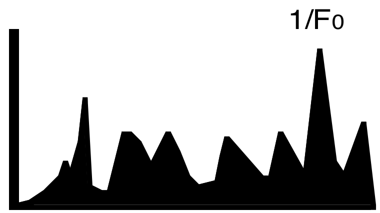
All-at-once





Population interval

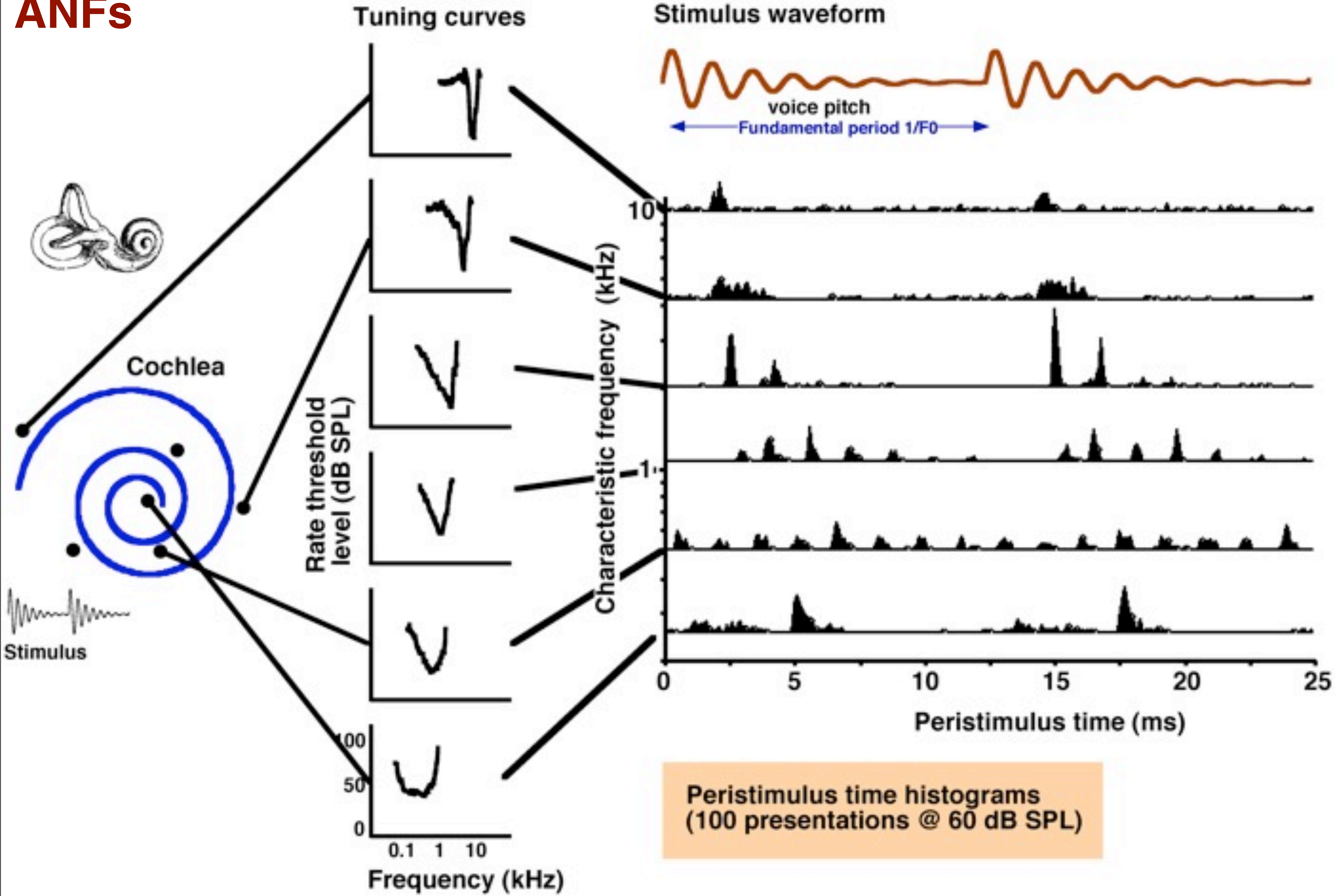
Population-interval



Complex tone pitch

Global

ANFs



Peristimulus time histograms
(100 presentations @ 60 dB SPL)

General theories of pitch

1. Distortion theories

– reintroduce F0 as a cochlear distortion component (Helmholtz)

–sound delivery equipment can reintroduce F0 through distortion

–however, masking F0 region does not mask the low pitch (Licklider)

–low pitch thresholds and growth of salience with level not consistent with distortion processes (Plomp, Small)

–binaurally-created pitches exist

2. Spectral pattern theories

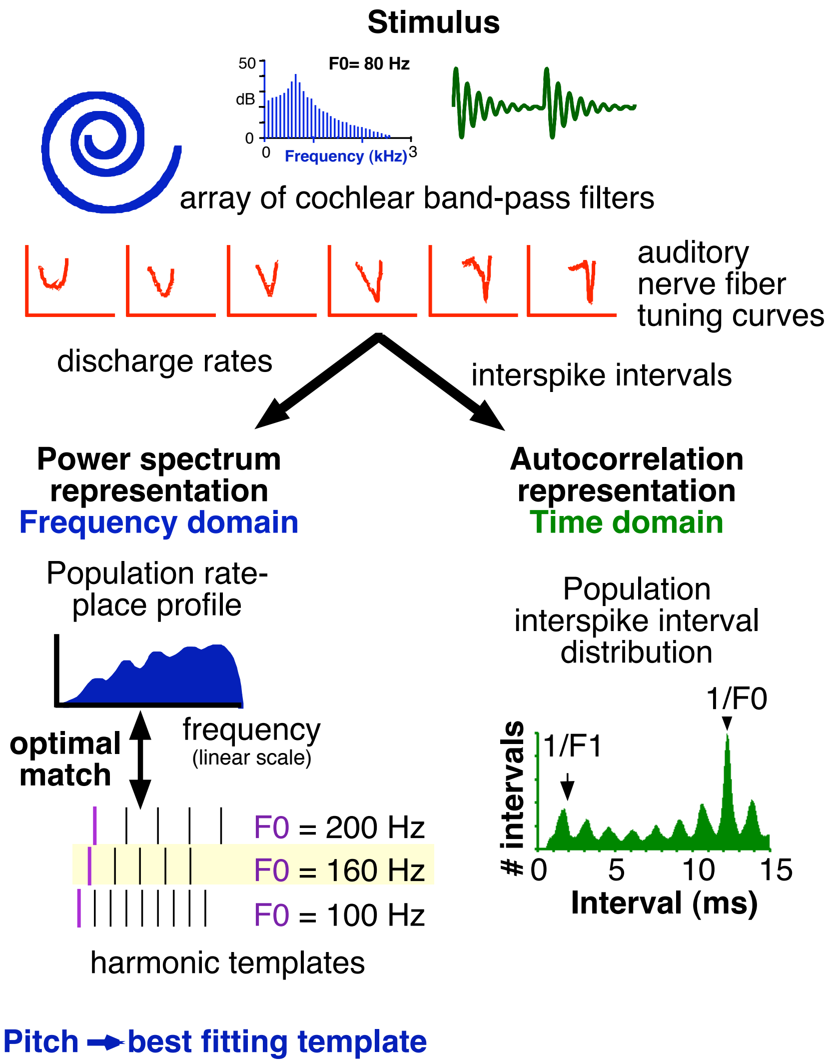
–Operate in frequency domain

–Recognize harmonic relations on resolved components

3. Temporal pattern theories

–Operate in time domain

–Analyze interspike interval dists.



“Virtual” pitch: F0-pitch as pattern completion

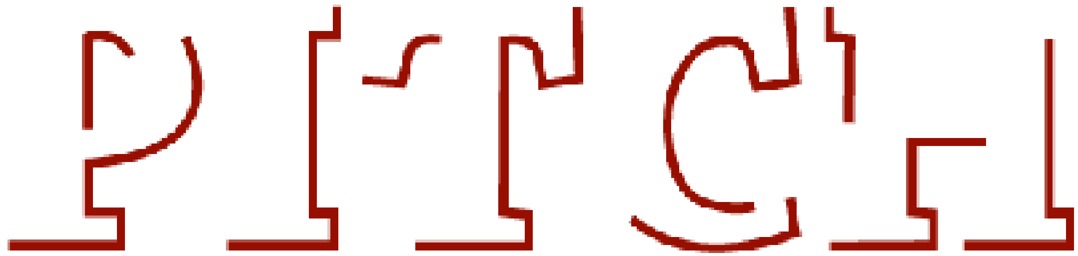
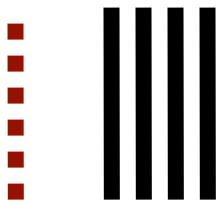


Figure by MIT OpenCourseWare.



“Missing fundamental“ analogy to illusory contour

Psychological perspectives on pitch

Analytical: **break sounds into frequencies** (perceptual atoms, features),

then analyze patterns (templates, combinations)

(British empiricism; machine perception)

Relational: **extract invariant relations from patterns**

(Gestaltists, Gibsonians, temporal models)

Nativist/rationalist: mechanisms for pitch are given by innate knowledge and/or computational mechanisms

differences re: how recently evolved these are

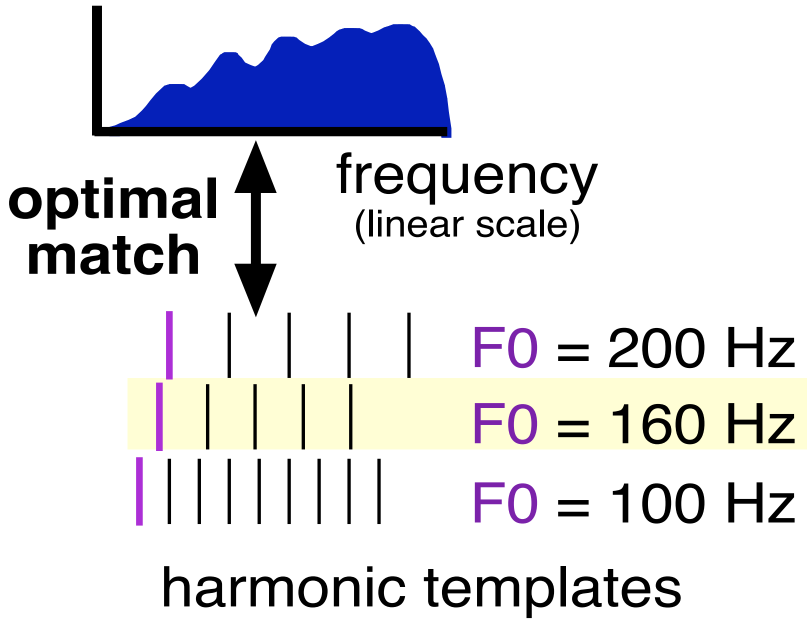
Associationist: mechanisms for pitch (e.g. templates) must be acquired through experience (ontogeny, culture)

Interactionist: (Piaget) interaction between native faculties and

Spectral pattern theories

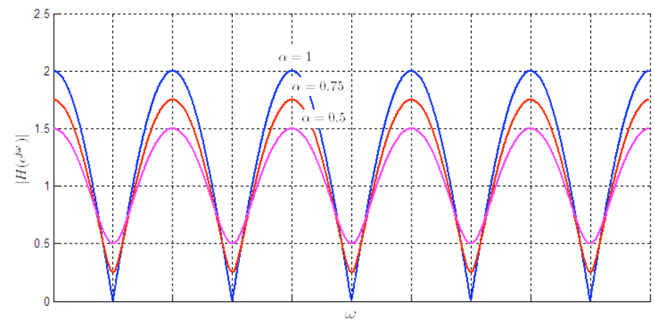
- Not the lowest harmonic
- Not simple harmonic spacings
- Not waveform envelope or peak-picking (pitch shift exps by Schouten & de Boer)
- Must do a real harmonic analysis of spectral fine structure to find common denominator, which is the fundamental frequency (comb filtering works)
- Terhardt: find common subharmonics
- Wightman: autocorrelation of spectra
- Goldstein, Houtsma: match spectral excitation pattern to harmonic templates
- SPINET: Use lateral inhibition/center-surround then fixed neural net to generate equivalence classes
- Barucha: adaptive connectionist networks for forming harmonic associations (hear many harmonic exemplars; problems with F0 range --

Central spectrum



Pitch → best fitting template

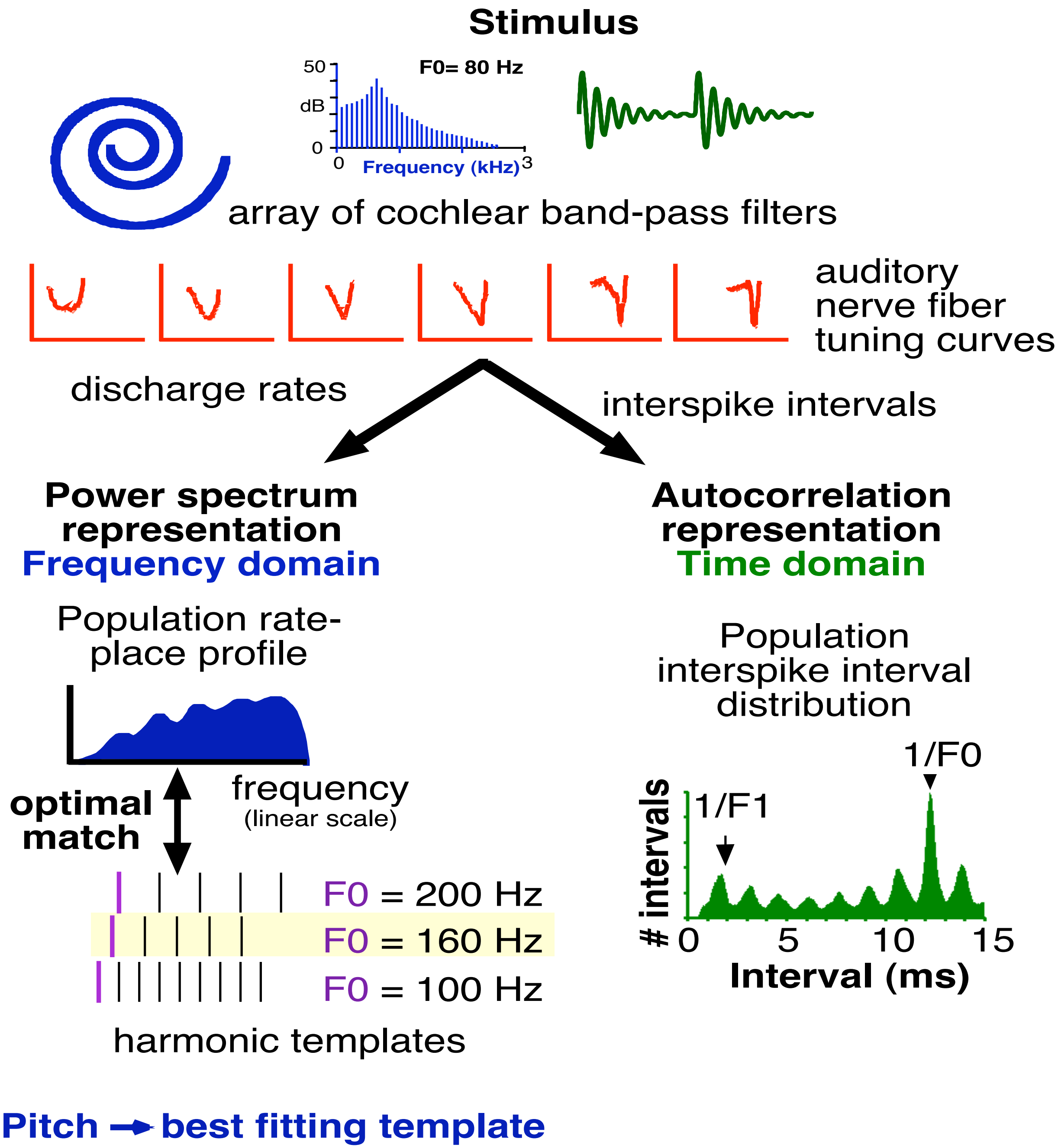
Output of comb filters



Spectral pattern analysis vs. temporal pattern analysis

Note: Some models, such as Goldstein's use interspike interval information to first form a Central Spectrum which is then analyzed using harmonic spectral templates.

There are thus dichotomies
1) between use of time and place information as the basis of the central representation, and
2) use of spectral vs. autocorrelation-like central representations



Resolution of harmonics

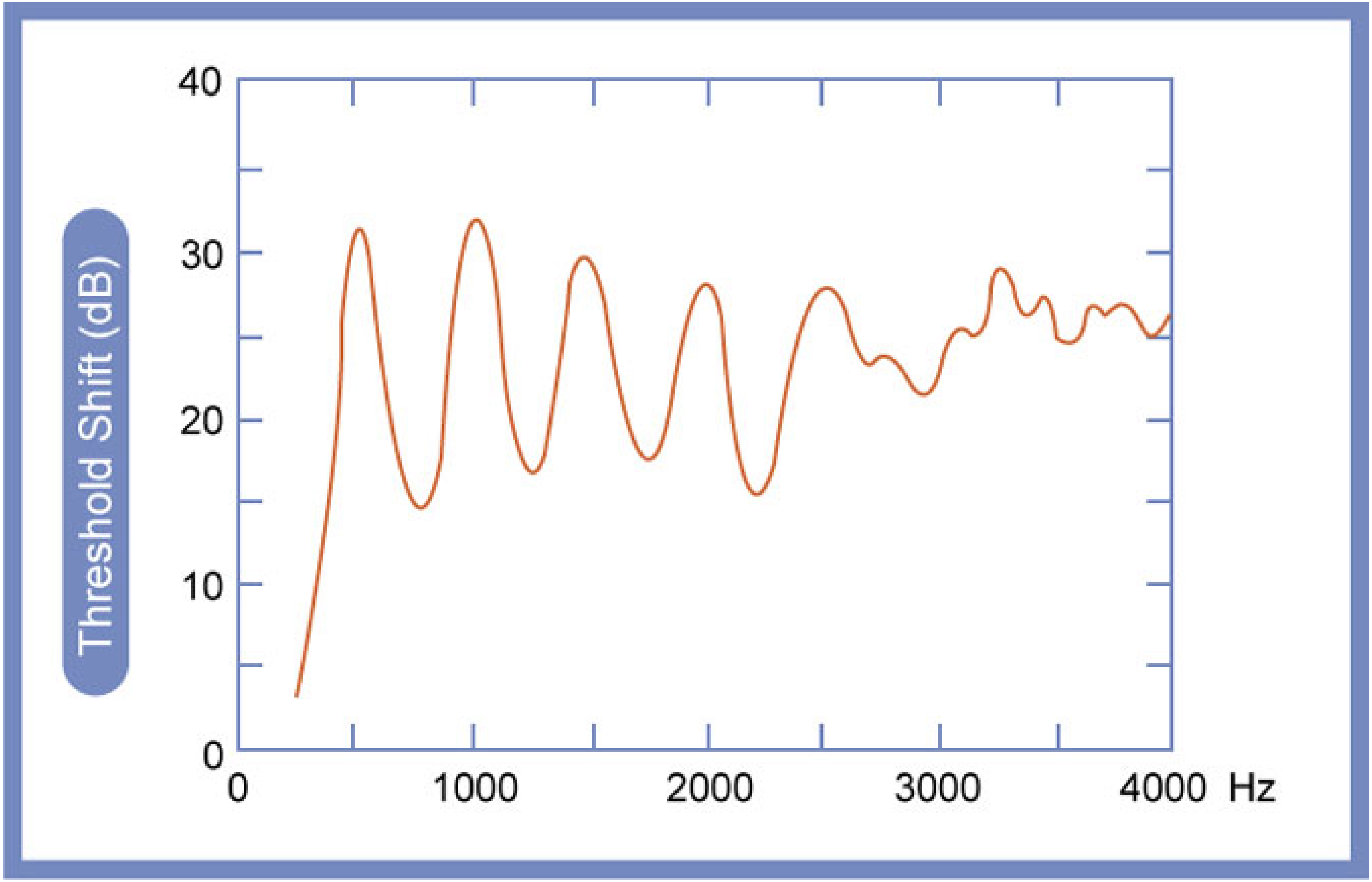


Figure by MIT OpenCourseWare.

Periodic sounds produce distinct pitches

Many different sounds produce the same pitches

Strong pitches

Strong
- Pure tones
- Harmonic complexes
- Iterated noise

Weaker
- High harmonics
- Narrowband noise

Very weak
- AM noise
- Repeated noise

Weaker low pitches

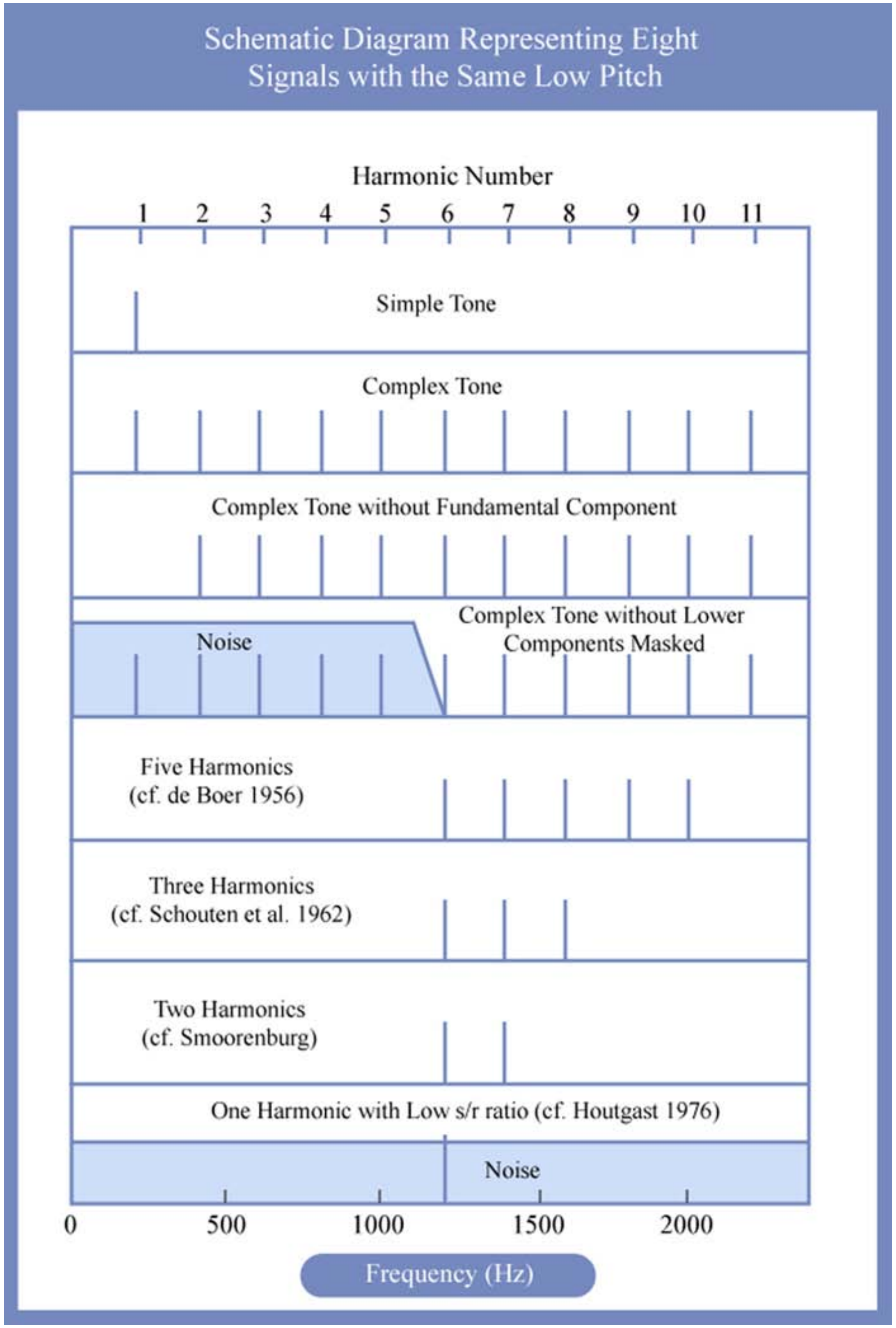


Figure by MIT OpenCourseWare.

Goldstein's

Figure removed due to copyright restrictions.
Diagram of periodicity pitch as harmonic frequency pattern recognition.
figure 3 in Goldstein, J. L., et al. "Verification of the Optimal Probabilistic Basis of Aural Processing in Pitch of Complex Tones." *J Acoust Soc Am* 63 (1978): 486-510.
http://dx.doi.org/10.1121/1.381749

Julius Goldstein references

Models for pure tone pitch discrimination, low pitches of complex tones, binaural pitches, and aural distortion products

Goldstein JL (1970) Aural combination tones. In: Frequency Analysis and Periodicity Detection in Hearing (Plomp R, Smoorenburg GF, eds), pp 230-247. Leiden: A. W. Sijthoff.

Goldstein JL (1973) An optimum processor theory for the central formation of the pitch of complex tones. J Acoust Soc Am 54:1496-1516.

Goldstein JL, Kiang NYS (1968) Neural correlates of the aural combination tone $2f_1$-f_2. IEEE Proc 56:981-992.

Goldstein JL, Srulovicz P (1977) Auditory-nerve spike intervals as an adequate basis for aural frequency measurement. In: Psychophysics and Physiology of Hearing (Evans EF, Wilson JP, eds). London: Academic Press.

Goldstein JL, Buchsbaum G, First M (1978a) Compatibility between psychophysical and physiological measurements of aural combination tones. J Åcoust Soc Am 63:474-485.

Goldstein JL, Buchsbaum G, Furst M (1978b) Compatibility between psychophysical and physiological measurements of aural combination tones... Journal of the Acoustical Society of America 63:474-485.

Goldstein JL, Gerson A, Srulovicz P, Furst M (1978c) Verification of the optimal probabilistic basis of aural processing in pitch of complex tones. J Acoust Soc Am 63:486-510.

H. L. Duifhuis and L. F. Willems and R. J. Sluyter (1982,) Measurement of pitch in speech: An implementation of Goldstein's theory of pitch perception,. jasa, 71,: 1568--1580.

Houtsma AJM, Goldstein JL (1971) Perception of musical intervals: Evidence for the central origin of the pitch of complex tones. In: M.I.T./R.L.E.

Houtsma AJM, Goldstein JL (1972) The central origin of the pitch of complex tones: Evidence from musical interval recognition. J Acoust Soc Am 51:520-529.

P. Srulovicz and J. Goldstein (1983) A central spectrum model: A synthesis of auditory nerve timing and place cues in monoaural communication offrequency spectrum,. jasa, 73,: 1266--1276,.

Srulovicz P, Goldstein JL (1977) Central spectral patterns in aural signal analysis based on cochlear neural timing and frequency filtering. In: IEEE, p 4 pages. Tel Aviv, Israel.

Srulovicz P, Goldstein JL (1983) A central spectrum model: a synthesis of auditory-nerve timing and place cues in monaural communication of frequency spectrum. J Acoust Soc Am 73:1266-1276.

Terhard's method of common subharmonics

Spectral vs. virtual pitch: duplex model

Virtual pitch computation:

1. Identify frequency components, e.g. 1000, 1200, 140
2. Find common subharmonics

3. Strongest common subharmonic after F0 weighting is the virtual pitch

Terhardt's model has been extended by Parncutt to cover pitch multiplicity and fundamental bass of chords

Terhard's method

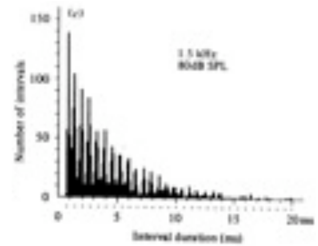
1. Identify frequency components, e.g. 1000, 1200, 1400
2. Find common subharmonics, f/n for n = 1, 2, 3, ...
 f=1000: 500, 333, 250,   200, 166, 143, 125, 111, 100, ...
 f=1200: 600, 400, 300, 240,   200, 171, 150, 133, 109, 100, f=1400: 700, 466, 350, 280, 233, 200, 175, 155, 140, ...100, ...
3. Strongest common subharmonic after F0 weighting, which biases against low F0s, is the virtual pitch

Parallels with all-order interspike interval models

Each harmonic generates intervals at its subharmonics

Adding together all the intervals and finding the most common intervals therefore finds the common subharmonics (F0/n)

F0-weighting is achieved by limiting interval length



Terhardt references

Terhardt E (1970) Frequency analysis and periodicity detection in the sensations of roughness and periodicity pitch. In: Frequency Analysis and Periodicity Detection in Hearing (Plomp R, Smoorenburg GF, eds). Leiden: A. W. Sijthoff.

Terhardt E (1974a) On the perception of periodic sound fluctuations (roughness). Acustica 30:201-213.

Terhardt E (1974b) Pitch, consonance, and harmony. J Acoust Soc Am 55:1061-1069.

Terhardt E (1977) The two-component theory of musical consonance. In: Psychophysics and Physiology of Hearing (Evans EF, Wilson JP, eds), pp 381-390. London: Academic Press.

Terhardt E (1979) Calculating virtual pitch. Hearing Research 1:155-182.

Terhardt E (1984) The concept of musical consonance: a link between music and psychoacoustics. Music Perception 1:276-295.

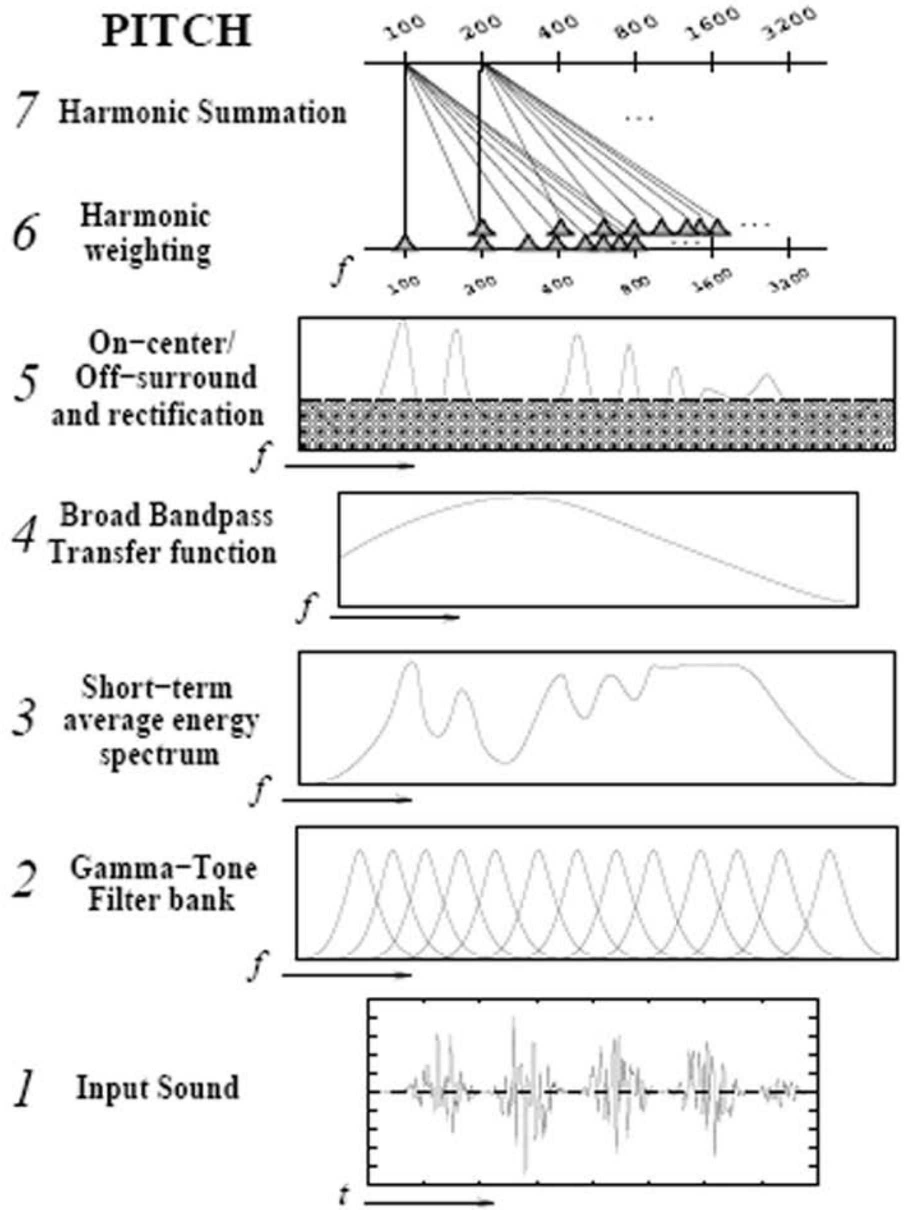
Terhardt E, Stoll G, Seewann M (1982a) Pitch of complex signals according to virtual-pitch theory: test, examples, and predictions. J Acoust Soc Am 71:671-678.

Terhardt E, Stoll G, Seewann M (1982b) Algorithm for extraction of pitch and pitch salience from complex tonal signals. J Acoust Soc Am 71:679-688.

Parncutt R (1989) Harmony: A Psychoacoustical Approach. Berlin: Springer-Verlag

SPINET:
Cohen Grossberg, Wyse JASA

Fixed neural network: connection weights arranged so as to form pitch-equivalence classes



Source: Cohen, M. A., S. Grossberg, and L. L. Wyse. "A Spectral Network Model of Pitch Perception." Technical Report CAS/CNS TR-92-024, Boston University. Also published in *J Acoust Soc Am* 98, no. 2 part 1 (1995): 862-79.

Neural networks

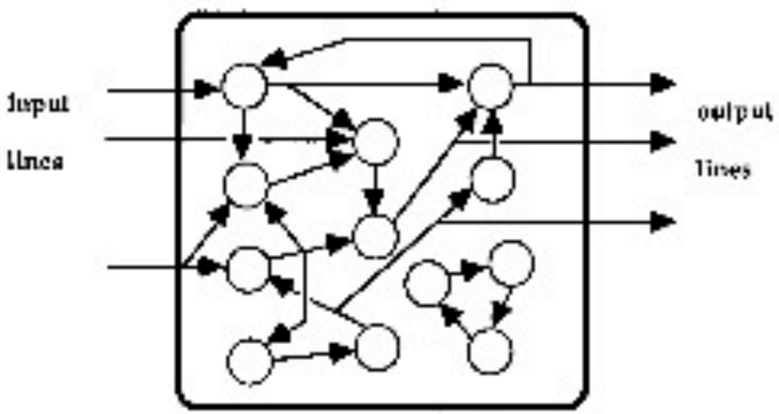
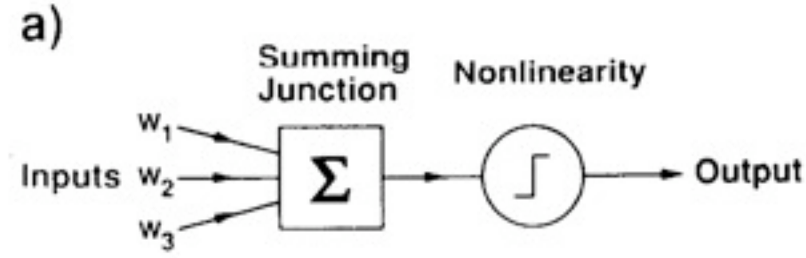
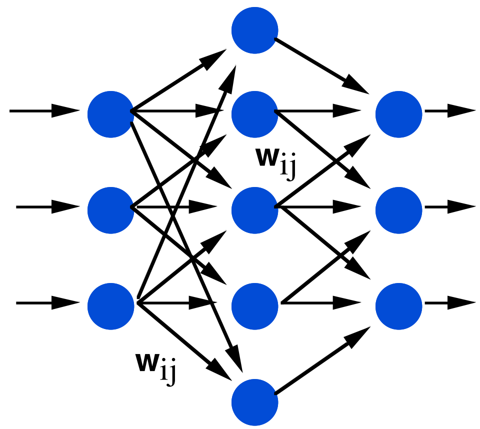


Figure 5. A neural network viewed as a system. The input at time t is the pattern of firing on the input lines, the output is the pattern of firing on the output lines; and the internal state is the vector of firing rates of all the neurons of the network.

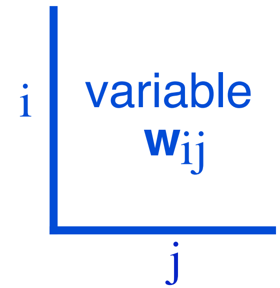
Source: Arbib, M. A., ed. *The Handbook of Brain Theory and Neural Networks*.
2nd ed. Cambridge MA: MIT Press, 2003. ISBN: 9780262011976.

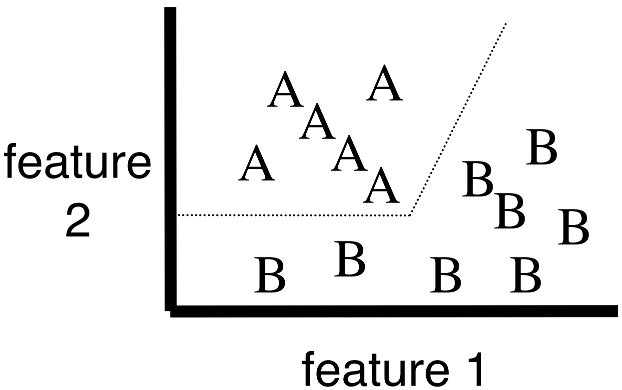




Purpose: group combinations of features into equivalence classes

Rate integrators





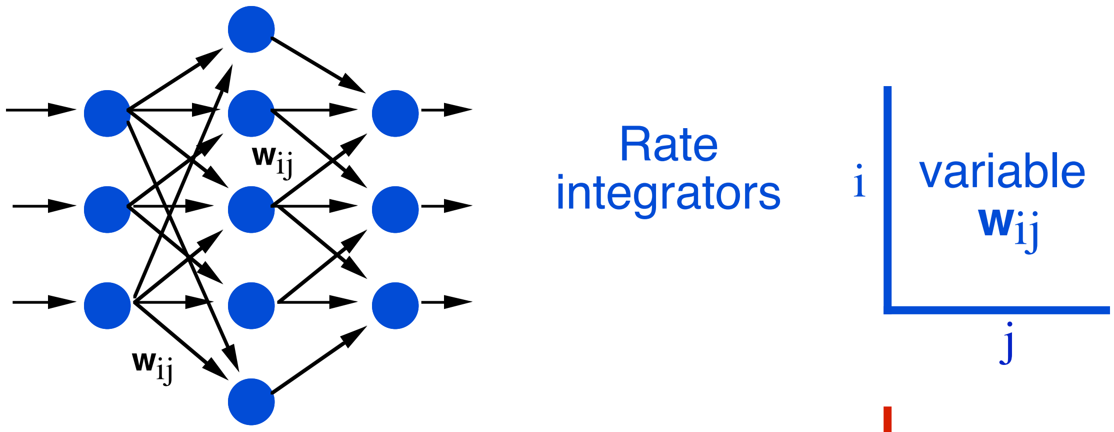
Adaptive adjustment of synaptic weights so as to properly classify objects by their feature combinations

Neural networks

<u>Connectionist networks</u>

Purely spatial correlators

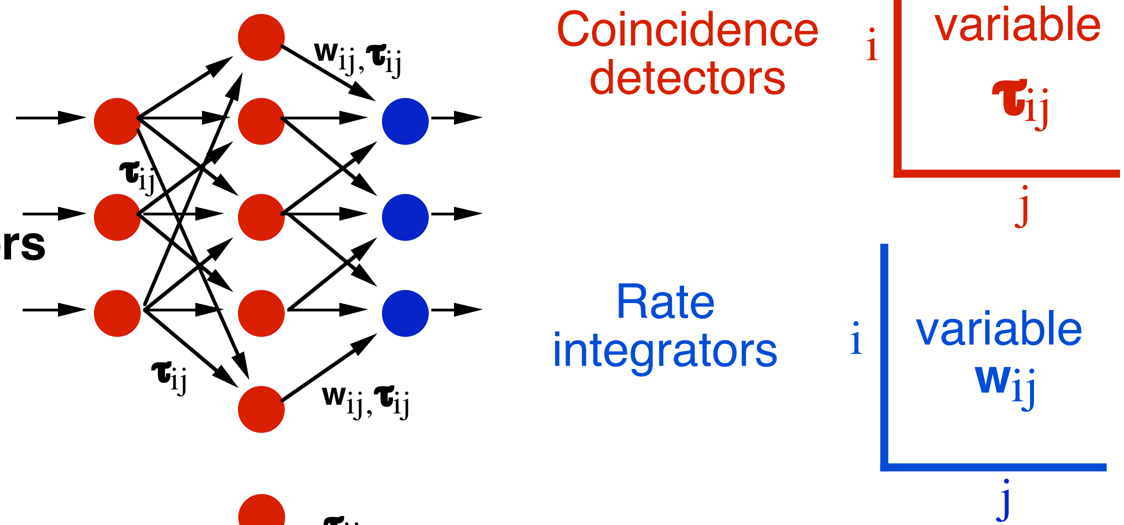
Place-Place mappings



<u>Time-delay networks</u>

Spatio-temporal correlators

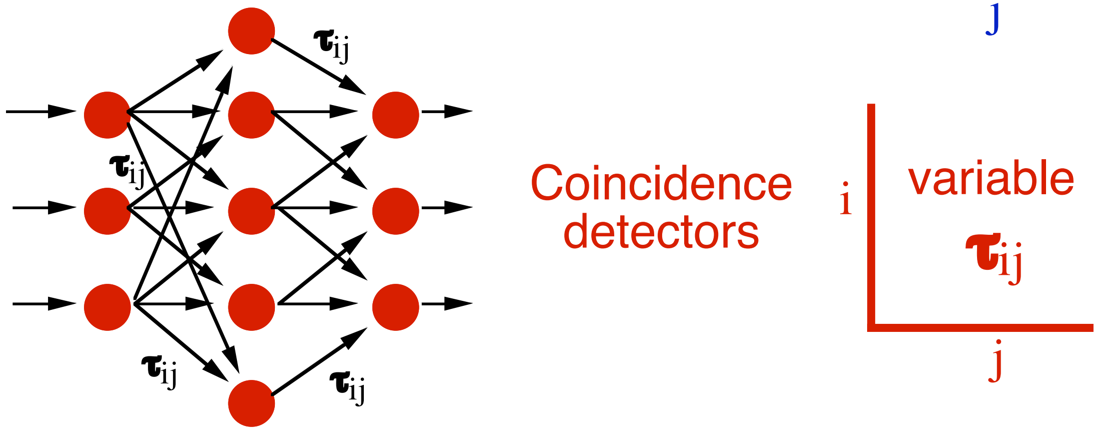
Time-Place mappings



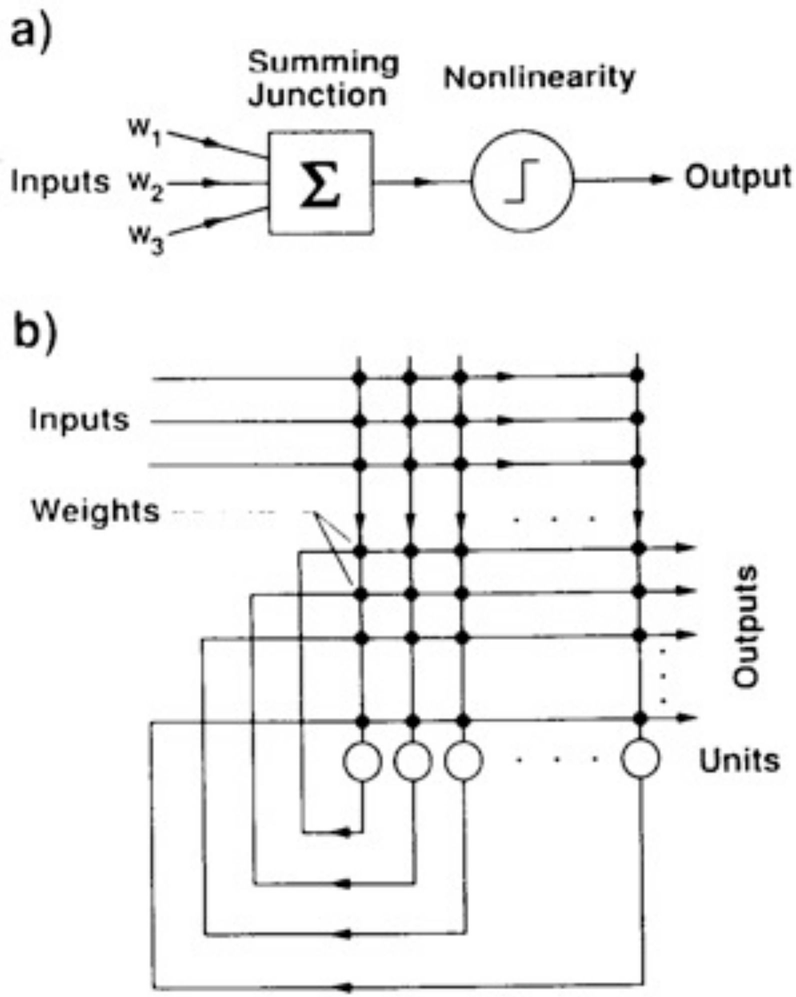
<u>Timing nets</u>

Temporal correlators

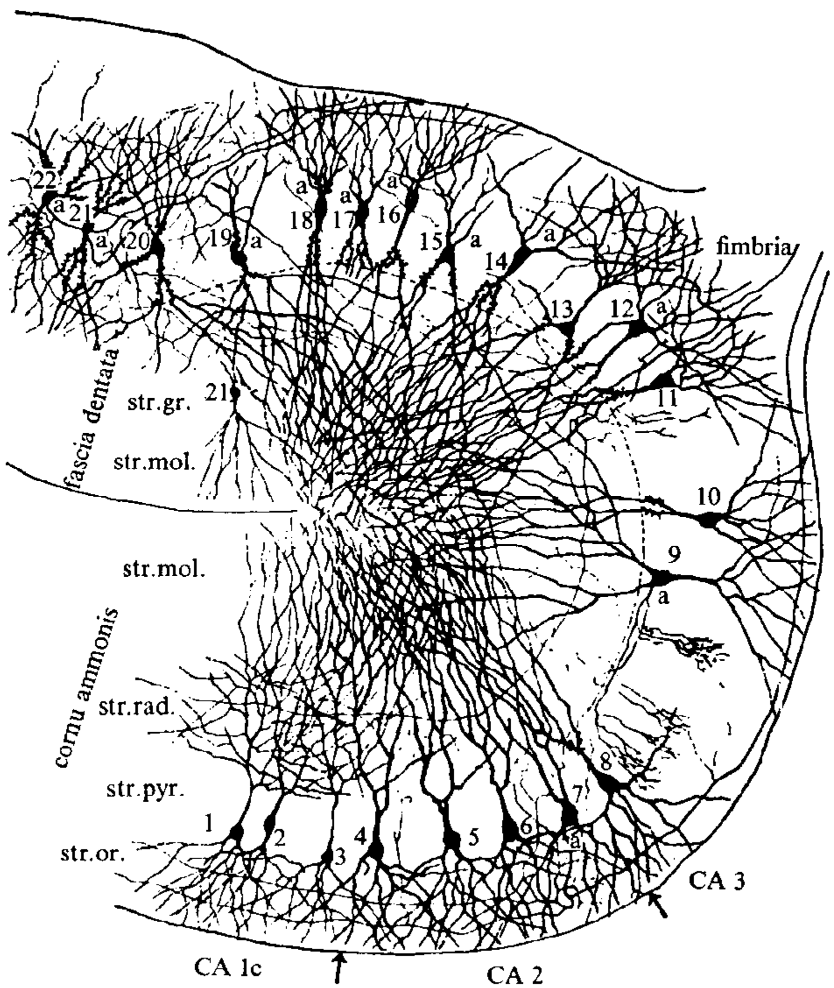
Time-Time mappings



Hippocampus as a connectionist architecture



Courtesy of the MIT Press. Used with permission.
Source: Fig 3.14 in Churchland, P. and T. Sejnowski. *The Computational Brain*. Cambridge, MA: MIT Press, 1992. ISBN: 9780262531207.

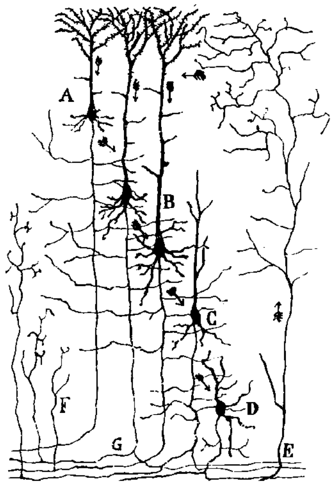


Lorente de No

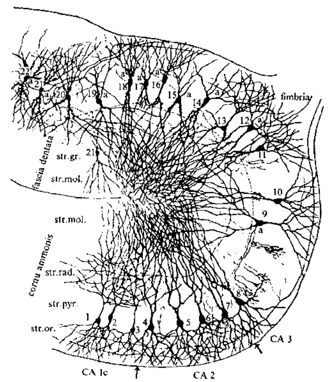
Auto-associative network

(rate-channel code)

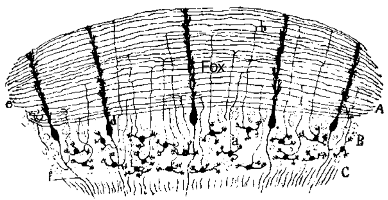
Cerebral cortex



Hippocampus



Cerebellar cortex



Cajal

CORTICAL STRUCTURES

RATE CODES

TIME CODES

PURELY SPATIAL CORRELATORS

effective connectivity

SPATIO-TEMPORAL CORRELATORS

effective connectivity
&
timing relations

Spectral pattern theories - pros & cons

Do make use of frequency tuning properties of auditory elements

No clear neural evidence of narrow (< 1/3 octave) frequency channels in low-BF regions (< 2 kHz) (re: mistuning detection)

Operate on perceptually-resolved harmonics

Do not explain low pitches of unresolved harmonics

Require templates or harmonic pattern analyzers

Little neural evidence for resolved low harmonics or req. analyzers

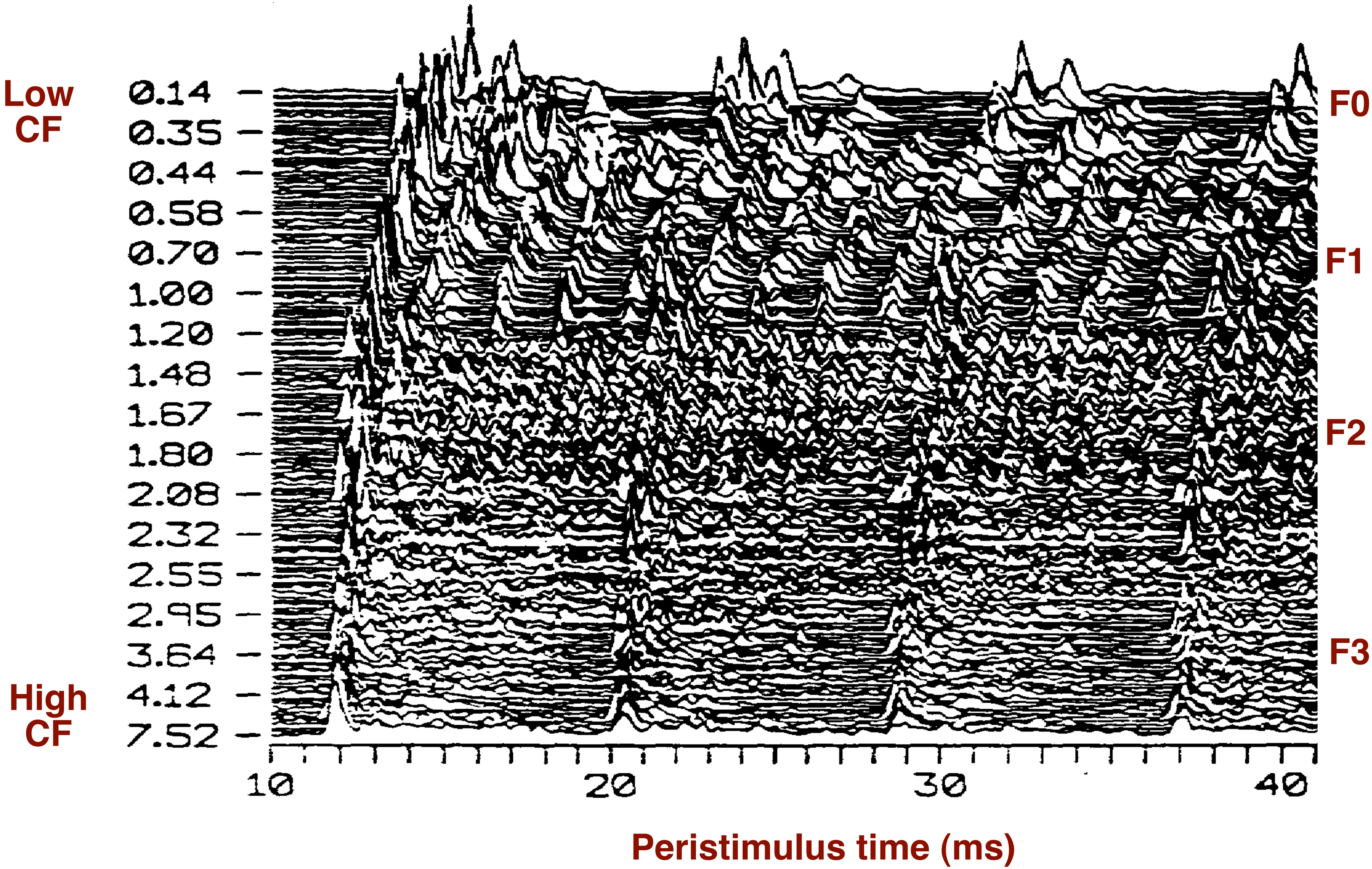
Possible evidence for F0-detectors (Bendor & Wang(2005)

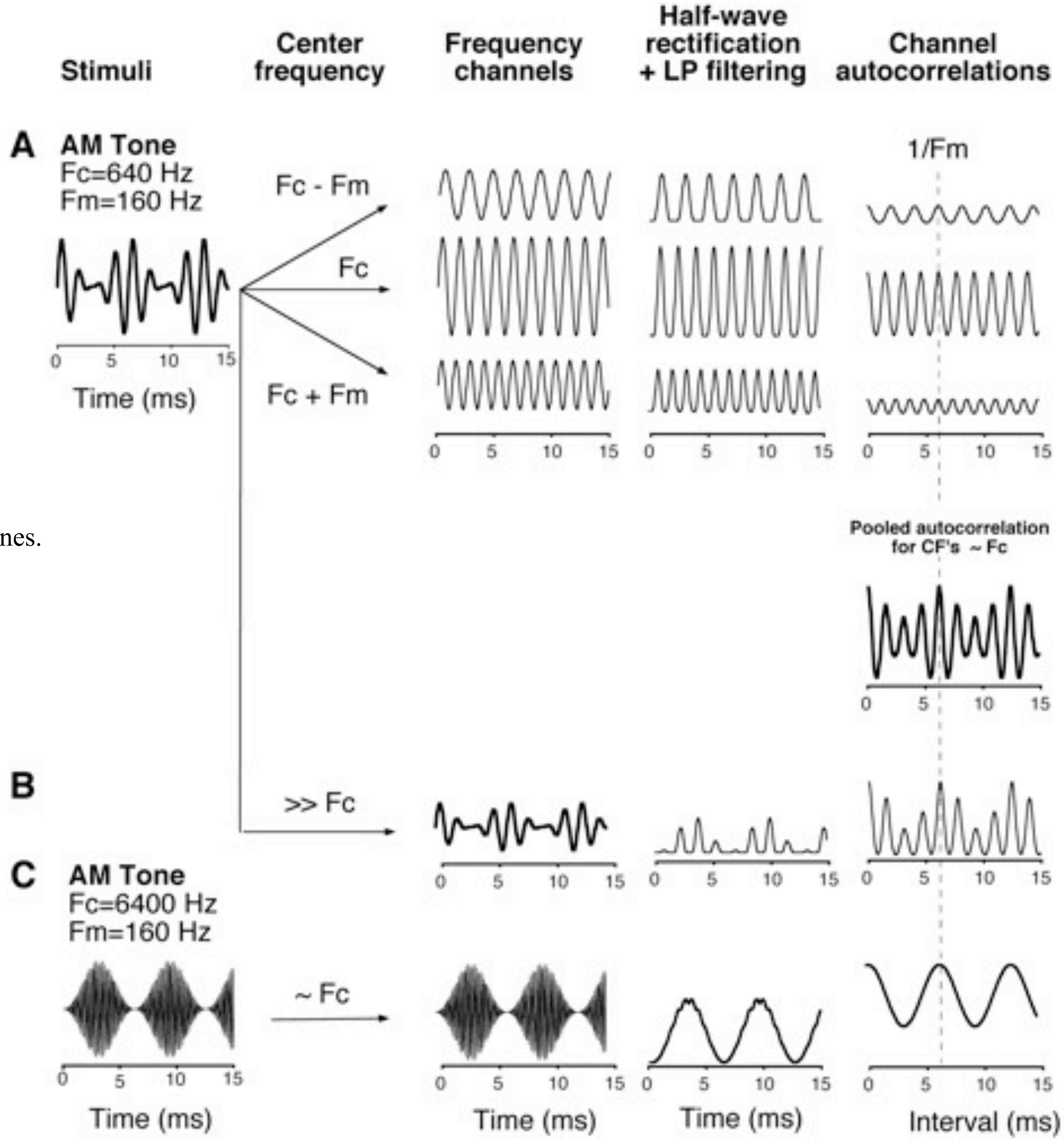
Problems w. templates: relative nature of pitch

Do not explain well existence region for F0

Vowel Formant Regions

Time domain analysis of auditory-nerve fiber firing rates.
Hugh Secker-Walker & Campbell Searle, J. Acoust. Soc. 88(3), 1990
Neural responses to /da/ @ 69 dB SPL from Miller and Sachs (1983)





Source: Cariani, P. A., and B. Delgutte. "Neural Correlates of the Pitch of Complex Tones. I. Pitch and Pitch Salience." *J Neurophysiol* 76 (1996): 1698-1716. [0022-3077/96].

Temporal pattern theories

Image removed due to copyright restrictions.

See Fig. 2, "Schematic representation of the origination of low pitch." In van Noorden, L. "Two Channel Pitch Perception." Clynes, M., ed. *Music, Mind and Brain.* New York, NY: Plenum, 1982.

Schouten's temporal theory (1940's) depended on interactions between unresolved (high) harmonics. It was displaced by discovery of dominance region and binaural combination pitches in the 1960's. The idea persists, however in the form of spectral mechanisms for resolved harmonics and temporal ones for unresolved harmonics.

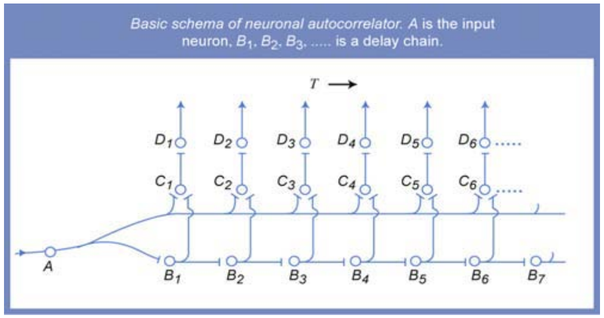
van Noorden (1982)

Σ First-order intervals (renewal density)

Σ All-order intervals (temporal autocorrelation)

Licklider (1951)

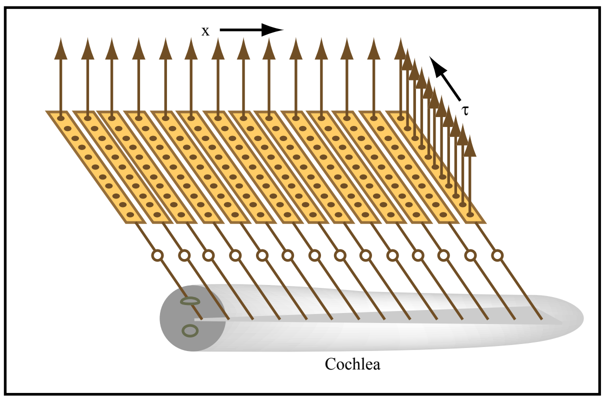
Meddis & Hewitt (1991)



Please see Figure 1 in Meddis, R., and M. J. Hewitt. Virtual Pitch and Phase Sensitivity of a Computer Model of the Auditory Periphery. I. Pitch identification. *J Acoust Soc Am* 89, no. 6 (1991): 2866-2882.

Image removed due to copyright restrictions.
See Moore, B. C. J. *An Introduction to the Psychology of Hearing*. 5th ed. San Diego, CA: Academic Press, 2003.

Moore (1982)
Σ First-order intervals



Images by MIT OpenCourseWare.

Interval-based theories of pitch

Image removed due to copyright restrictions.
Moore, B. C. J. *An Introduction to the Psychology of Hearing*. 5th ed.
San Diego, CA: Academic Press, 2003.

Moore (1982)

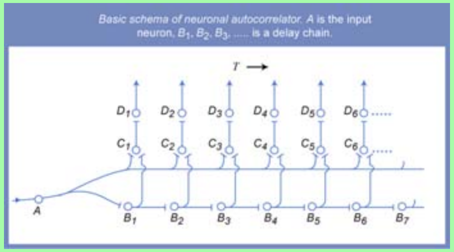
Image removed due to copyright restrictions.
See Fig. 2, "Schematic representation of the origination of low pitch." In van Noorden, L. "Two channel pitch perception." In Clynes, M. ed. *Music, Mind and Brain*
New York, NY: Plenum Press, 1982.

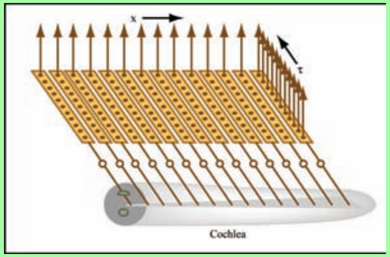
van Noorden (1982)

First-order intervals
(renewal density)

All-order intervals
(temporal autocorrelation)

Licklider (1951)





Figures by MIT OpenCourseWare.

Meddis & Hewitt (1991)

Image removed due to copyright restrictions.
Figure 1 in Meddis, R., and M. J. Hewitt.
"Virtual Pitch and Phase Sensitivity of a Computer Model of the Auditory Periphery. I. Pitch identification."
J Acoust Soc Am 89, no. 6 (1991): 2866-2882.

Licklider's (1951) duplex model of pitch perception

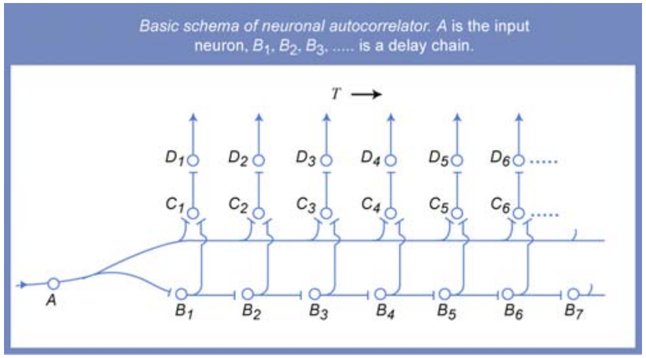


Figure by MIT OpenCourseWare.

Licklider's binaural triplex model

Image removed due to copyright restrictions.
Figure 5, "Schematic illustration of hypothetical auditory system."

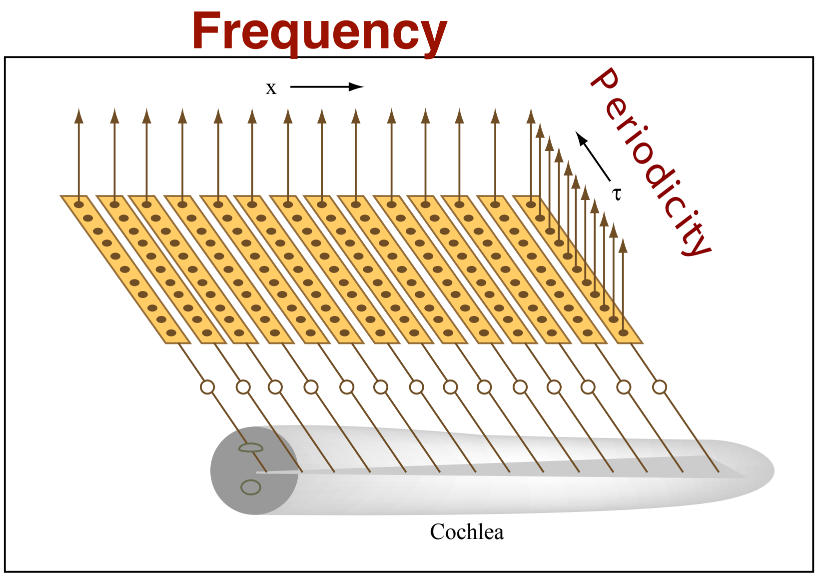
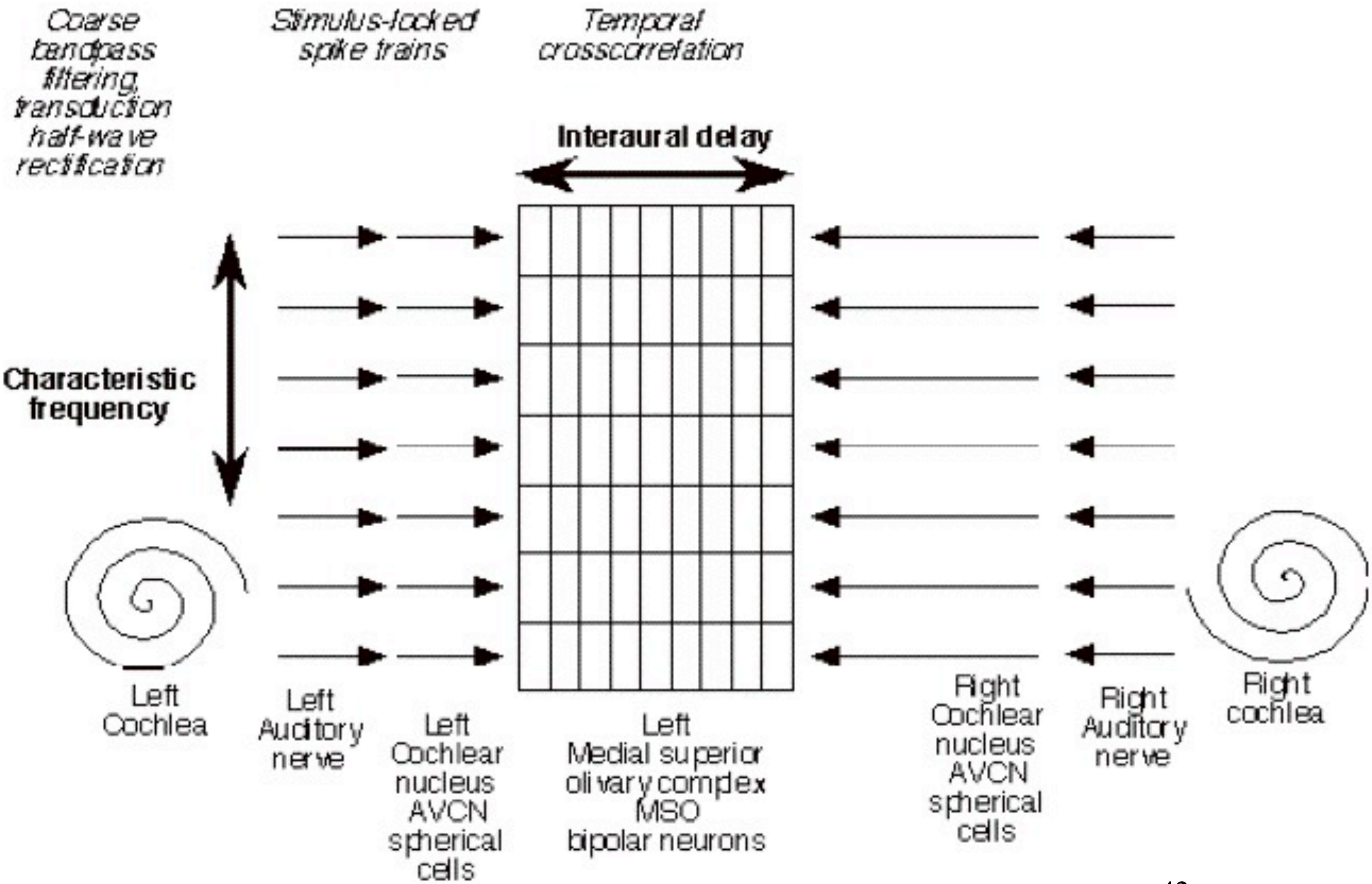


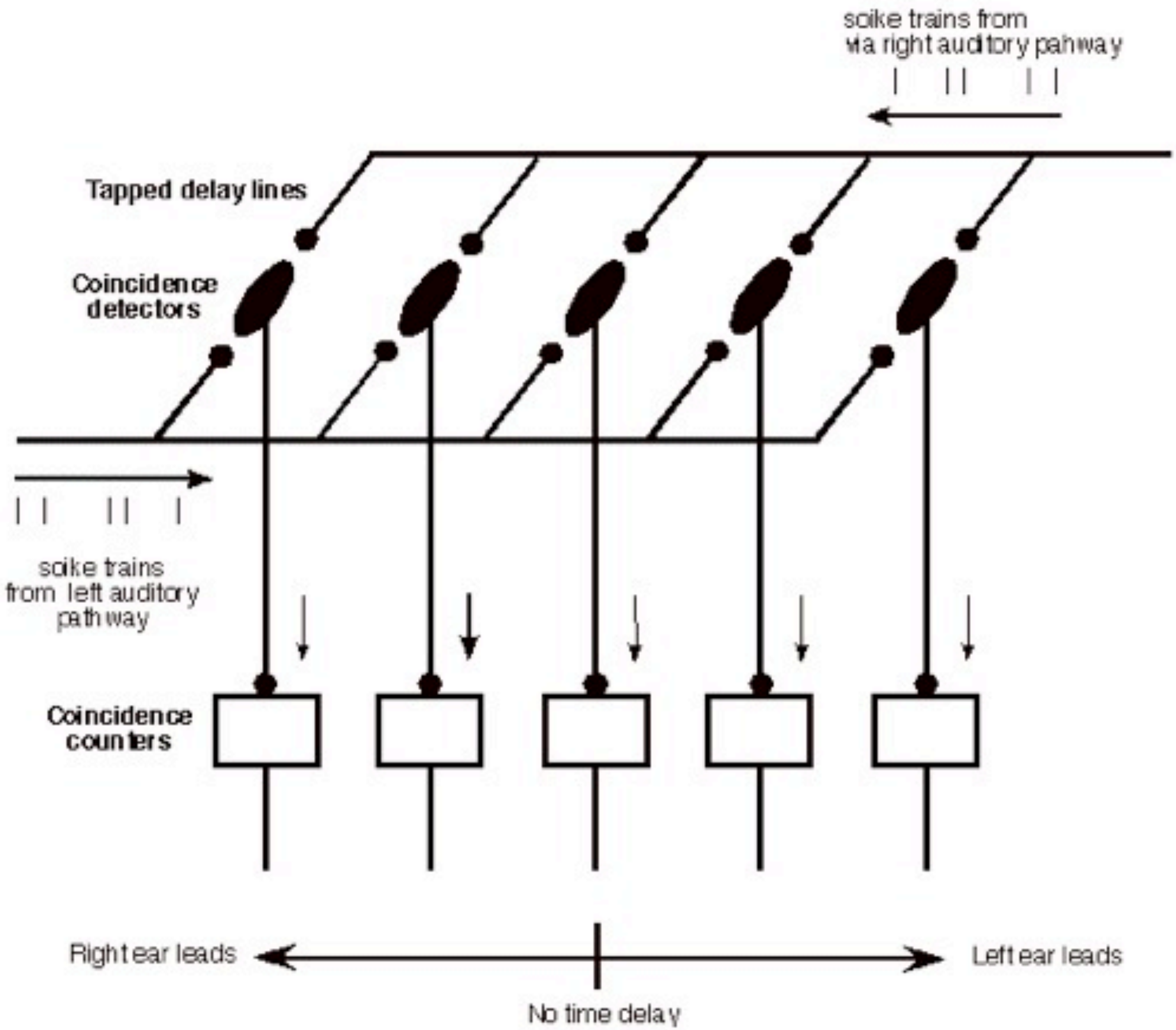
Figure by MIT OpenCourseWare.

J.C.R. Licklider (1959) "Three Auditory Theories" in Psychology: A Study of a Science, Vol. 1, S. Koch, ed., McGraw-Hill, pp. 41-144.

Basic plan of the Jeffress binaural crosscorrelator



Jeffress temporal correlation model for sound localization (1948)



Tapped delay lines: synaptic and transmission delays

Basic schema of neuronal autocorrelator. A is the input neuron, B_1, B_2, B_3, ... is a delay chain.

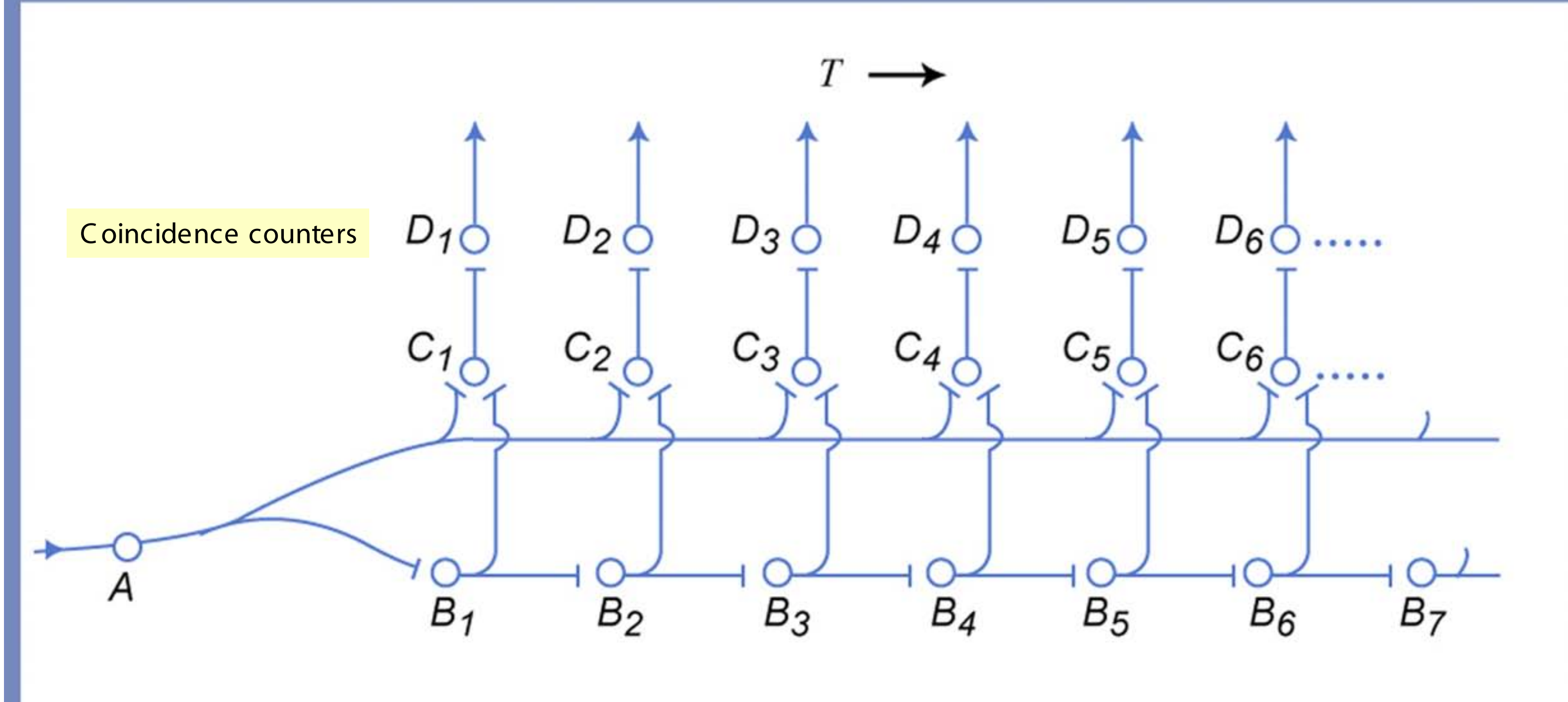


Figure by MIT OpenCourseWare.

Frequency

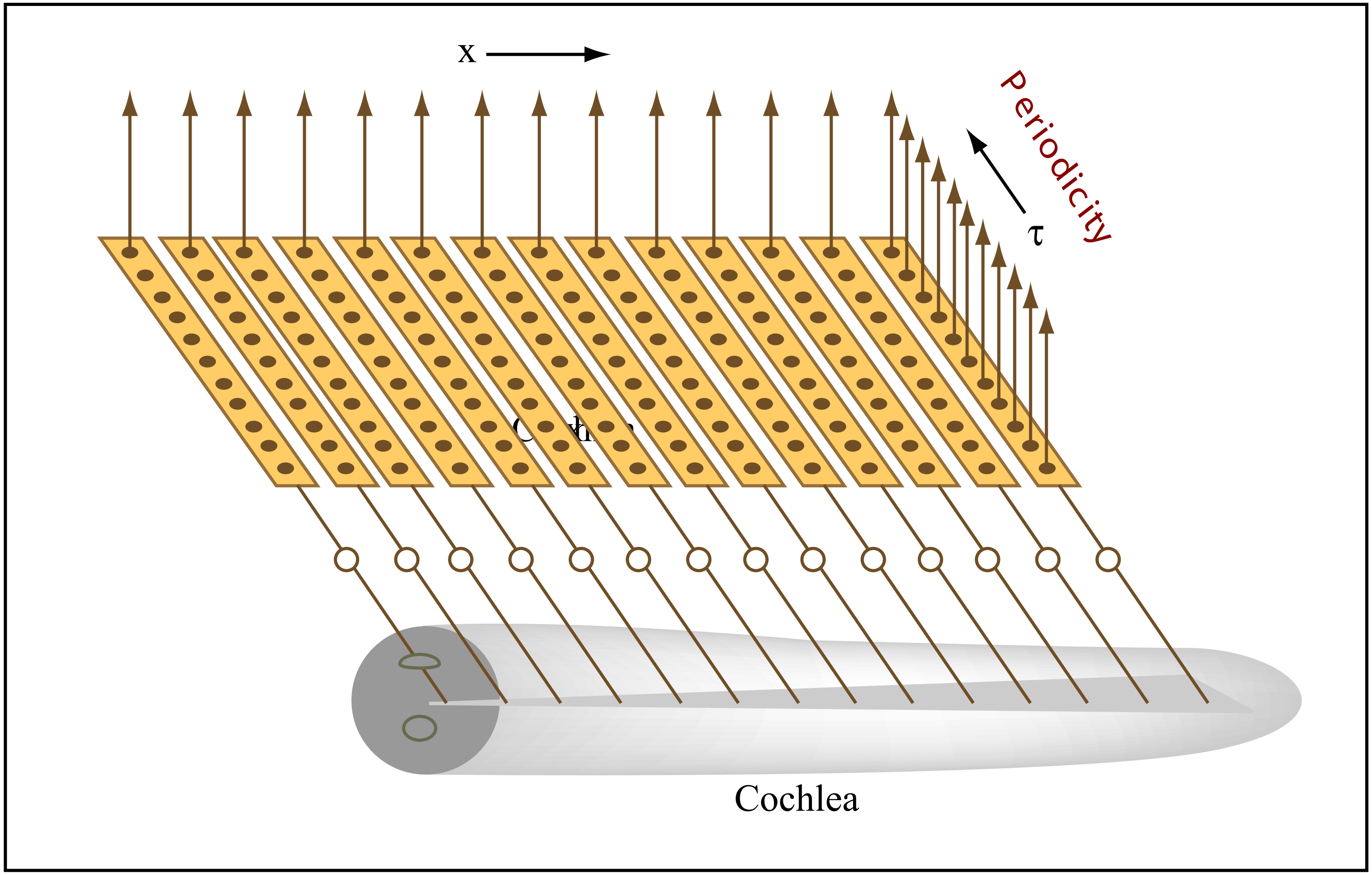
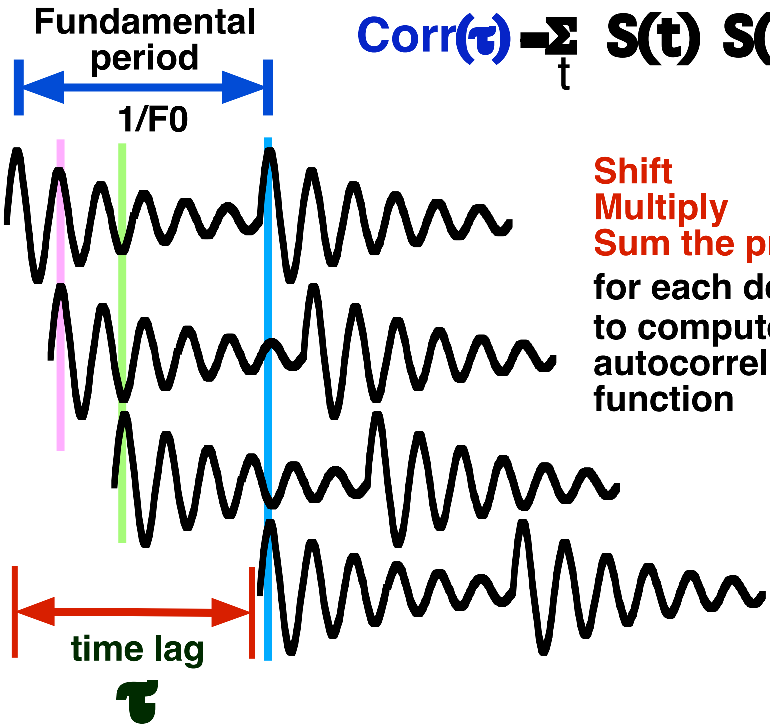


Figure by MIT OpenCourseWare.

Autocorrelation and interspike intervals

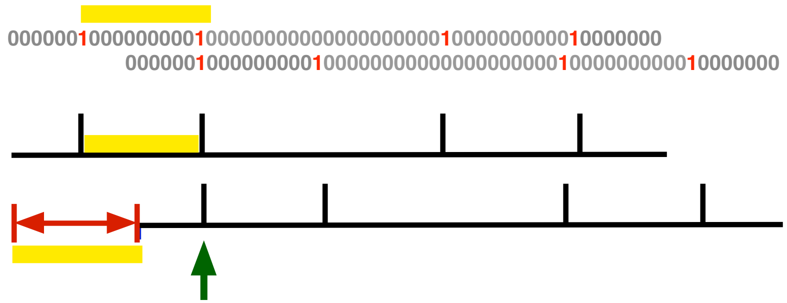
Autocorrelation functions



$$\mathrm{Corr}(\tau) = \sum_t S(t)\, S(t-\tau)$$

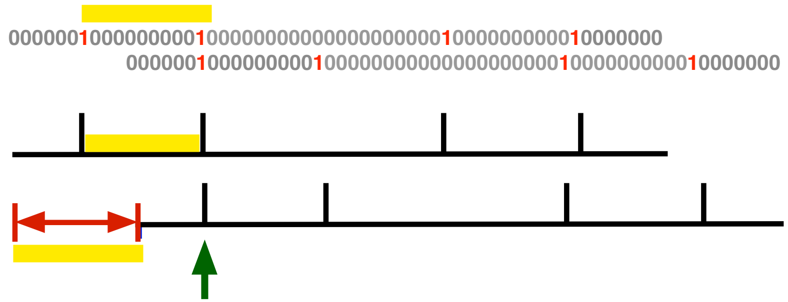
Shift
Multiply
Sum the products
for each delay τ
to compute
autocorrelation
function

Autocorrelations of spike trains = Histograms of all-order intervals



Delay lines, coincidence detectors, and autocorrelation

Autocorrelations of spike trains = Histograms of all-order intervals



Basic schema of neuronal autocorrelator. A is the input neuron, B_1, B_2, B_3, ... is a delay chain.

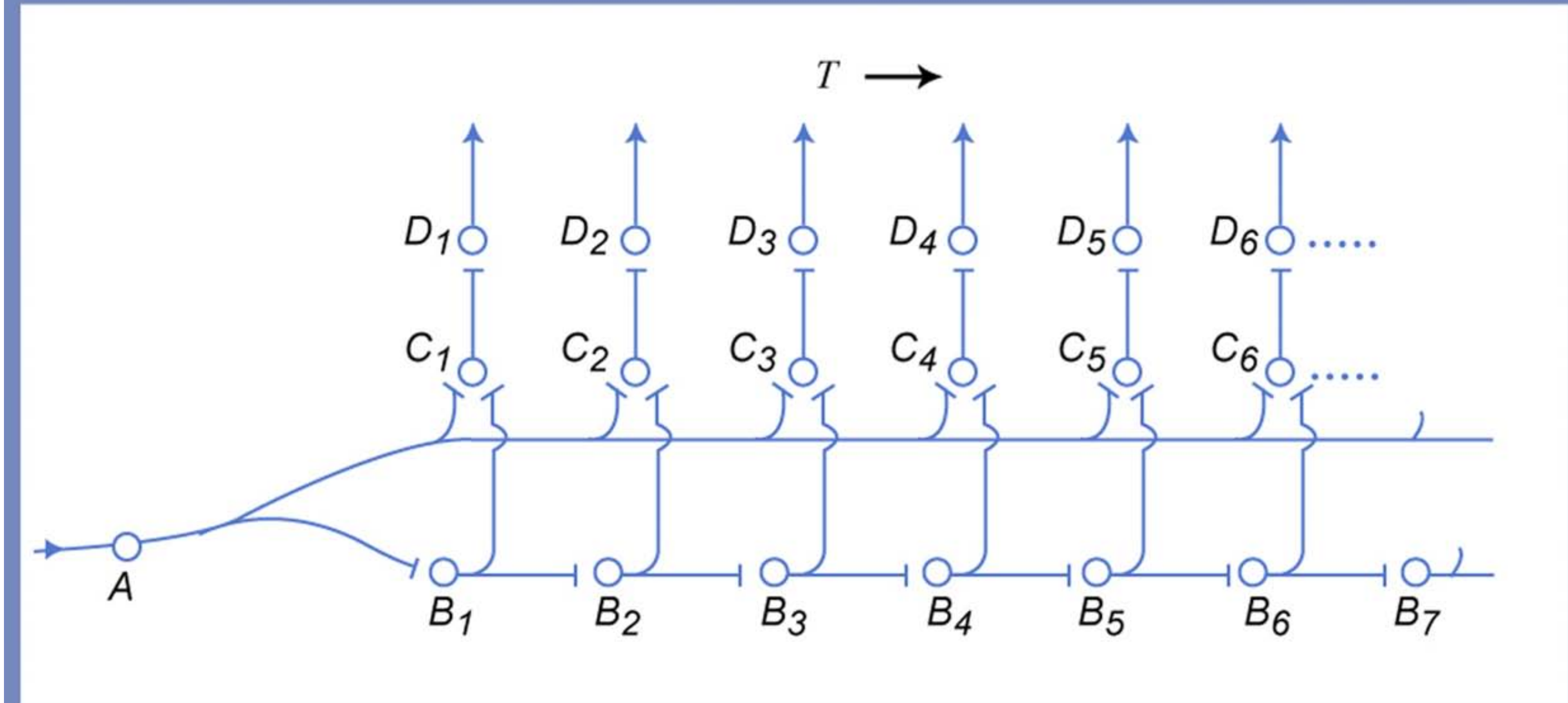


Figure by MIT OpenCourseWare.

Images removed due to copyright restrictions.
See Figure 6.16A-D in Lyon, R., and S. Shamma. "Auditory Representations of Timbre and Pitch." In *Auditory Computation*. Edited by R. R. Fay. New York, NY: Springer, 1996.

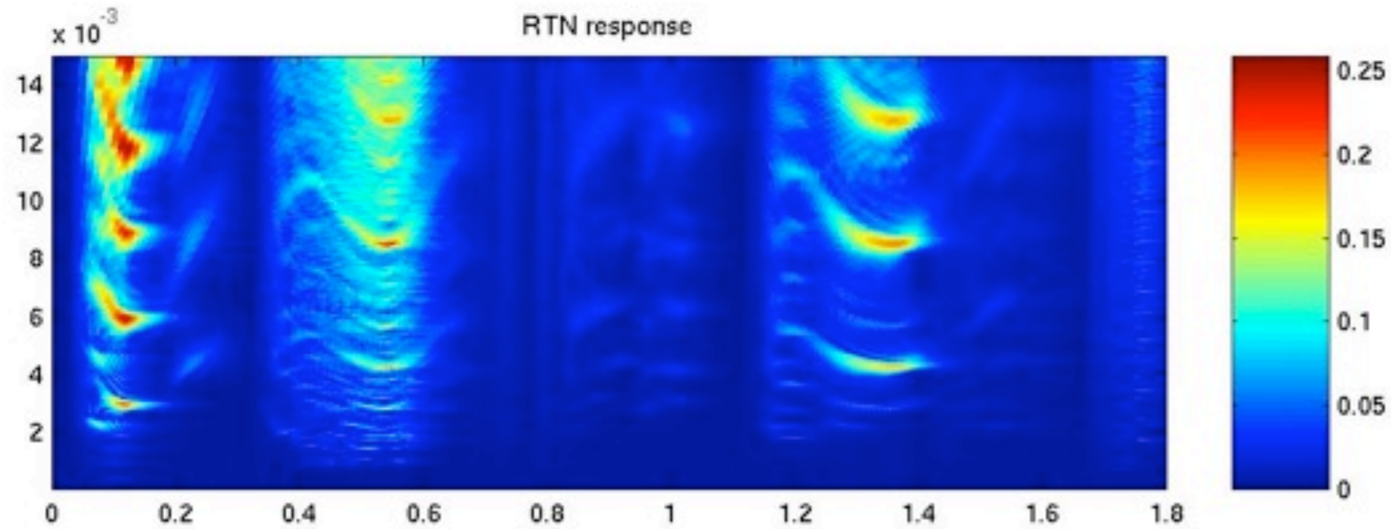
1950's Tape autocorrelator

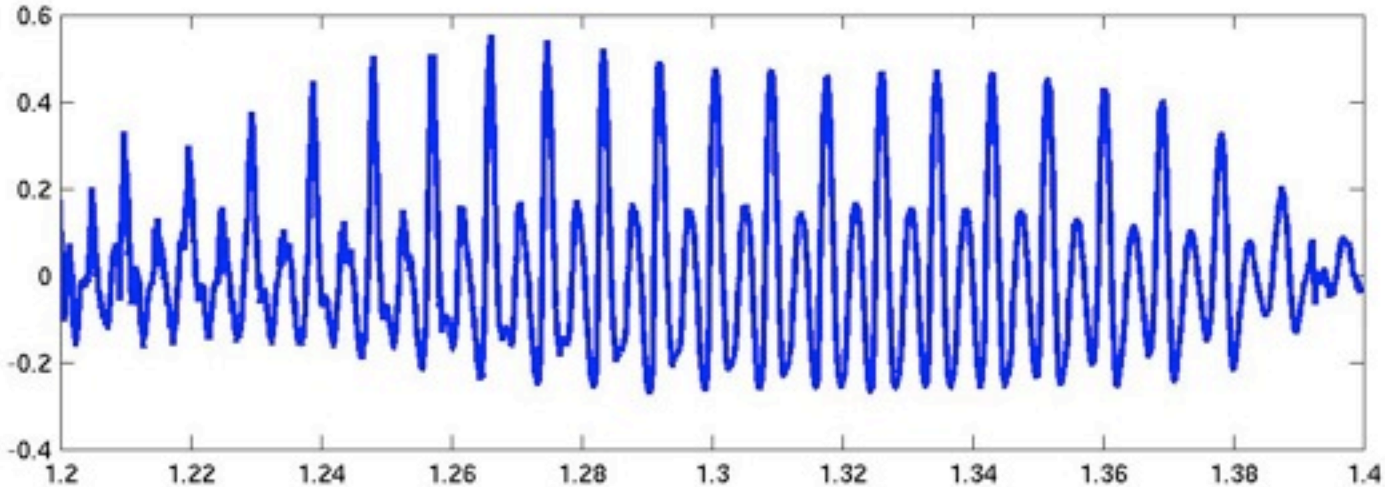
Images removed due to copyright restrictions.
Two photos of a tape autocorrelator machine (magnetic correlatograph).
See Plates 3.1 and 3.2 in Lange, F. H. *Correlation Techniques: Foundations and Applications of Correlation Analysis*. Iliffe, 1967.

Biddulph's speech autocorrelograms **(from Lange, *Correlation Techniques*)**

"Biddulph's Correlatogram 29" showing various sounds.See Plate 3.4 in Lange, F. H. *Correlation Techniques:Foundations and Applications of Correlation Analysis*. Iliffe, 1967.

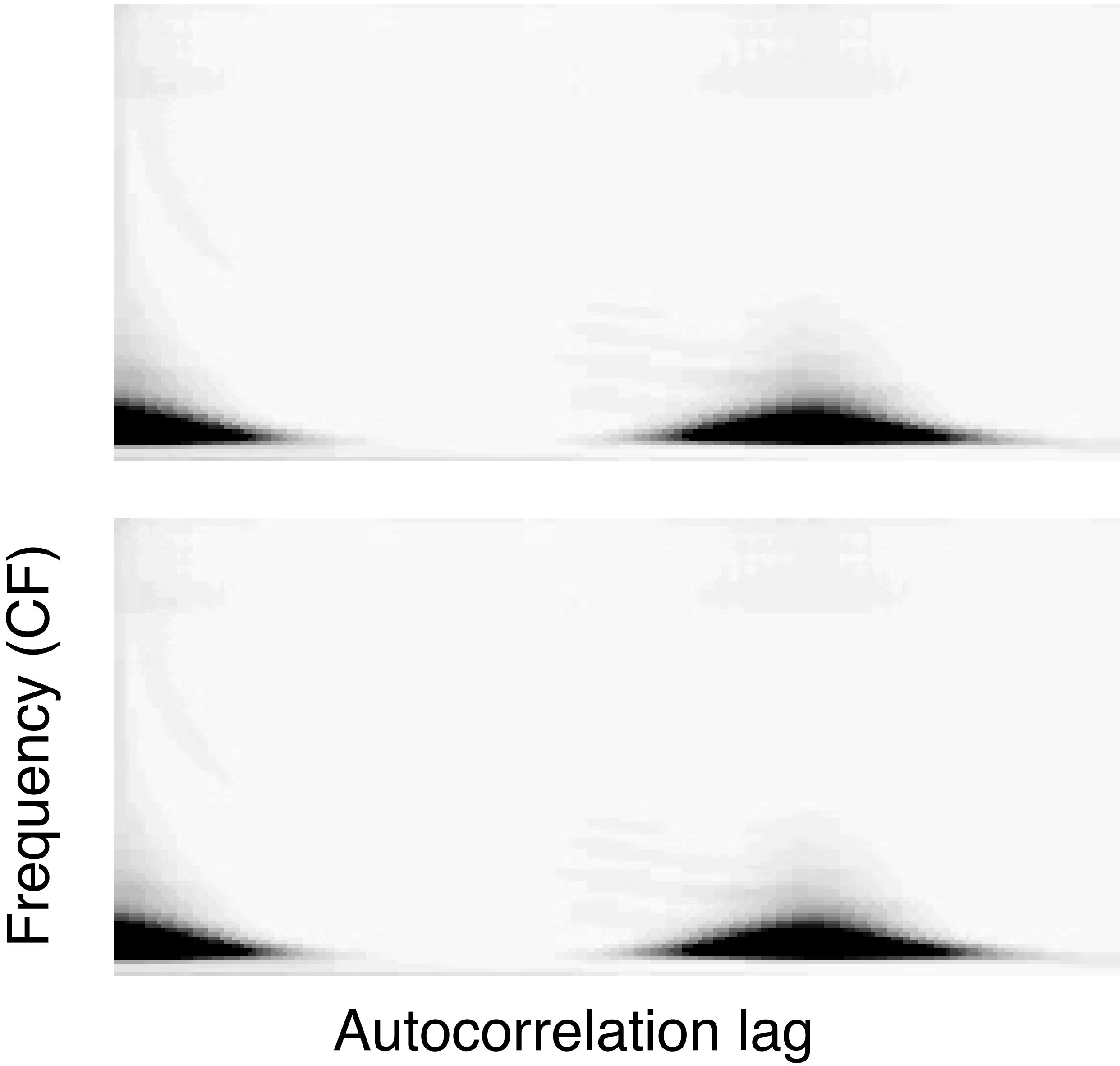
‘Big dogs can be dangerous.’





See Cariani, P. "Recurrent Timing Nets for F0-based Speaker Separation." Paper for Proceedings of Perspectives on Speech Separation, Montreal, October 30-November 2, 2003.

Correlograms: interval-place displays (Slaney & Lyon)



Correlograms

Images removed due to copyright restrictions.
See Figure 6.17 in Lyon, R. and S. Shamma. "Auditory Representations of Timbre and Pitch."
In *Auditory Computation*. Edited by R. R. Fay. New York, NY: Springer, 1996.

INTERVAL DISTRIBUTIONS AND OCTAVE SIMILARITY

This image is from the article Cariani, P. "Temporal Codes, Timing Nets, and Music Perception." *Journal of New Music Research* 30, no. 2 (2001): 107-135. DOI: 10.1076/jnmr.30.2.107.7115.
This journal is available online at http://www.ingentaconnect.com/content/routledg/jnmr/

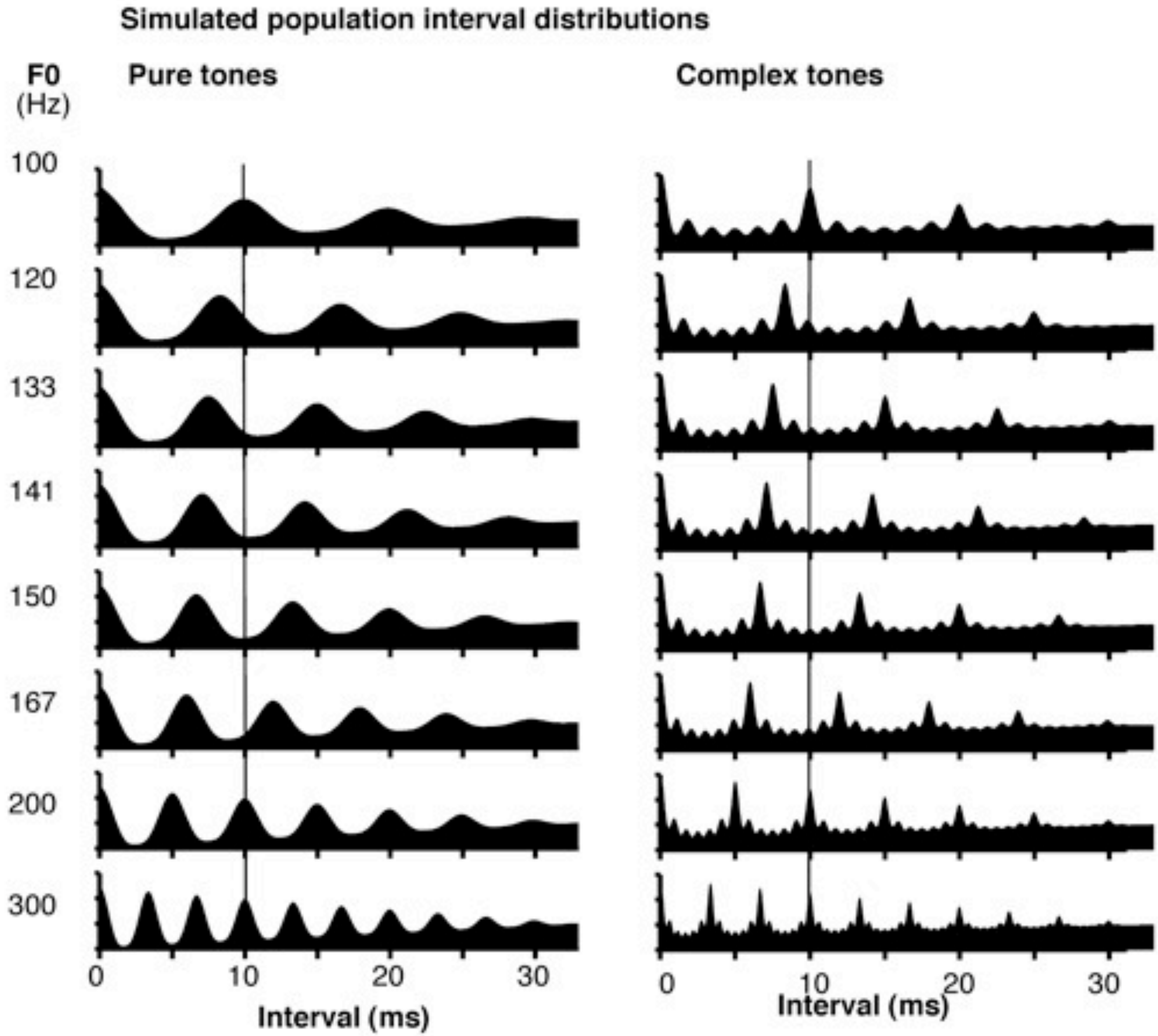
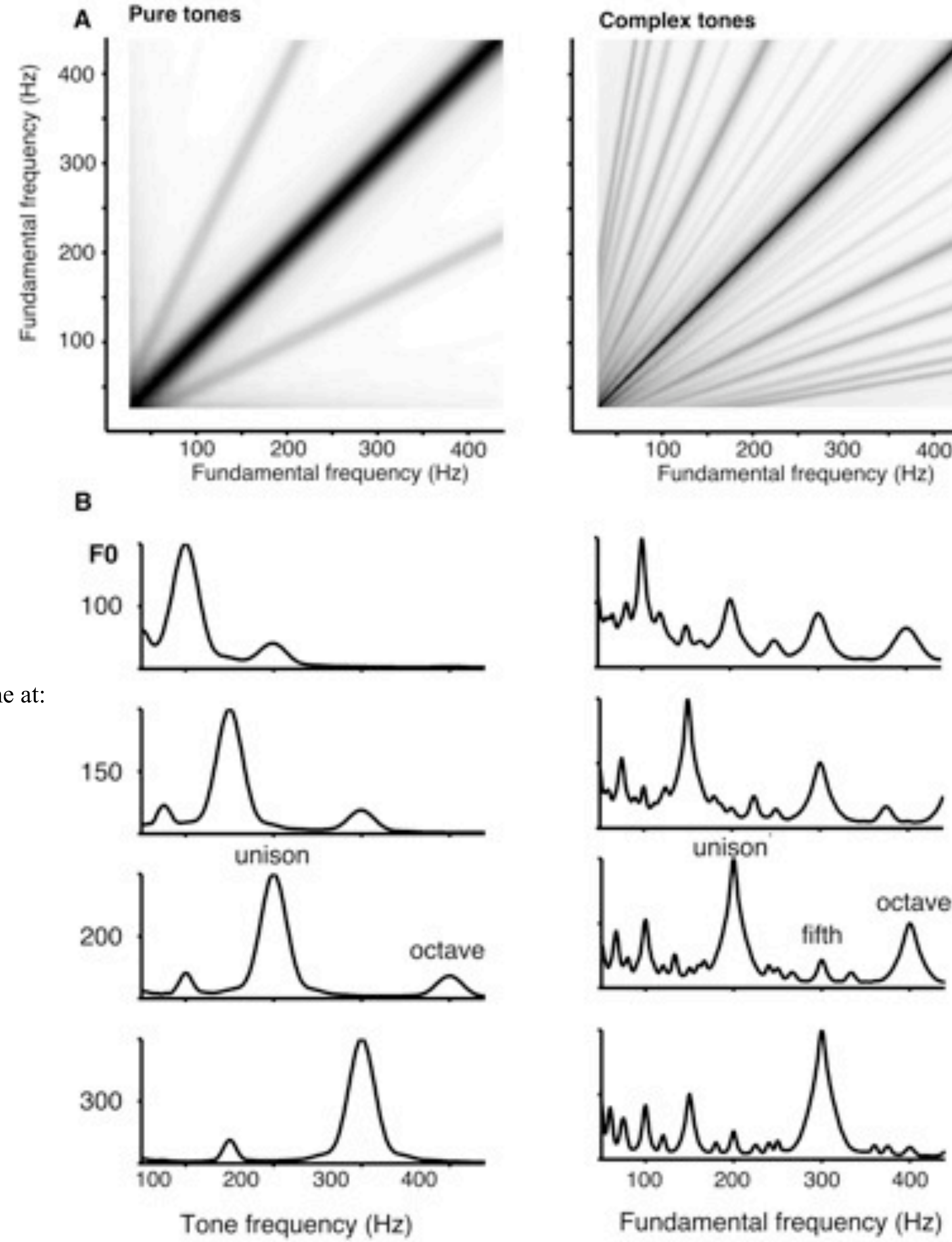


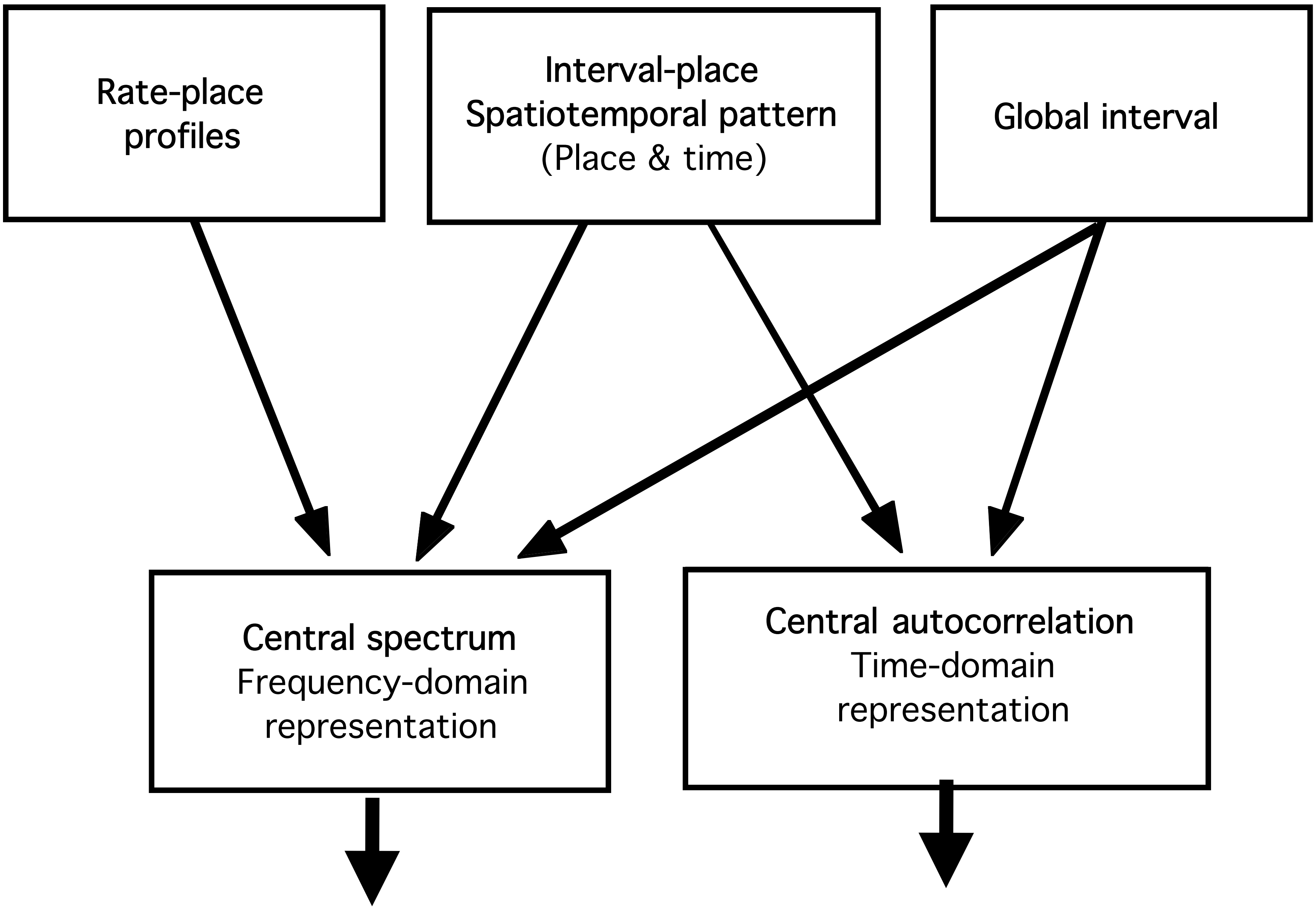
Figure 4. Similarities between population-interval representations associated with different fundamental frequencies. Simulated population-interval distributions for pure tones (left) and complex tones (right) consisting of harmonics 1-6.

Octave similarity

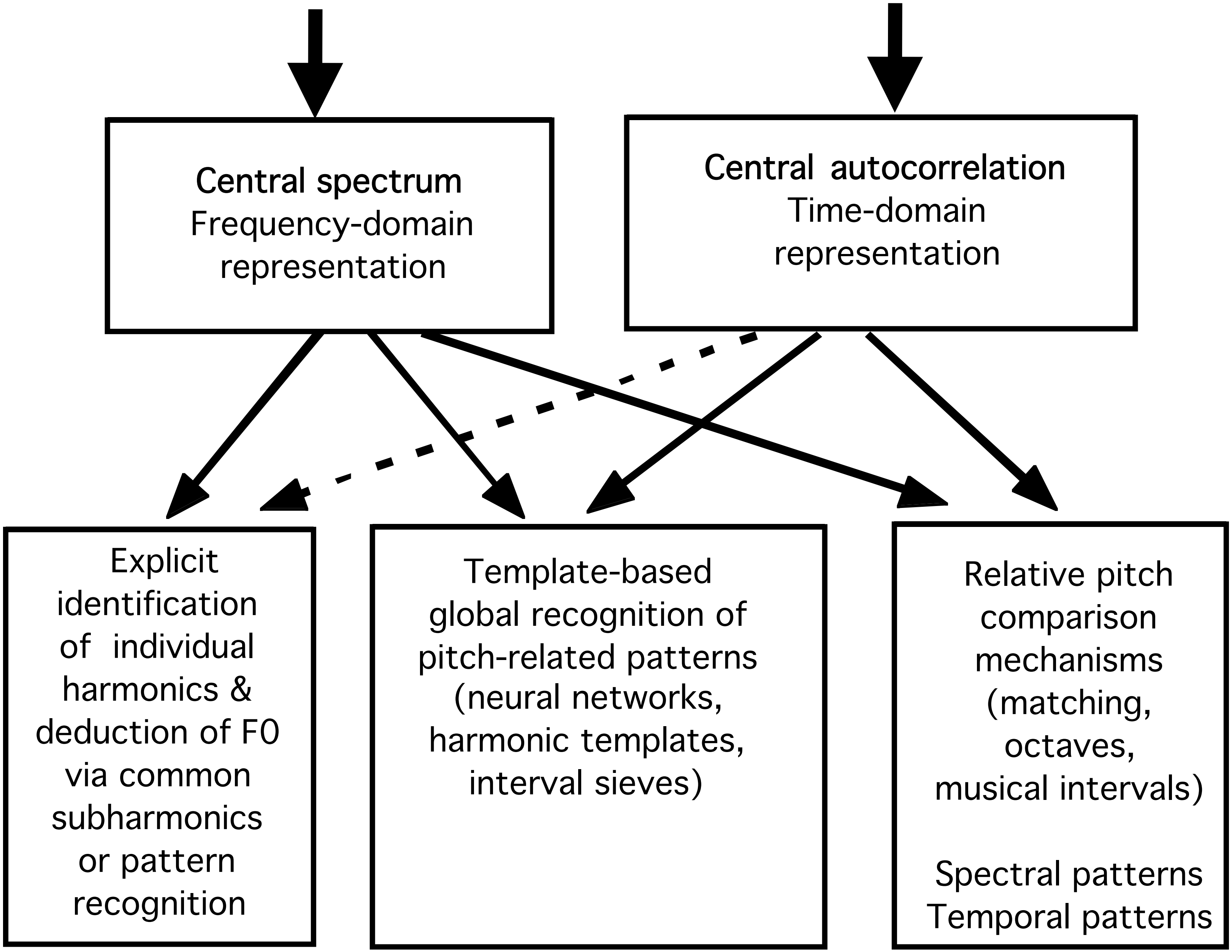


This image is from the article Cariani, P. "Temporal Codes, Timing Nets, and Music Perception." *Journal of New Music Research* 30, no. 2 (2001): 107-135.
DOI: 10.1076/jnmr.30.2.107.7115. This journal is available online at:
http://www.ingentaconnect.com/content/routledg/jnmr/

Physiological and functional representations



Different representations can support analogous strategies for pitch extraction, recognition, and comparison



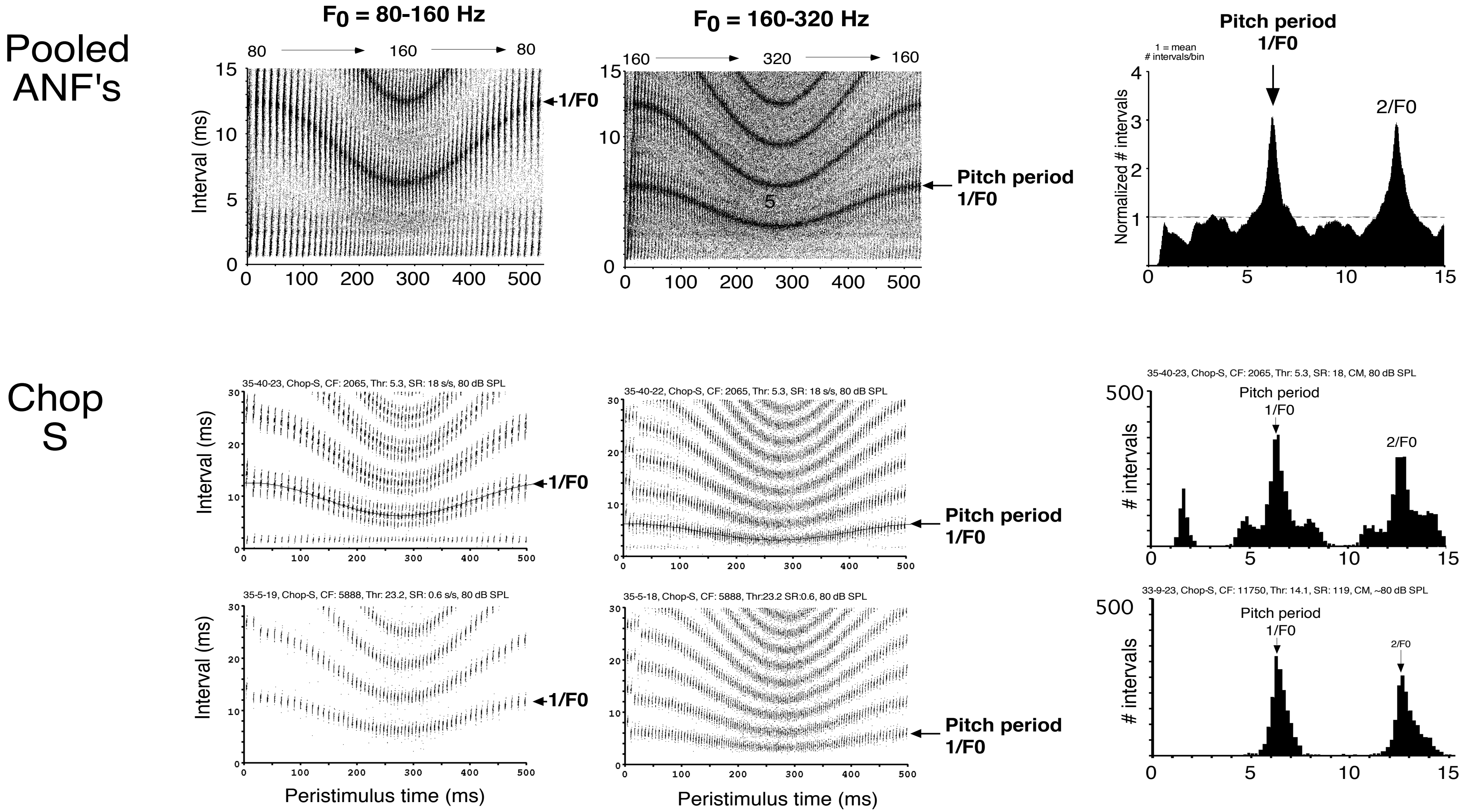
Cochlear nucleus I

Unipolar click trains
Variable F0

Pitch period ~1/F0

Unipolar click train
Fixed F0 (160 Hz)

Pitch ~160 Hz



Pitch height and pitch chroma

Images removed due to copyright restrictions. Figures 1, 2 and 7 in this paper.

Roger N. Shepard
Geometrical approximations to the
structure of musical pitch.
Psychological Review
89(4):305-322, 1982

Inharmonic complex tones (inharmonic AM tones)

Were used to falsify spectral models based on simple f-spacings and simple temporal models based on waveform envelopes.

Rules of thumb:

Low harmonics (perceptually resolved):

pitch is phase-insensitive

pitch follows fine structure of waveform, not envelope (pitch shifts, de Boer's rule)

High harmonics (unresolved)

pitch can be phase-sensitive (octave shifts)

Temporal theories - pros & cons

Make use of spike-timing properties of elements in early processing (to midbrain at least)

Interval-information is precise & robust & level-insensitive

No strong neurally-grounded theory of how this information is used

Unified model: account for pitches of perceptually-resolved & unresolved harmonics in an elegant way (dominant periodicity)

Explain well existence region for F0 (albeit with limits on max interval durations)

Do explain low pitches of unresolved harmonics

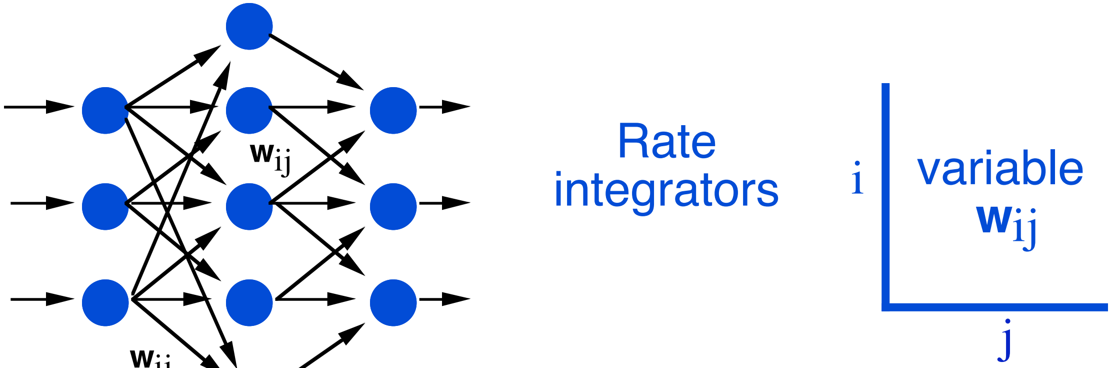
Interval analyzers require precise delays & short coincidence windows

Three networks

<u>**Connectionist networks**</u>

Purely spatial correlators

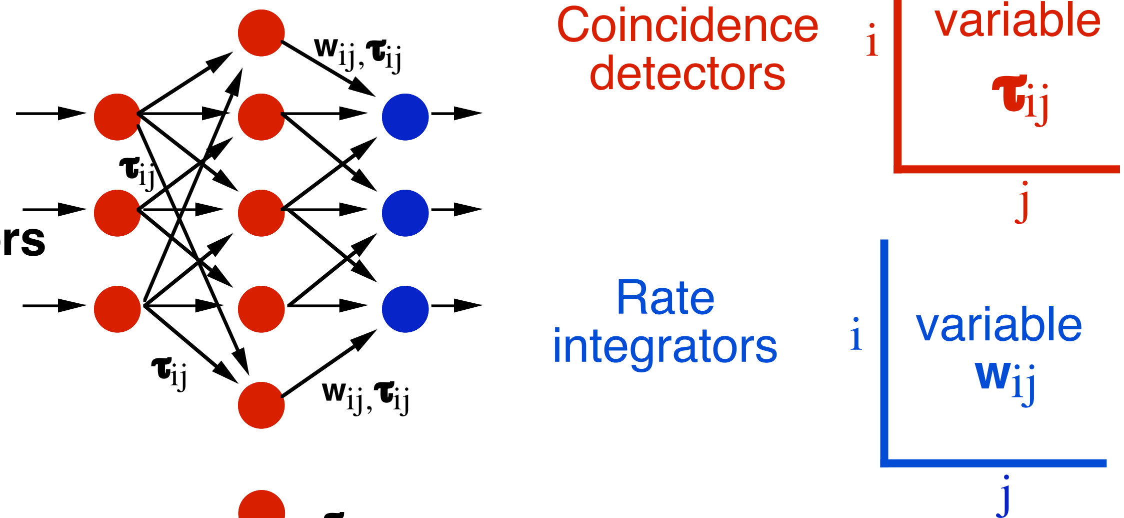
Place-Place mappings



<u>**Time-delay networks**</u>

Spatio-temporal correlators

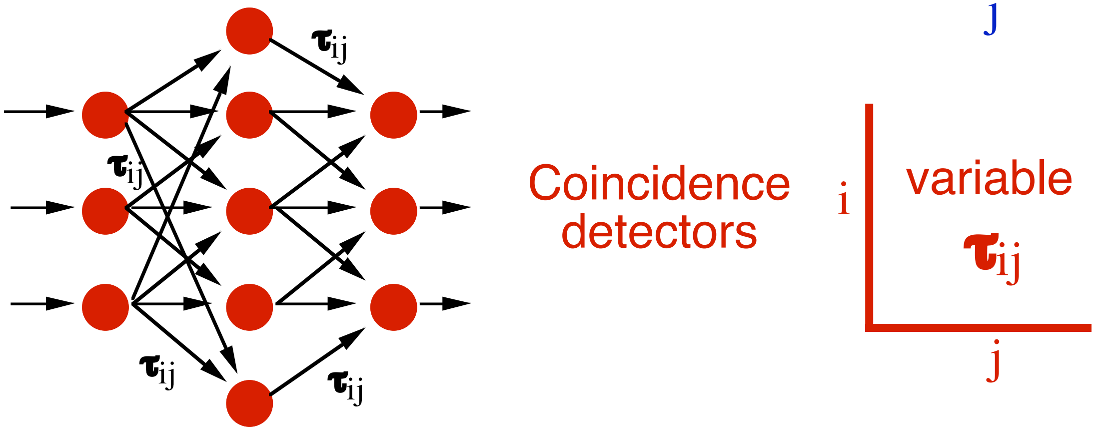
Time-Place mappings



<u>**Timing nets**</u>

Temporal correlators

Time-Time mappings



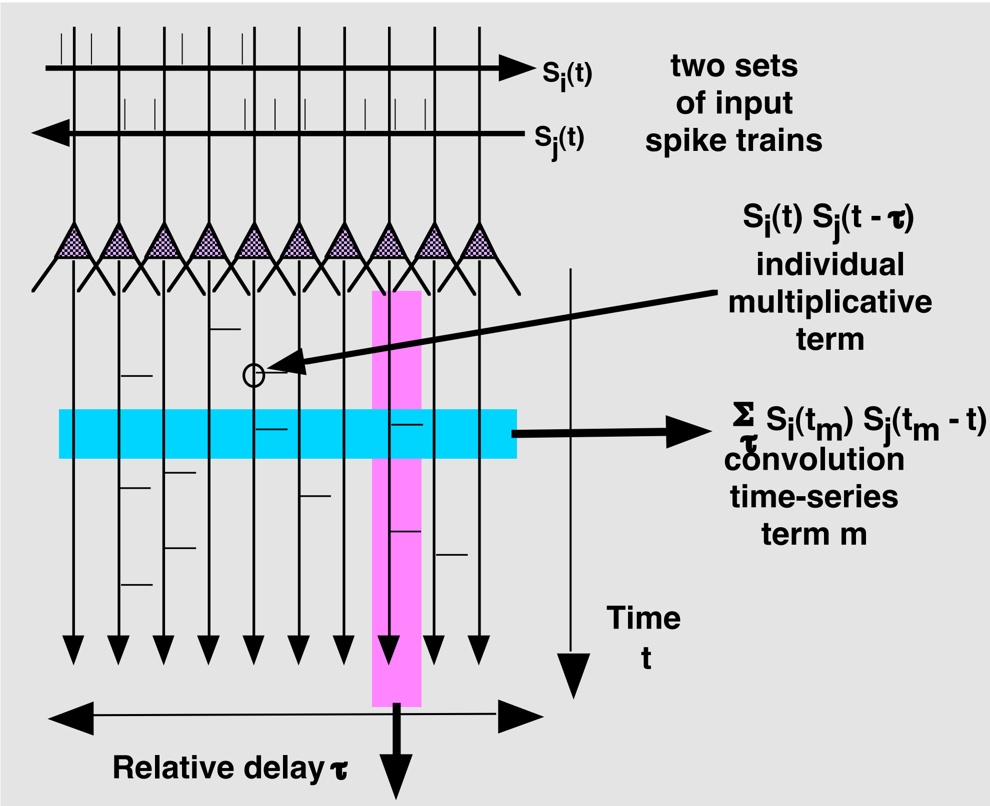
Neural timing nets

FEED-FORWARD TIMING NETS

- Temporal sieves
- Extract (embedded) similarities
- Multiply autocorrelations
- Pitch & timbre matching

RECURRENT TIMING NETS

- Build up pattern invariances
- Detect periodic patterns
- Separate auditory objects by F0
- Metric induction
- Time domain comb filters



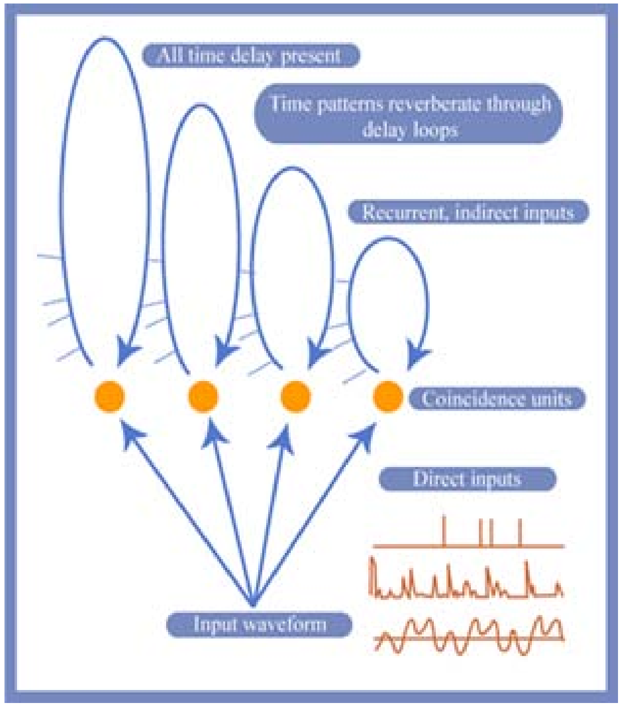
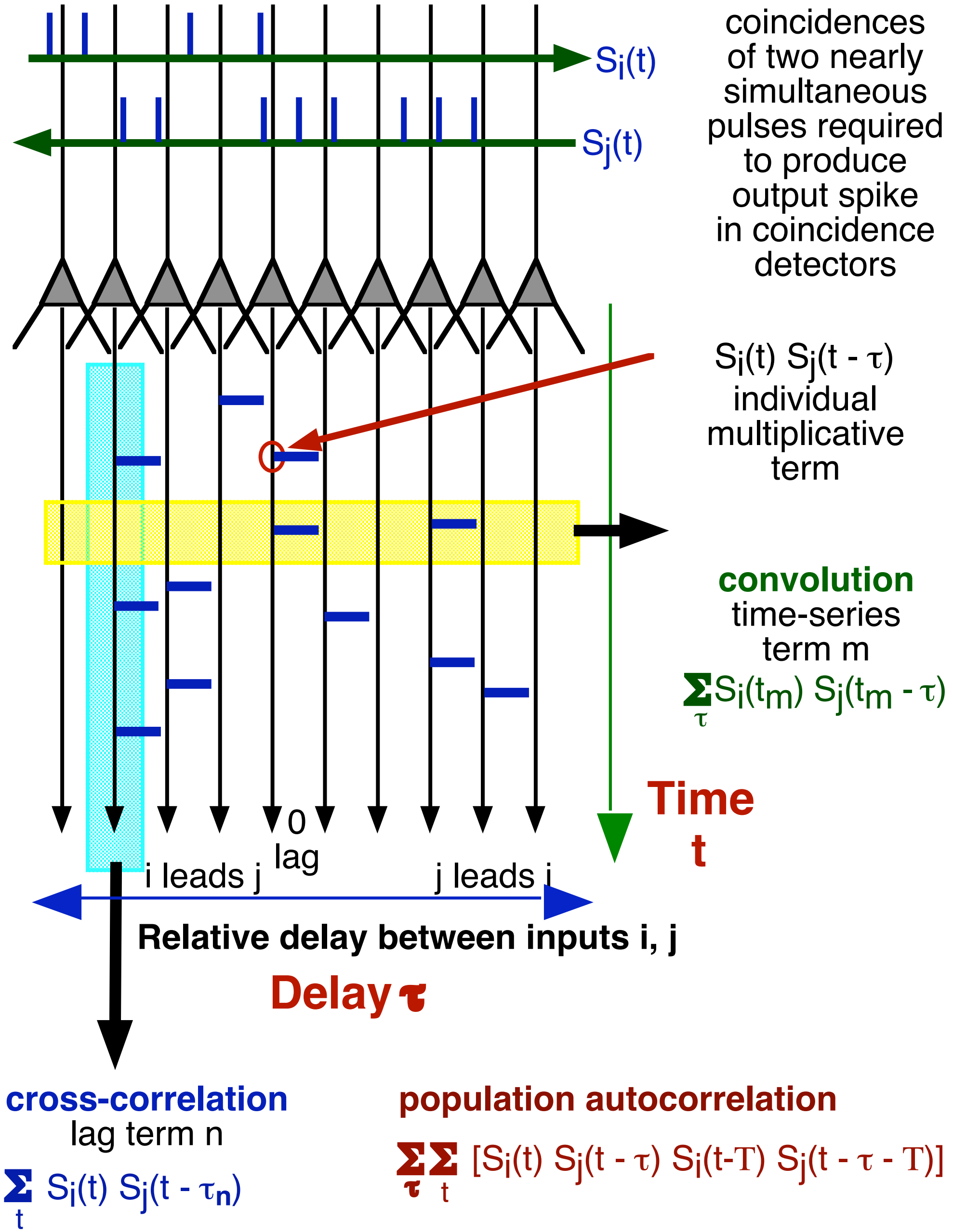


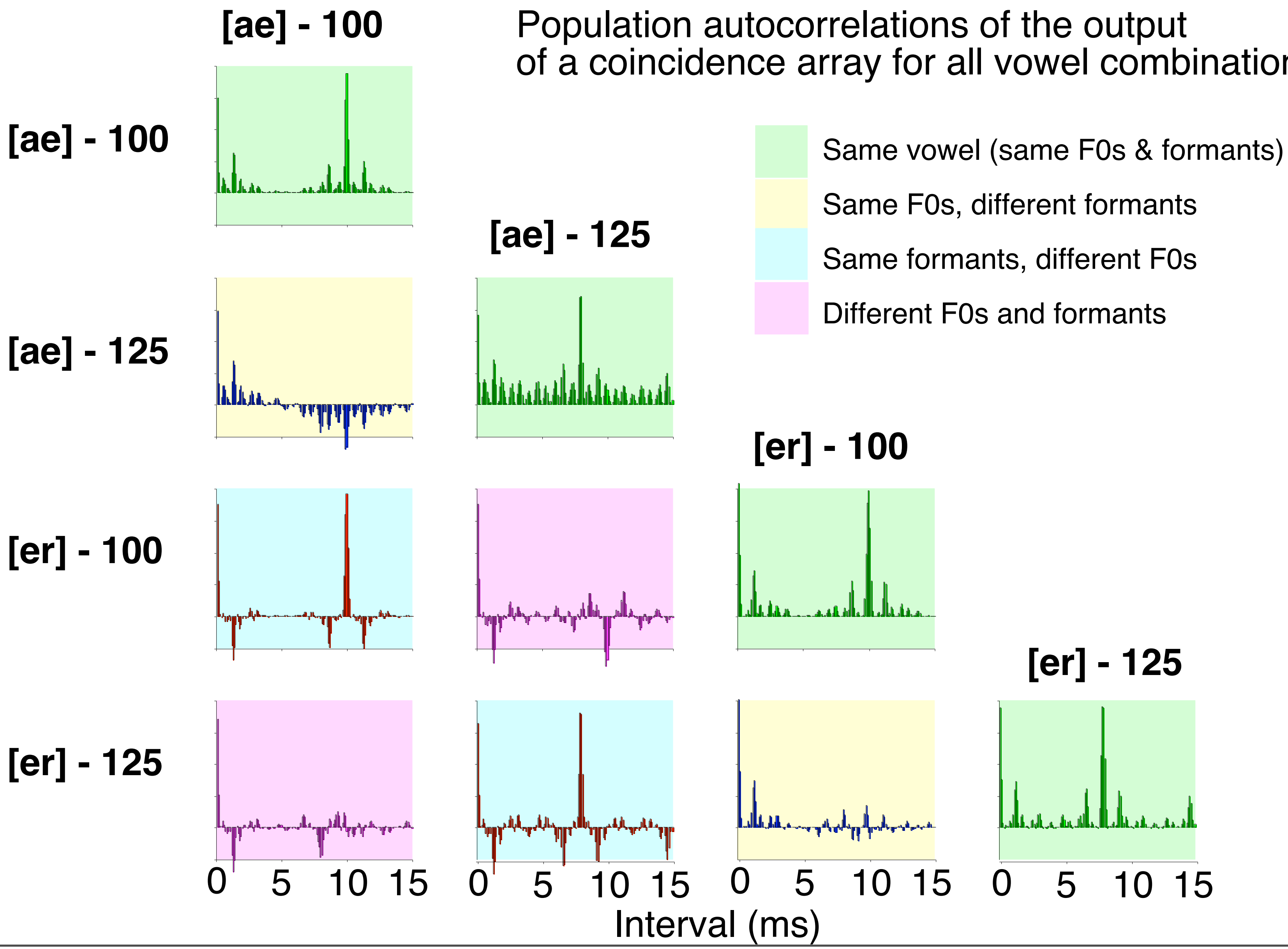
Figure by MIT OpenCourseWare.

Feedforward coincidence net



Common timbre

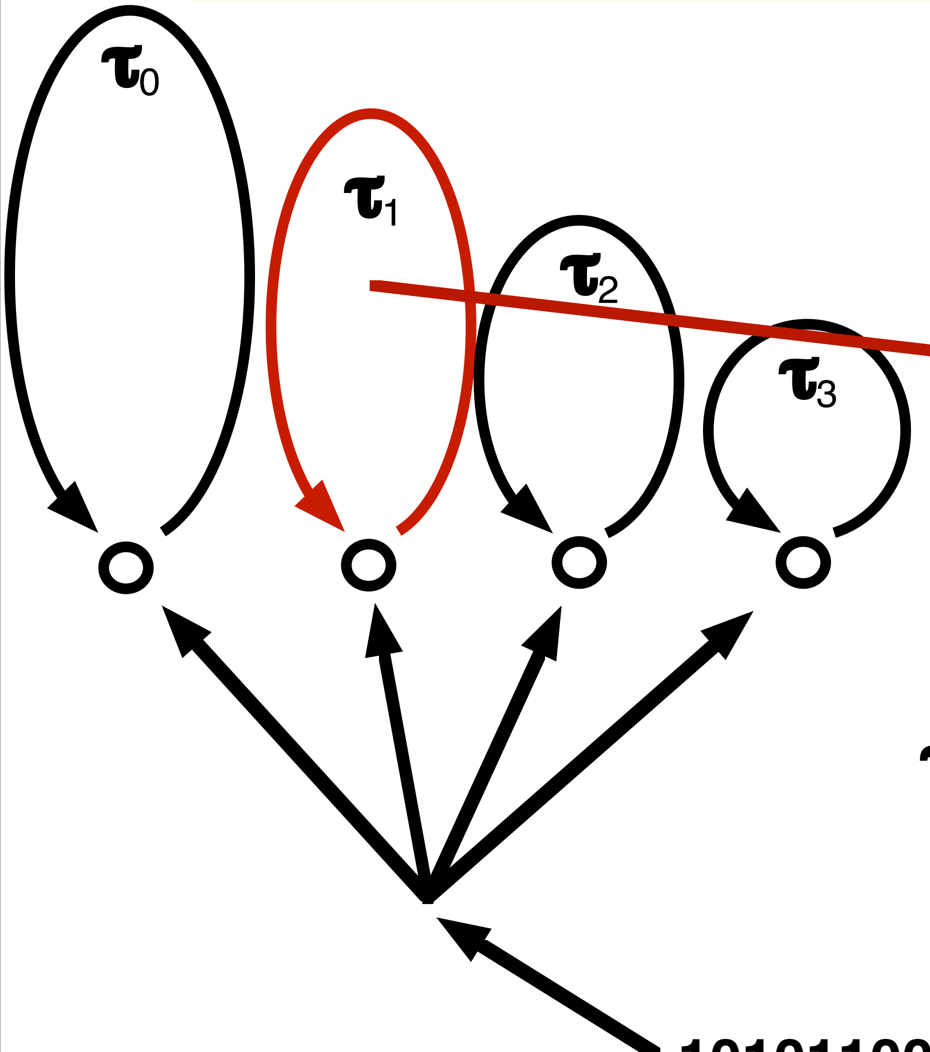
Population autocorrelations of the output of a coincidence array for all vowel combinations

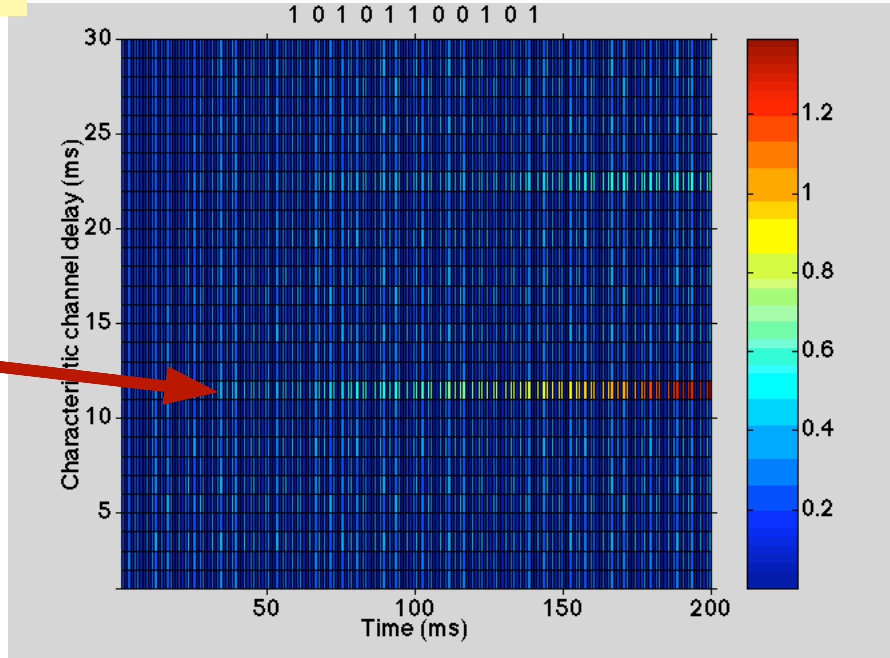


Detection of arbitrary periodic patterns

Periodic patterns invariably build up in delay loops whose recurrence times equals the period of the pattern and its multiples.

"cyclochronism" (Popov)

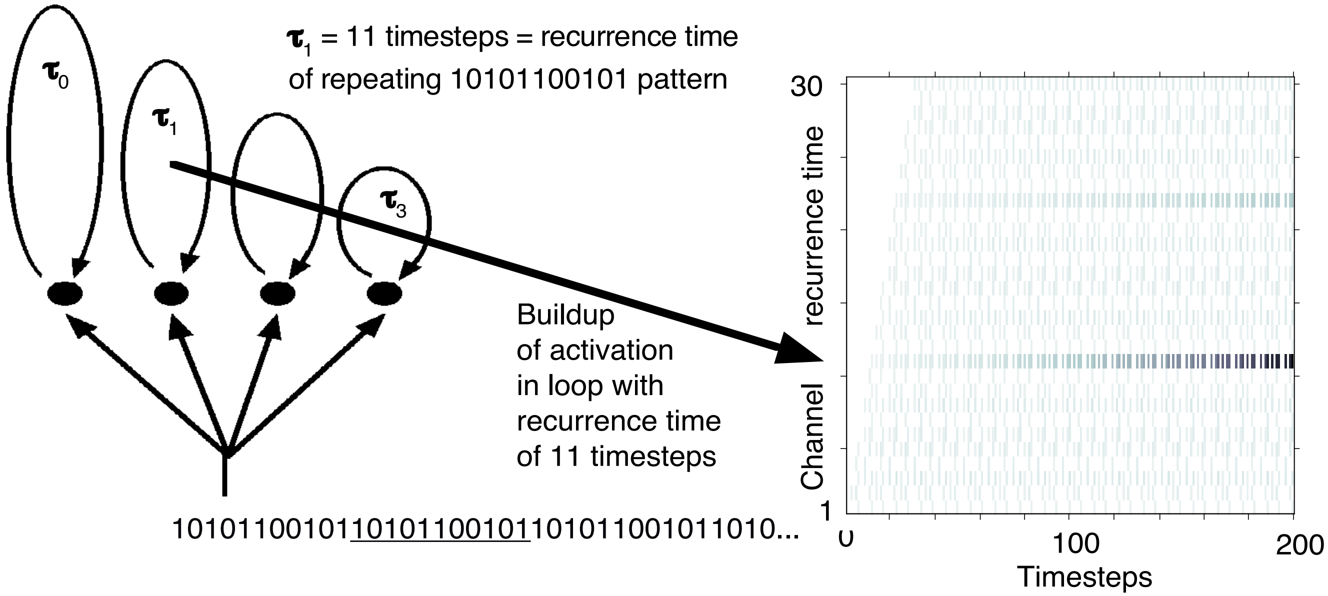




τ_1 = 11 ms = recurrence time of input pattern10101100101

Input pattern

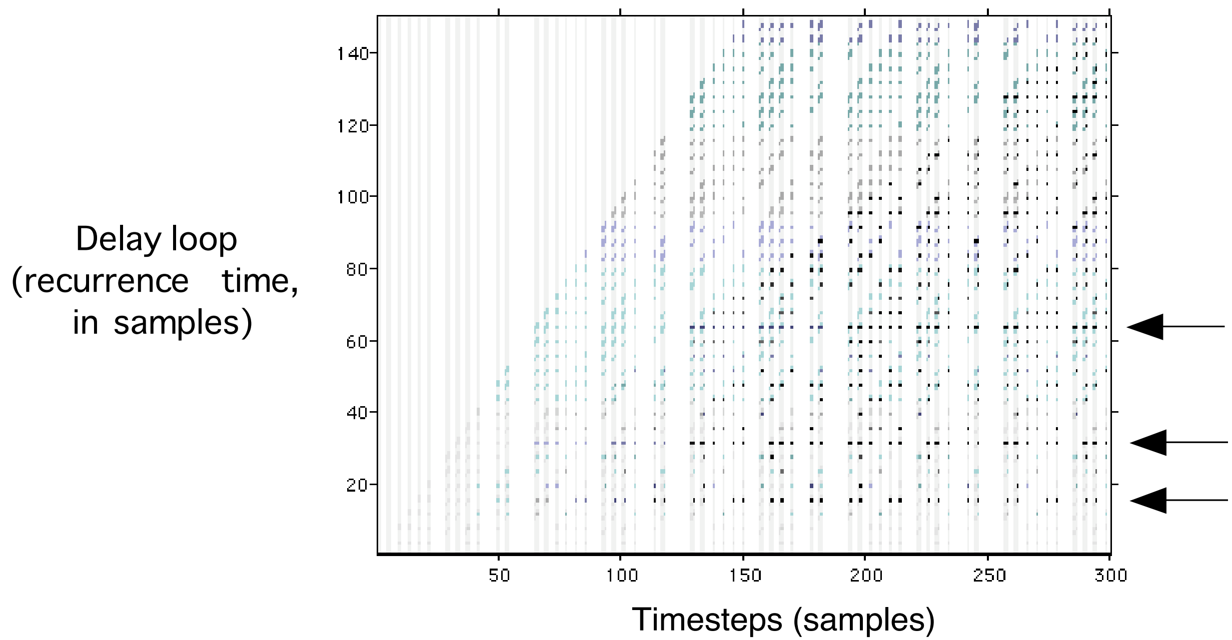
1010110010110101100101**10101100101 1010...**



Periodic pattern builds up

La Marseillaise rhythm

110011000100010001000100000011001100110001000000010011000000000...

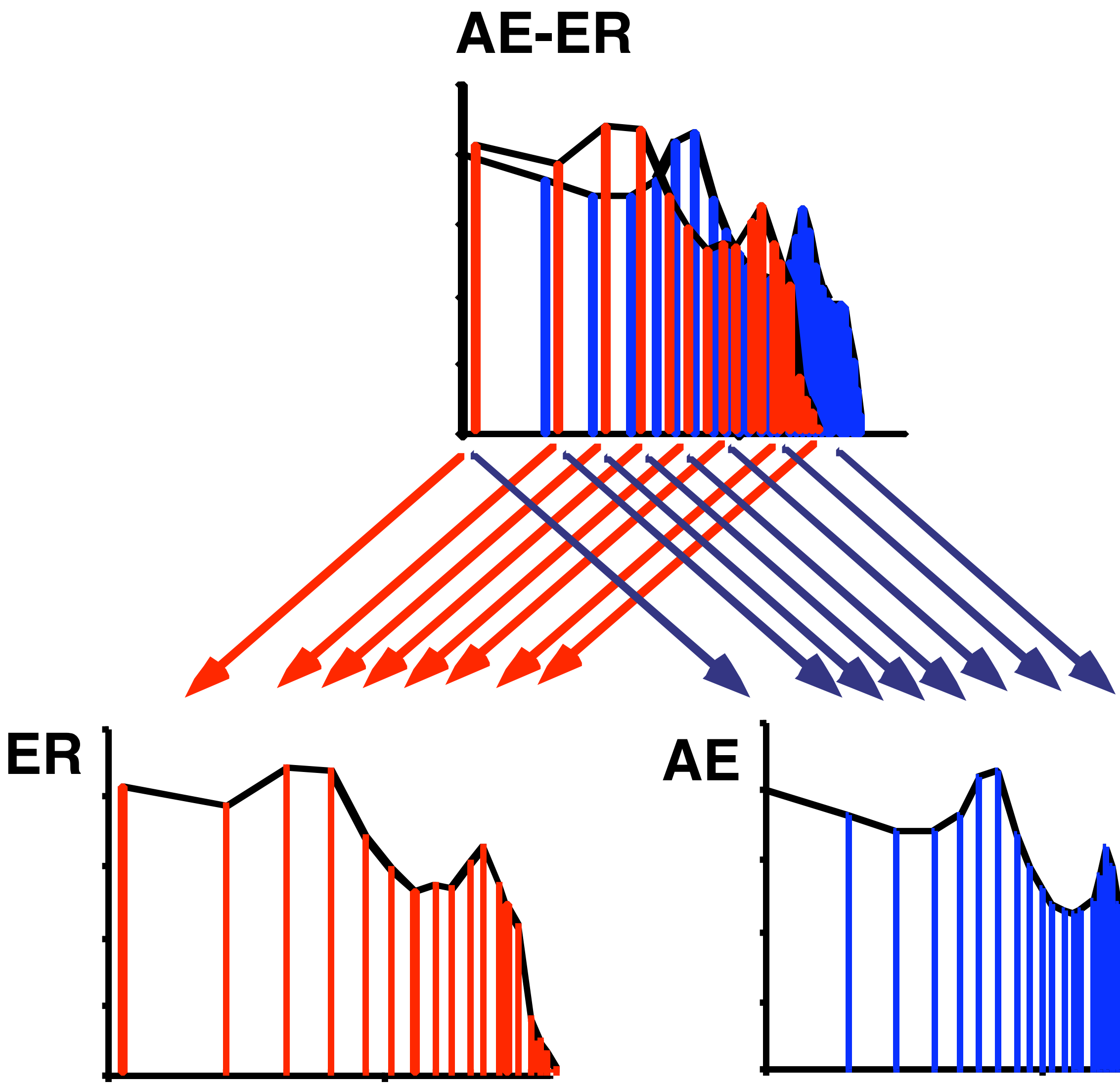


Rhythmic subpatterns

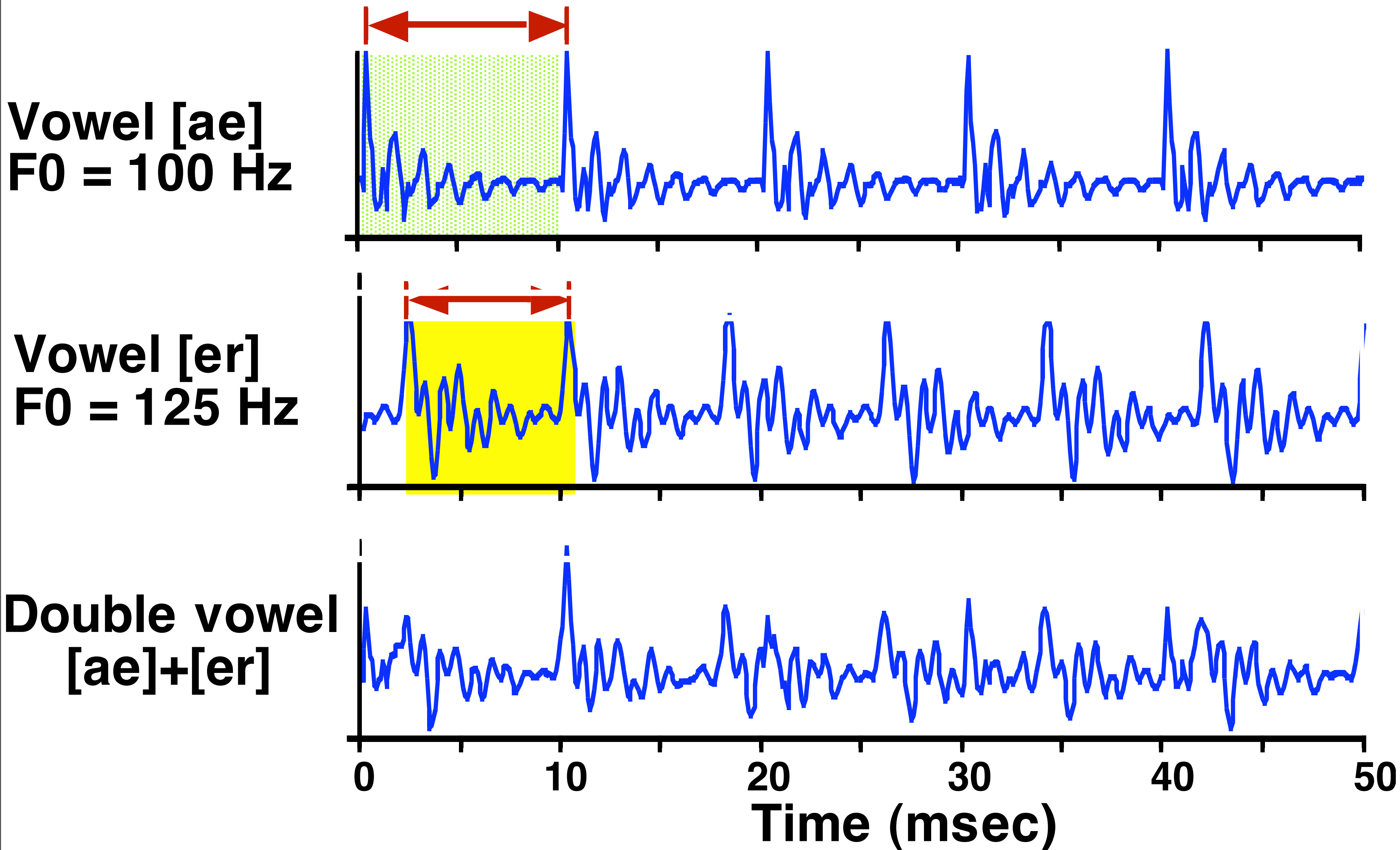
Traditional approach (Frequency domain)

Segregate frequency channels

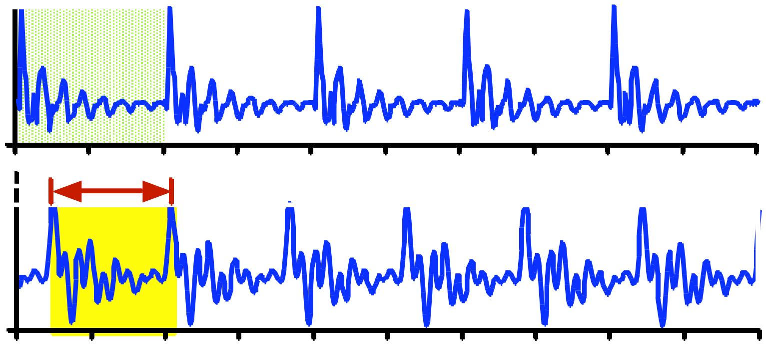
Assign channels to objects



Is a time-domain strategy possible?
Effect of different F0s in the time domain

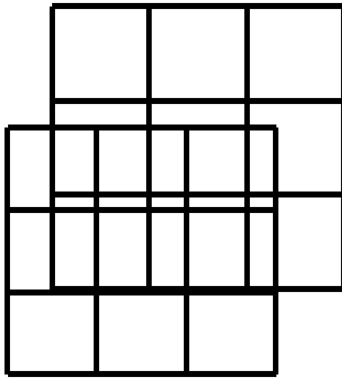


A general hypothesis re phase relations & grouping



1. Constant temporal relations fuse

2. Changing temporalrelations separate



3. The build-up mechanism is indifferent to particular stationary phase relations, but sensitive to changes in phase.

4. After stable objects are formed, they are analyzed via representations & mechanisms that are phase-insensitive (pitch, timbre, loudness)

Reading/assignment for next meeting

- **Pitch models & mechanisms**

Pitch classes
and
perceptual similarity

Build up harmonic
associations
from repeated
exposure to harmonic
complex tones

Harmonic similarity relations
are direct consequences
of the inherent structure
of interval codes

From cochlea to cortex

10,000k	Primary auditory cortex (Auditory forebrain)
	Auditory thalamus
500k	Inferior colliculus (Auditory midbrain)
	Lateral lemniscus
	Auditory brainstem
30k	Auditory nerve (VIII)
3k	Cochlea

Afferent Auditory Pathways

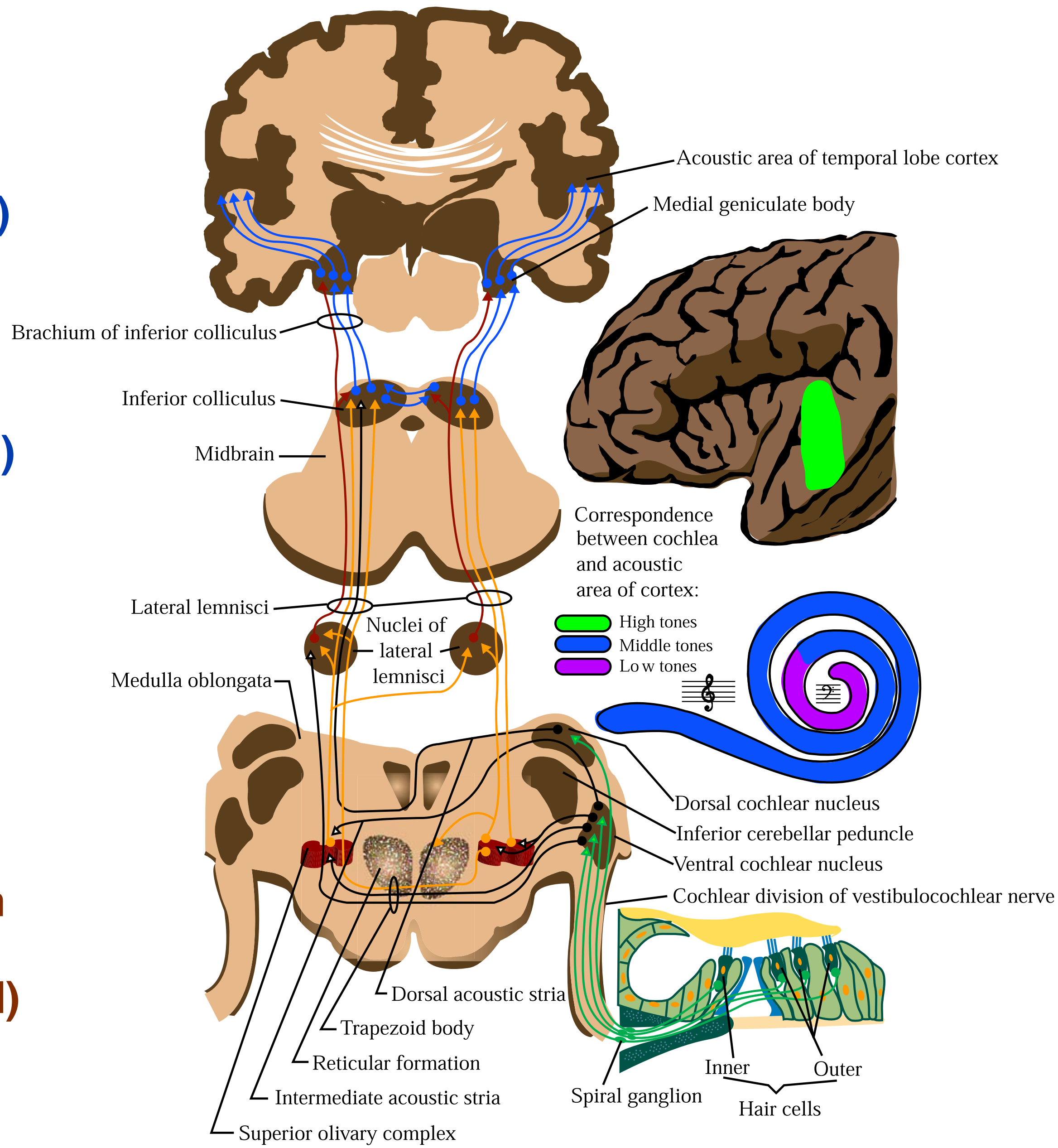


Figure by MIT OpenCourseWare.

Basic problems to be solved

• "Hyperacuity problem"

• Account for the precision of pitch discriminations given the relatively coarse tunings of auditory neurons (at all levels), especially lower-frequency ones (BFs < 2 kHz)

• "Dynamic range problem"

• Account for the ability of listeners to discriminate small fractional changes ($\Delta I/I$) in intensity over a large dynamic range, and especially at high SPLs, where the vast majority of firing rates are saturated.

• "Level-invariance problem"

• Account for the invariance (and precision) of auditory percepts over large dynamic ranges given the profound changes in neural response patterns that occur over those ranges (rate saturation, rate non-montonicities).

•Pitch equivalence

•Account for the ability to precisely match pitches of pure and complex tones (pitch equivalence, metamery) given differences in spectra and under conditions where stimulus intensities are roved 20 dB or more

•Relative nature of pitch & transpositional invariance

•Account for the ability to precisely match pitches an octave apart (and/or to recognize patterns of pitch sequences) in the absence of an ability to identify absolute frequencies/ periodicities. Account for ability to recognize transposed melodies as similar.

Some generalities about the auditory system

• Rough cochleotopy is found at all levels, but not necessarily in all pops

• Orderly tonotopic spatial maps exist only at low tone levels, near neural thresholds

As one ascends the afferent pathway:

• Numbers of neurons at each level increases (usually 2x or more)

• Fine timing information exists in great superabundance in lower stations, but becomes successively sparser

• Firing rates (spontaneous & driven) decline (usually 2x or more)

• Inhibition increases; % nonmontonic rate-level functions increase

• Greater proportion of phasic responders, onset & offset responses

• Diversity and complexity of response increases

• History-dependence and contextual effects increase

• Some modulation tuning that suc. declines in periodicity

Typical BMFs: AN: 200-300 Hz; IC: 50-100 Hz; Ctx (< 16 Hz)

• No clear "pitch detectors" (Schwarz & Tomlinson, 1991);

–until, perhaps, recently (Bendor & Wang, 2005)

• No narrow (BW < 0.3 octaves) "frequency channels" for BFs < 2 kHz (thus far)

Brainstem stations involved in localization of sounds

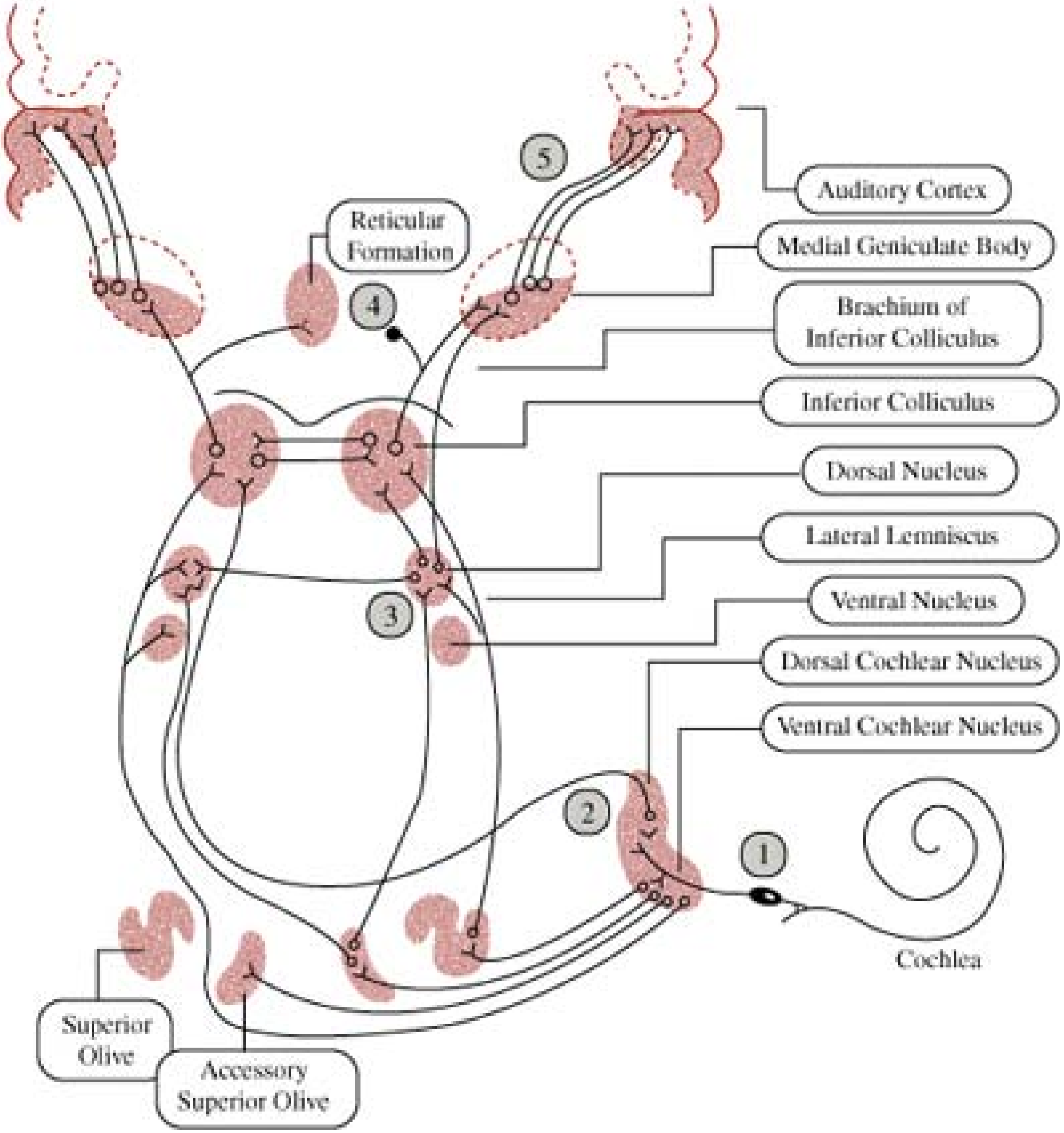


Figure by MIT OpenCourseWare.

Three cochlear nuclei :

AVCN

PVCN

DCN

Bifurcation of auditory nerve

Innervation of 3 major cochleotopically-organized

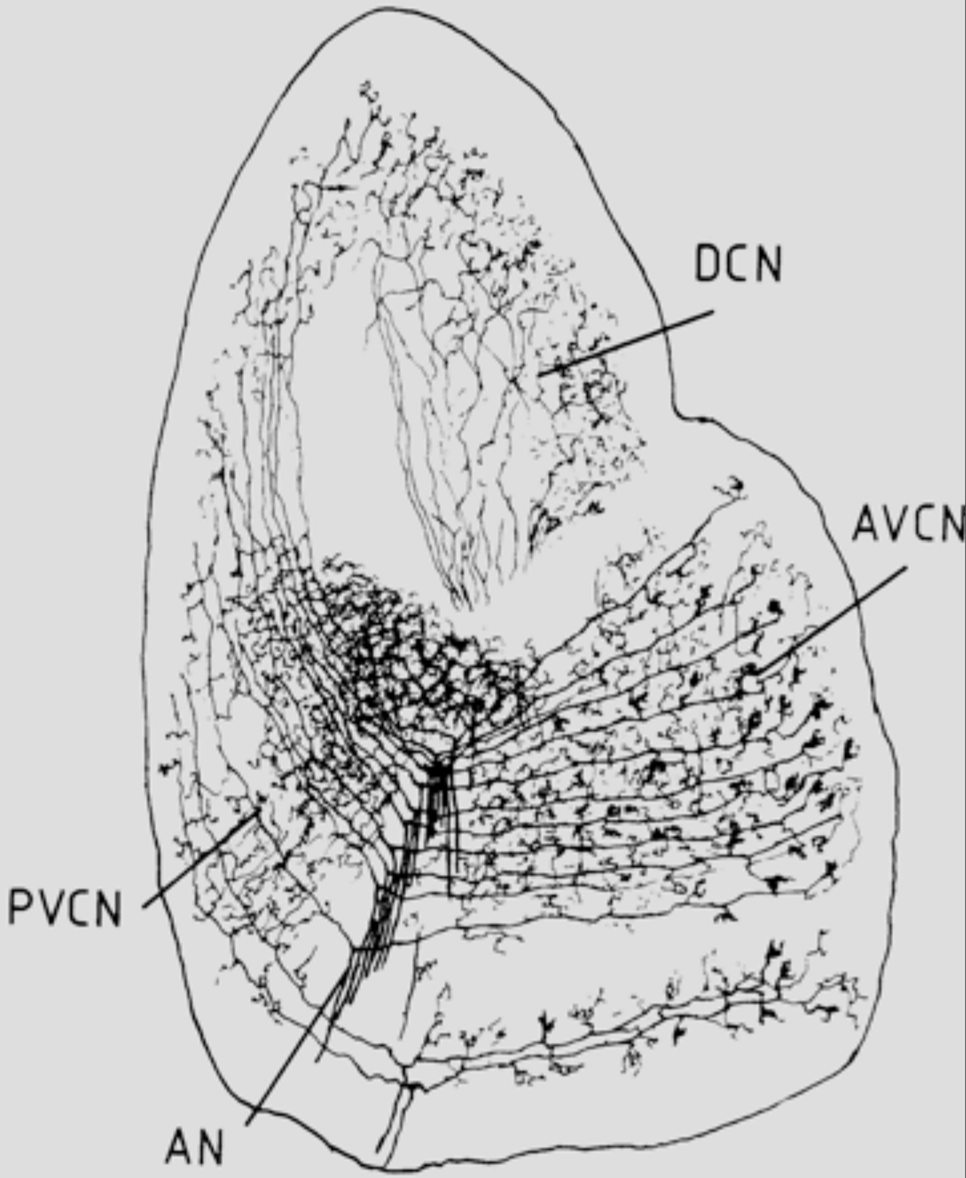


FIG. 2. Parasagittal section through cochlear nucleus of 4-day-old cat stained by Golgi method. Auditory nerve fibers, AN, bifurcate to yield ascending branch to AVCN and descending branch to PVCN and DCN. Ascending branch terminates in AVCN with large end bulbs of Held. [From Lorente de Nó (279).]

Source: public domain

Cochlear nuclei : first station in the auditory CNS

Images removed due to copyright restrictions.

Figures 1, 3 and 13 in Irvine, D. R. F. *The Auditory Brainstem*. New York, NY: Springer, 1986. ISBN: 9783540162995.

Cochlear nuclei : 3 major divisions (AVCN, PVCN, DCN)

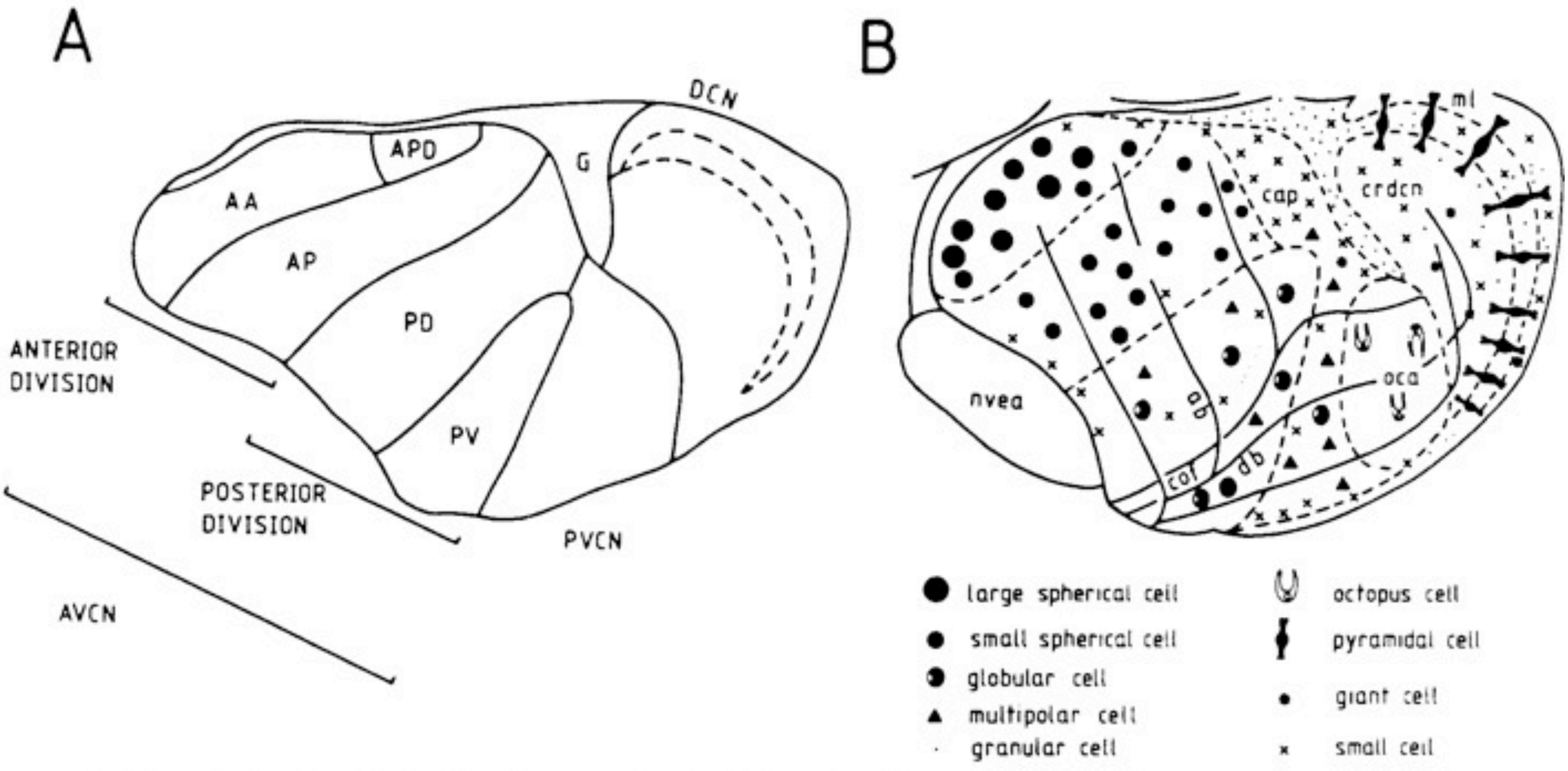
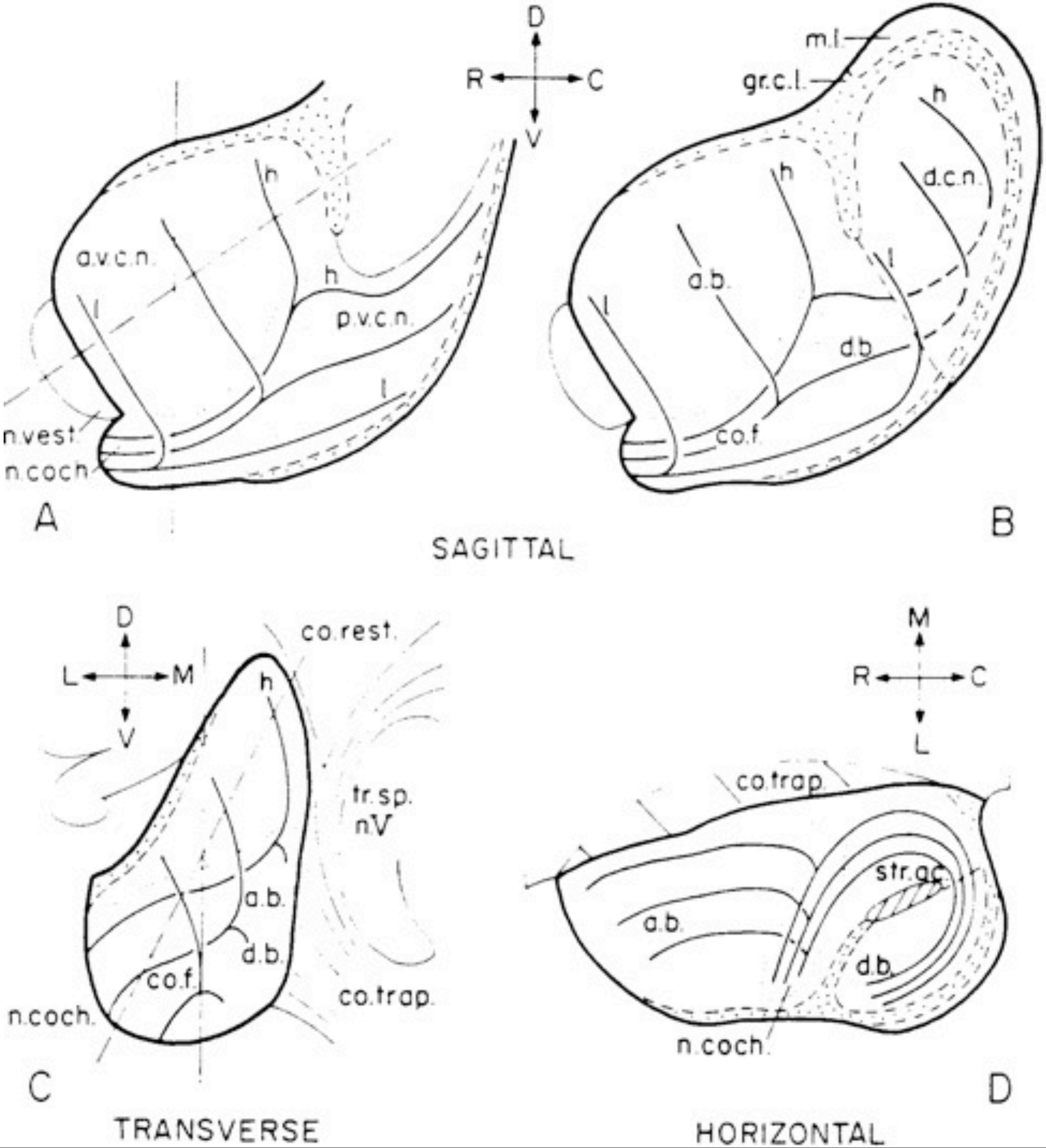


FIG. 4. Schematic drawings of sagittal sections comparing *A*: divisions of cochlear nucleus identified by Brawer et al. (59), and *B*: divisions of cochlear nucleus of Osen (343). In *A*: AVCN, anterior ventral cochlear nucleus, which is divided into anterior division comprising AA, anterior; AP, posterior; and APD, dorsal parts; and a posterior division comprising PD, dorsal; and PV, ventral parts; PVCN, posterior ventral cochlear nucleus; G, granule cell layer; DCN, dorsal cochlear nucleus; *dotted line* represents position of fusiform (pyramidal) cell layer. In *B*: nvea, vestibular nerve; oca, octopus cell area; cap, small cell cap; crdcn, central region of dorsal cochlear nucleus; ab, ascending cochlear branch; cof, cochlear nerve fiber; db, descending cochlear branch; ml, molecular layer.

Cochlear nuclei : 3 major divisions
(AVCN, PVCN, DCN)

FIG. 3. Cochlear nuclei in three dimensions showing bifurcation pattern in A, B: sagittal; C: transverse; and D: horizontal planes. Branches that represent high and low frequencies are labeled h and l, respectively. a.b., Ascending cochlear branch; a.v.c.n., anteroventral cochlear nucleus; C, caudal; co.f., cochlear nerve fiber; co. rest., restiform body; co. trap., trapezoid body; D, dorsal; d.b., descending cochlear branch; d.c.n., dorsal cochlear nucleus; gr.c.l., granular cell layer; m.l., molecular layer; L, lateral; M, medial; n.coch., cochlear nerve; n.vest., vestibular nerve; p.v.c.n., posteroventral cochlear nucleus; R, rostral; str.ac., acoustic striae; tr.sp.n.V., spinal fifth tract; V, ventral. [From Osen (346).]



Cochlear nuclei :

Types of responses seen

(to tone bursts at CF):

Primary-like (AVCN)
Primary-like w. notch (AVCN)
Phase-locked (PVCN)
Chopper (PVCN)
Pauser (DCN)
Build-up (DCN)
Onset (PVCN)

Image removed due to copyright restrictions.
See Fig. 2.18 in Romand, R., and P. Avan.
"Anatomical and Functional Aspects of the Cochlear Nucleus."
The Central Auditory System. Edited by G. Ehret and R. Romand.
New York, NY: Oxford University Press, 1997.
[Preview this image in Google Books.]

Most are linked to a particular neuronal morphological type

Figure 2.18. Types of responses that can be obtained in the cochlear nucleus to a 25 ms tone-burst stimulation. In the auditory nerve (AN), only a single response type exists for tone frequencies higher than 1 kHz, the so-called primary responses or primary-like responses. In the anteroventral cochlear nucleus (AVCN), mainly the following responses are obtained: A: primary-like; B: phase-locked; C: sustained chopper; D: onset chopper (O_C); E: onset. In the posteroventral (PVCN) and dorsal cochlear nucleus (DCN), the following responses are obtained: F: pauser; G: buildup; H: sustained chopper; I: onset sustained (O_L); J: onset transient (O_I).

(-) indicate main regions

Cochlear nucleus units: responses to tone bursts

Image removed due to copyright restrictions.
See Fig. 2.18 in Romand, R., and P. Avan.
"Anatomical and Functional Aspects of the Cochlear Nucleus."
The Central Auditory System. Edited by G. Ehret and R. Romand.
New York, NY: Oxford University Press, 1997.
[Preview this image in Google Books.]

Note: (C) & (H) "chopping" occurs for f > 1.5 kHz; phase-locking to fine structure for f < 1.5 kHz

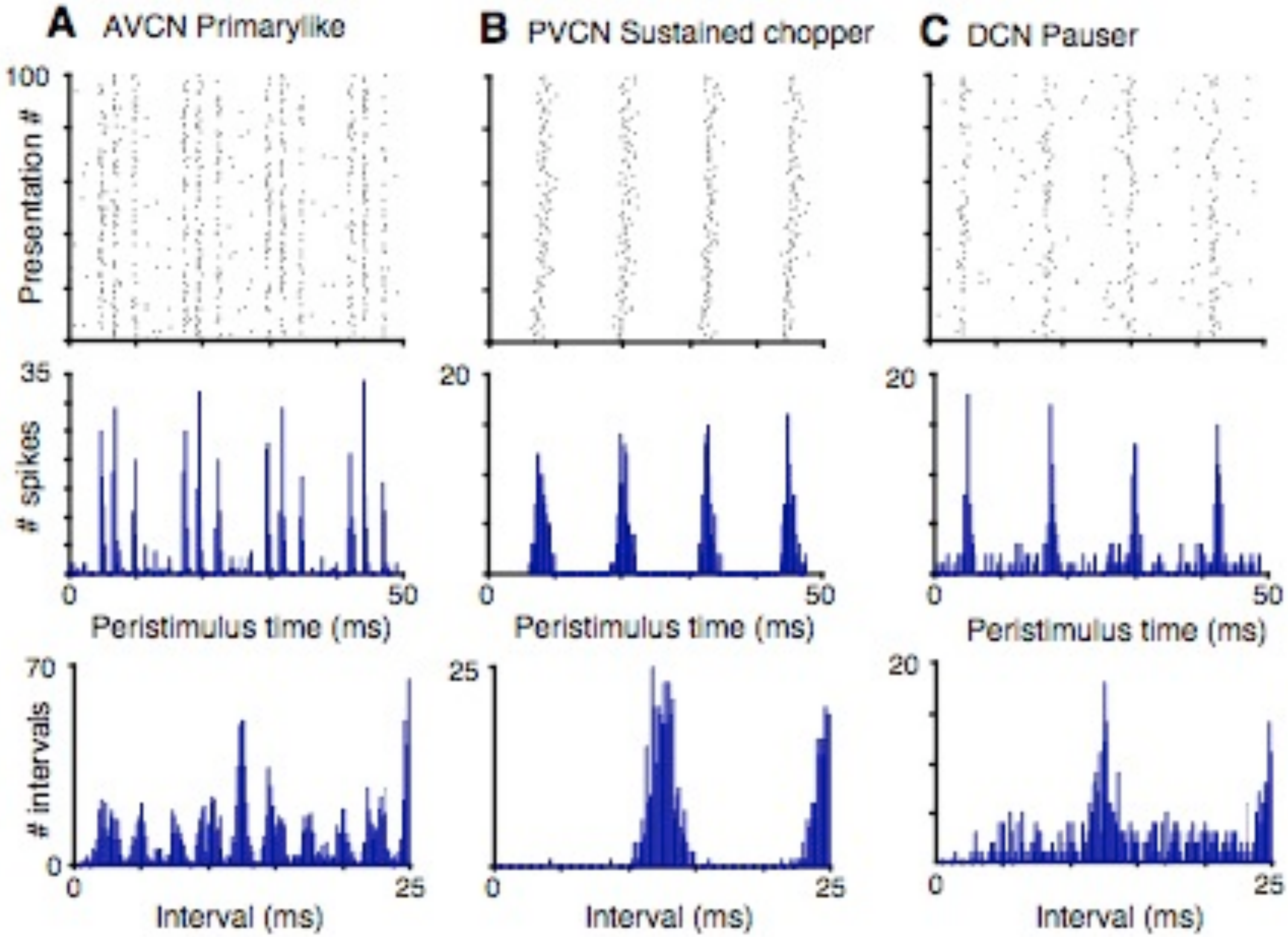
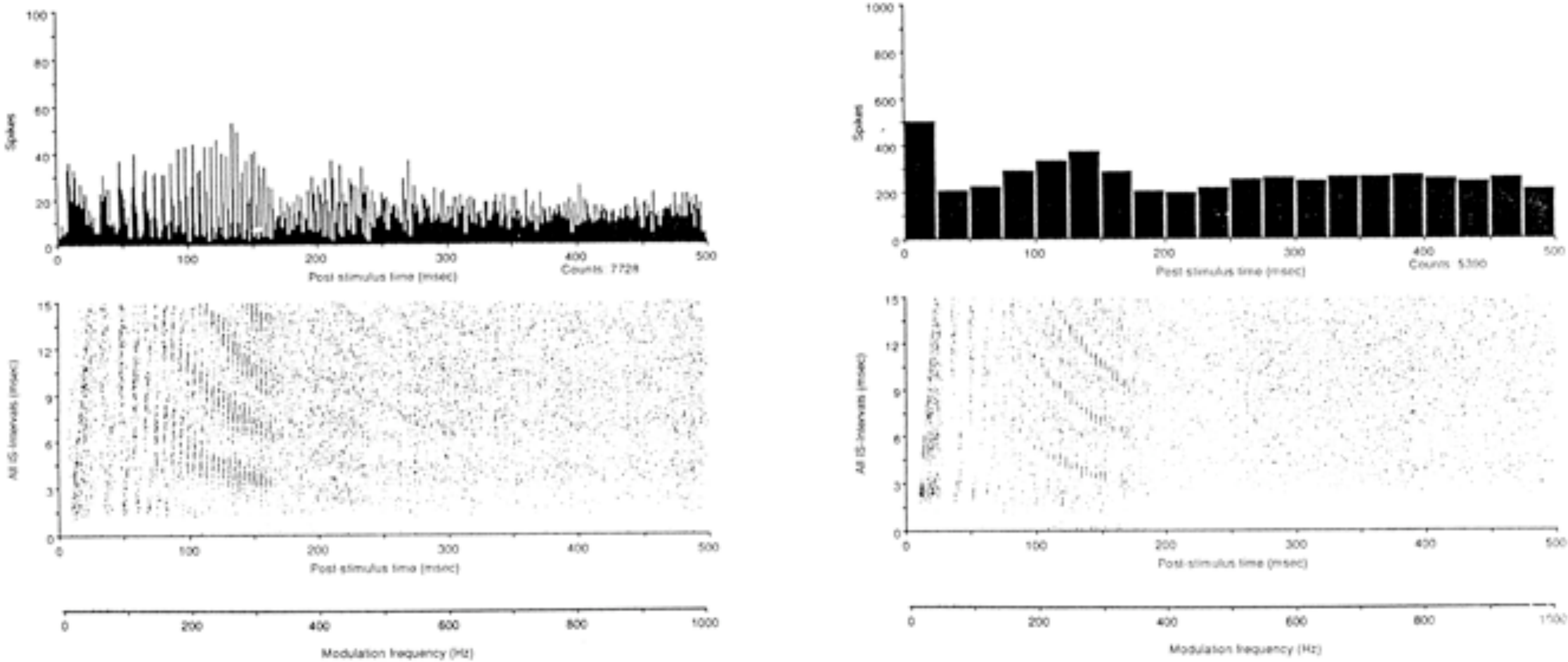
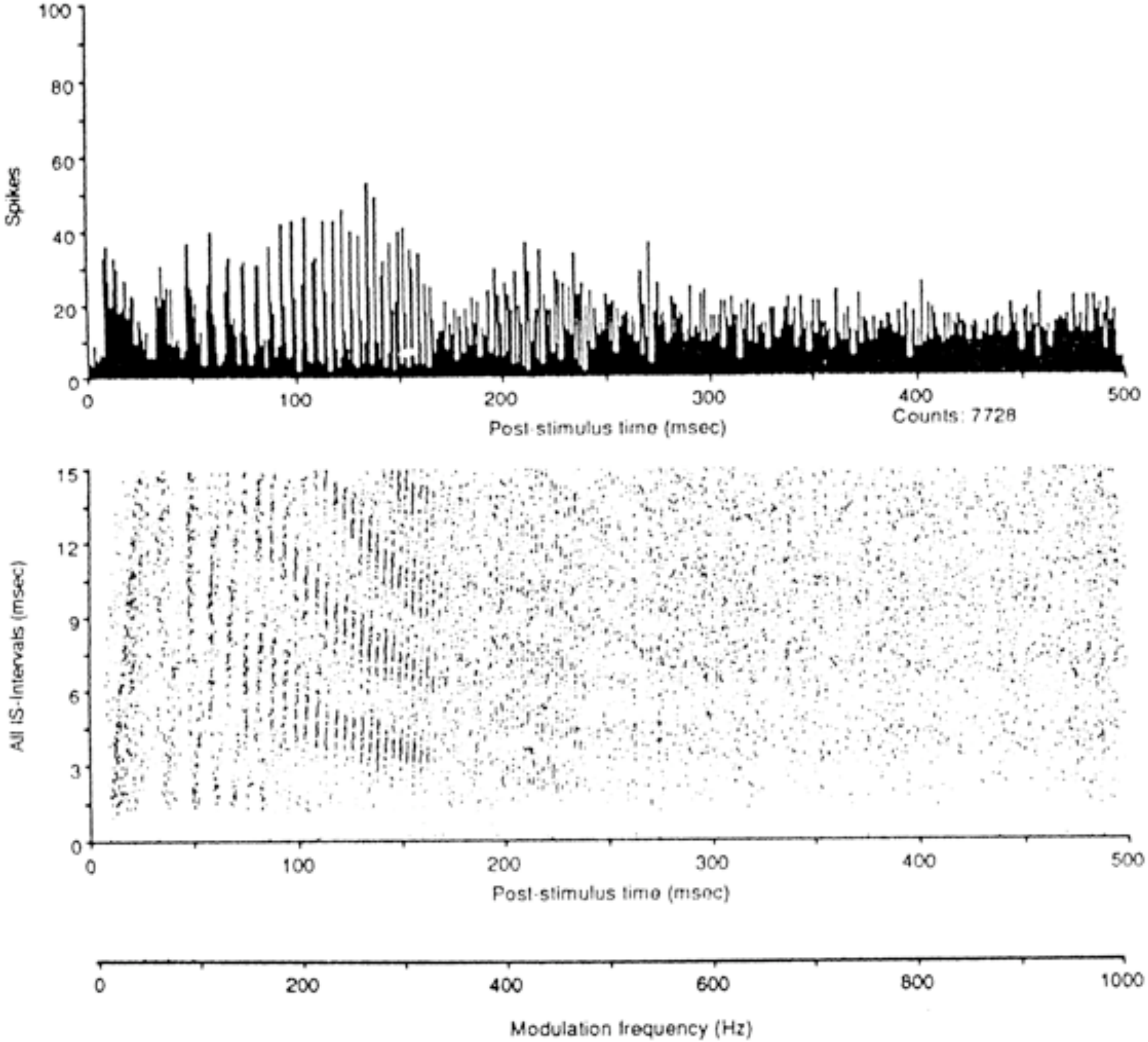


Figure 6. Responses of three units in the cochlear nucleus to 100 presentations of a single-formant vowel (F0= 80 Hz, F1 = 640 Hz, BW = 50) at 60 dB SPL. Units were classified according to their PSTH response to short tone bursts at CF. A. Dot-raster, PSTH, and all-order interval histogram for a primarylike unit in antero-ventral cochlear nucleus (AVCN), CF = 400 Hz. B. Response of a sustained chopper unit in posterior-ventral cochlear nucleus (PVCN), CF = 1.5 kHz. C. Response of a pauser unit in dorsal cochlear nucleus (DCN), CF = 4.4 kHz.

Cariani (1999)
Neural Plasticity

Responses of Two Posterior-Ventral Cochlear Nucleus (PVCN) Sustained Chopper (Chop-S) Units to an Amplitude Modulated Sweep at CF





Friday, March 13, 2009

Auditory central pathways: road map

Image removed due to copyright restrictions.
Figure 1 in Irvine, D. R. F. *The Auditory Brainstem*. New York, NY: Springer, 1986. ISBN: 9783540162995.

Brainstem stations involved in localization of sounds

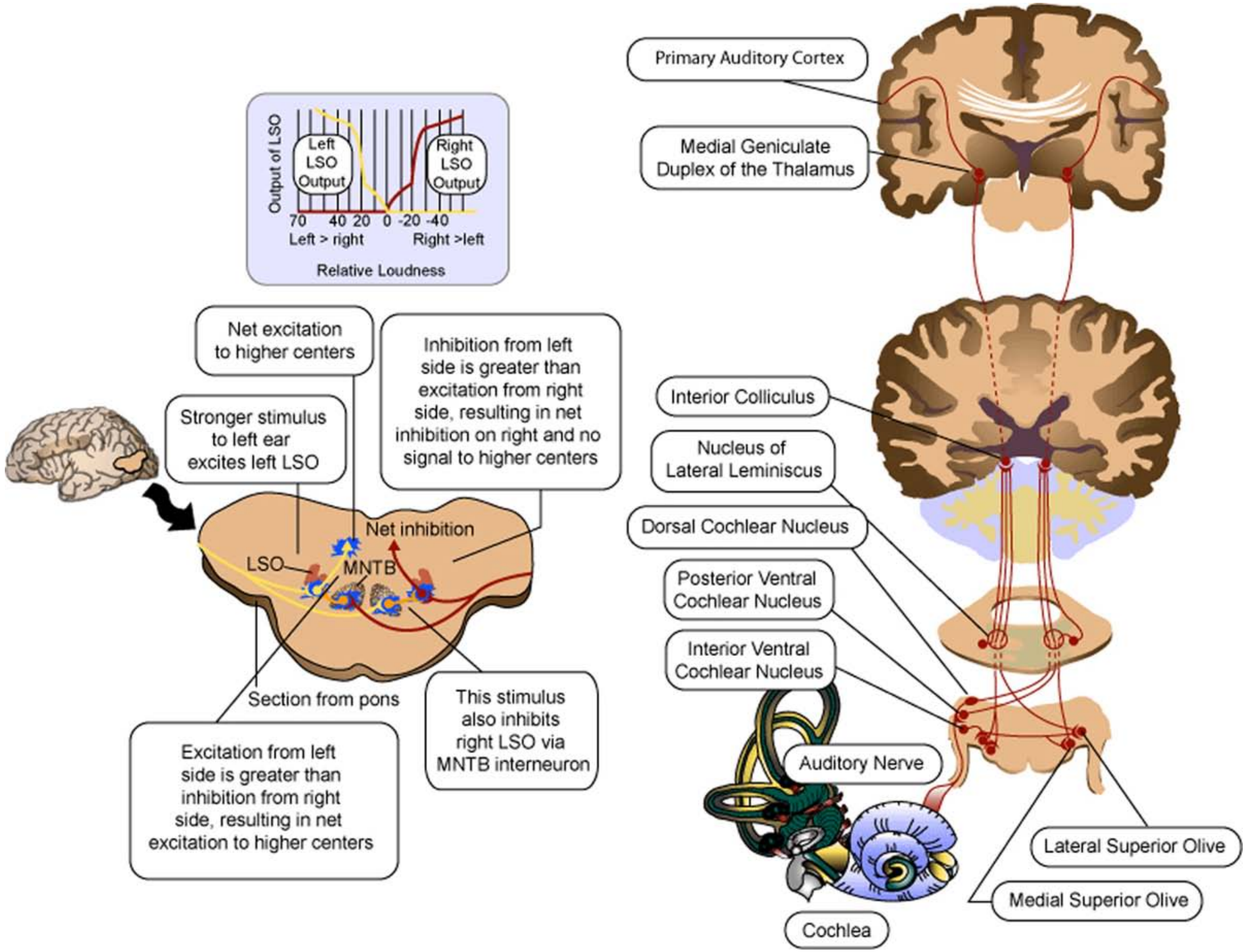
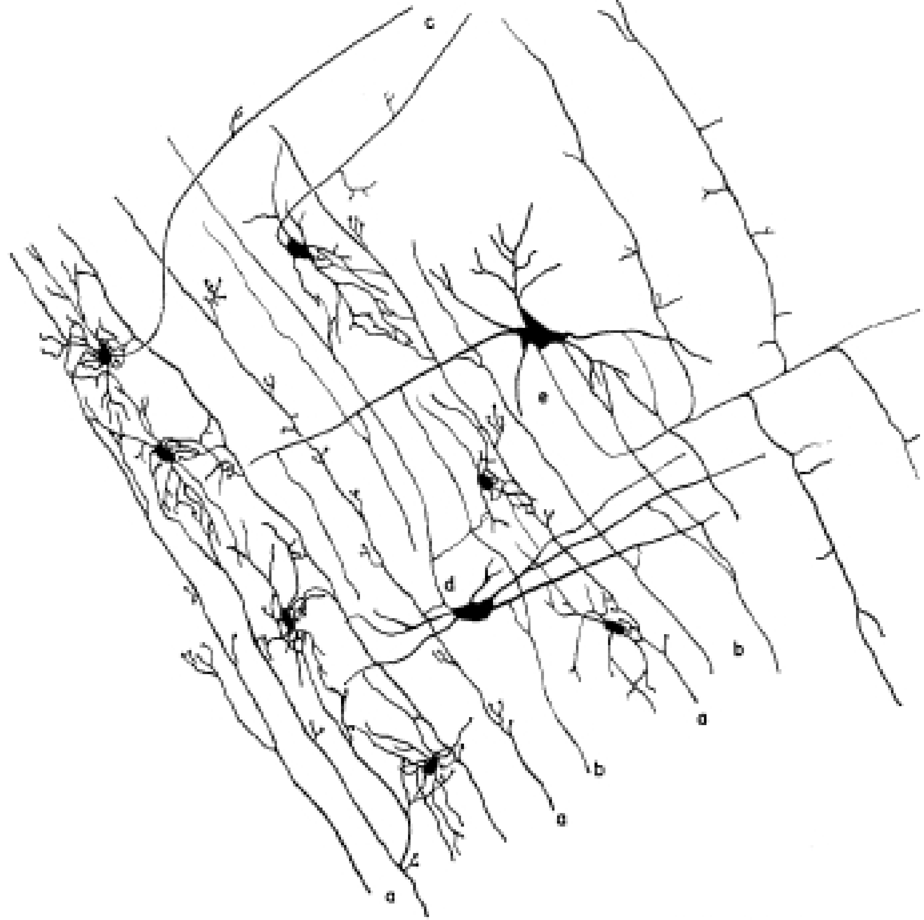


Figure by MIT OpenCourseWare.

Auditory midbrain: inferior colliculus



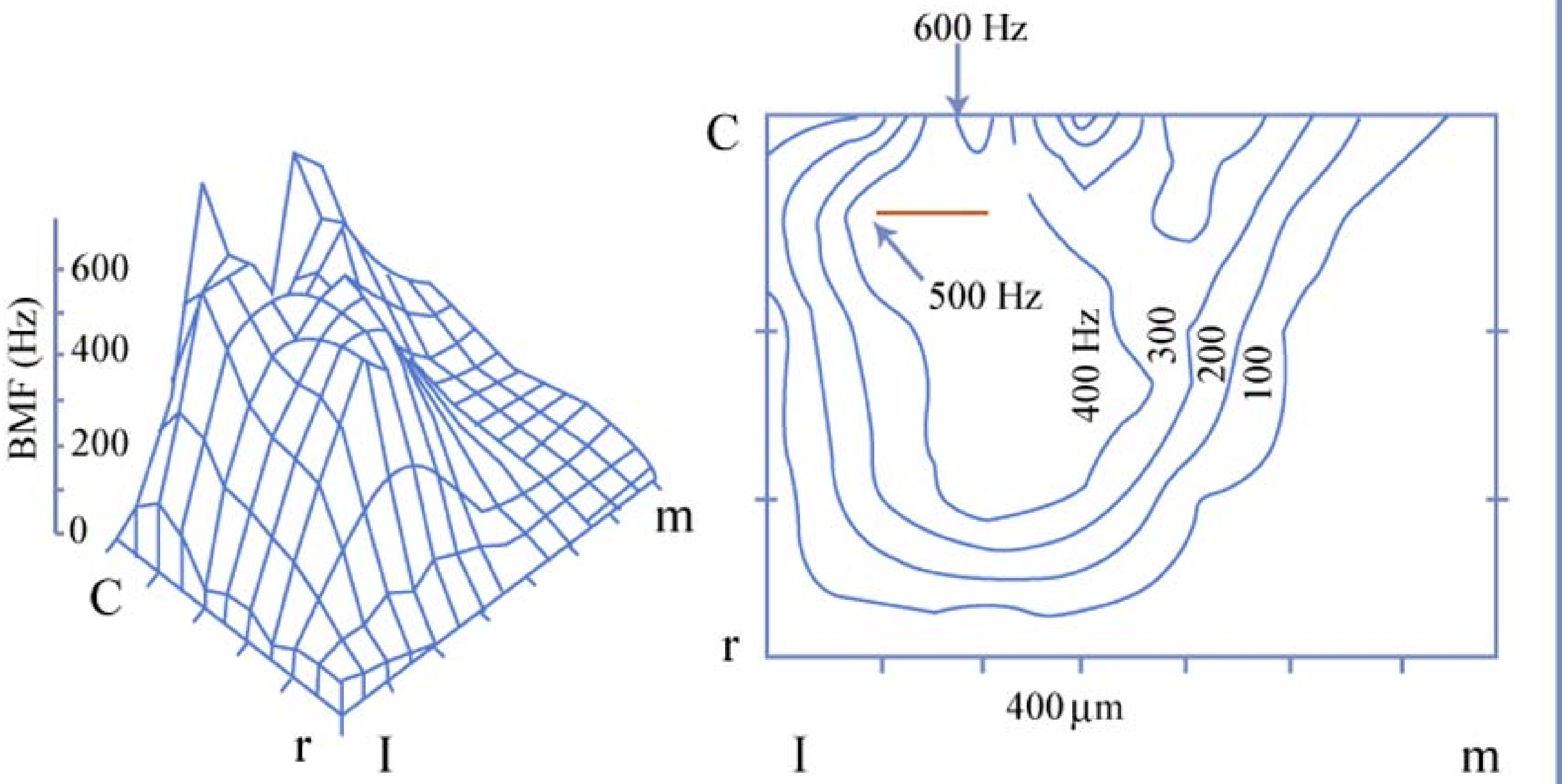
Sources: Fig. 3 in Morest, D. K., and D. L. Oliver. "The Neuronal Architecture of the Inferior Colliculus in the Cat: Defining the Functional Anatomy of the Auditory Midbrain." *J Comp Neurol* 222, no. 2 (1984): 209-236.

Fig 23 in Oliver, D. L., and D. K. Morest. "The Central Nucleus of the Inferior Colliculus in the Cat." *J Comp Neurol* 222, no. 2 (1984): 237-264.

Narrowly-tuned units in ICC (high BF)

Image removed due to copyright restrictions.
See Fig. 4.8 in Ehret, G. "The Auditory Midbrain..." in
The Central Auditory System. Edited by G. Ehret and R. Romand.
New York, NY: Oxford University Press, 1997.
[Preview this image in Google Books.]

Auditory midbrain: periodotopy?



View on the 3 kHz isofrequency plane of the inferior colliculus of a cat. Best modulation frequencies (BMF of amplitude modulation) are indicated as three-dimensional contour lines (left) and as iso-best modulation lines (right).

Figure by MIT OpenCourseWare.

Modulation detectors in the midbrain

Problems:

1) MTF tuning degrades at high SPLs & in noise

2) Wrong operation. Modulation tuning does not account for pitches of resolved harmonics of inharmonic tones (pitch-shift exps)

3) Representation will degrade when multiple F0s are present (doesn't support scene analysis)

4) Does not explain pitch equivalence of pure & complex tones

5) Structural. Could be due to ratio of excitation-inhibition rather than for specific function

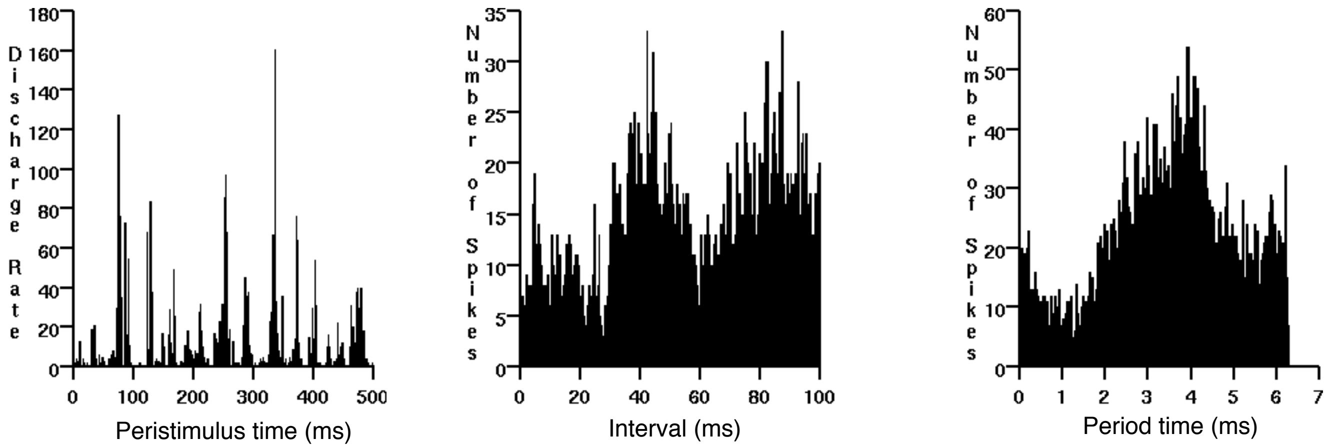
Sources for auditory CNS figures: Günter Ehret (1997) The auditory midbrain, a “shunting yard” of acoustical information processing. In: The Central Auditory System, Ehret, G. & Romand, R., eds. Oxford University Pres. Langner, G. and Schreiner, C.E. Periodicity coding in the inferior colliculus of the cat. I. Neuronal mechanisms. J. Neurophysiol. 60:1799-1822. See also Langner (1992) review, Periodicity coding in the auditory system. Hearing Research, 60:115-142.

Stimulus-related temporal discharge patterns in IC (PTs to ~4 kHz, F0s to 1200 Hz)

See Fig. 2 in Langner, G. and Schreiner, C.E.
"Periodicity coding in the inferior colliculus of the cat. I. Neuronal mechanisms."
J. Neurophysiol. 60 (1988): 1799-1822.

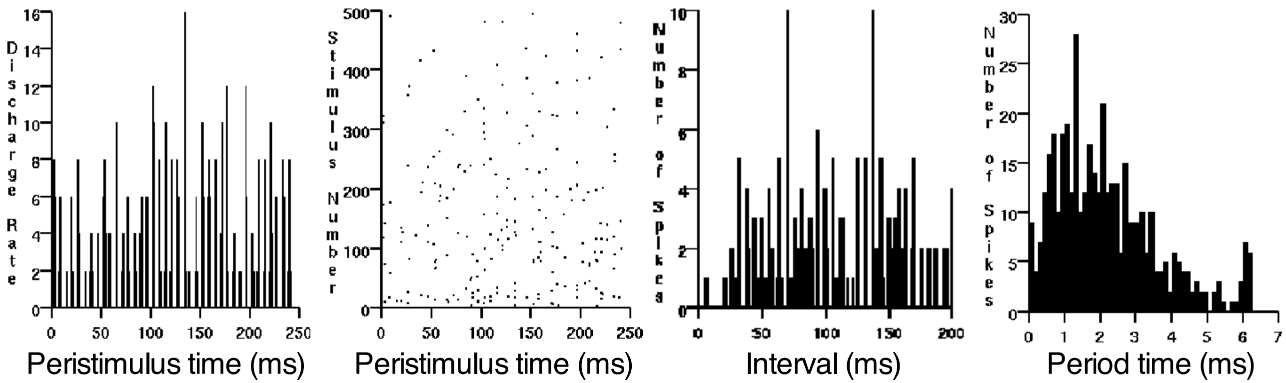
Coding of pitch in the inferior colliculus

AM broadband noise, Fm = 160 Hz, 1000 contralateral monotic presentations @ 80 dB SPL



PST histogram, all-order interval histogram, and period histogram (6.25 ms analysis period). Total number of spikes: 4421. Note the longer (~40 ms) preferred intervals for this unit and the pitch-related spacings (6.25 ms) between the individual interval peaks.

Click train, F0 = 160 Hz, 500 contralateral monotic presentations @ 80 dB SPL



Total number of spikes: 418. Patterns of longer intervals are pitch-related.

Upper limits of temporal pattern information (rough estimate)

Cochlear hair cells: no limit, but weakening AC component
Auditory nerve: < 4-5 kHz abundant & highly significant;
statistical significance depends on #spikes (> 5 kHz)
Cochlear nucleus: depending
Midbrain: 4-5 kHz in inputs (frequency-following response)
Interval information: 1/F0 up to ~1200 Hz
Thalamus: 10% of units lock to 2-3 kHz with SI > 0.3
(deRibaupierre, lightly anesthetized preps)
Primary cortex: 200 Hz averaged gross surface potentials
(unanesthetized, 100 Hz anesthetized; Goldstein
& Kiang, 1959);
300 Hz averaged gross potentials (CSD, input layers,
Steinschneider et al); anecdotal reports of locking
to 1 kHz in single units, but these are very rare
Rule-of-thumb: anesthesia decreases fmax by factor of 2

Auditory thalamus: medial geniculate body

Image removed due to copyright restrictions.
See Figure 1 in Morest, D. K. "The Neuronal Architecture of the Medial Geniculate Body of the Cat."
J Anat 98 (October 1964): 611-30. Available online at http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1261345/.

Gyri: hills
Sulci: valleys

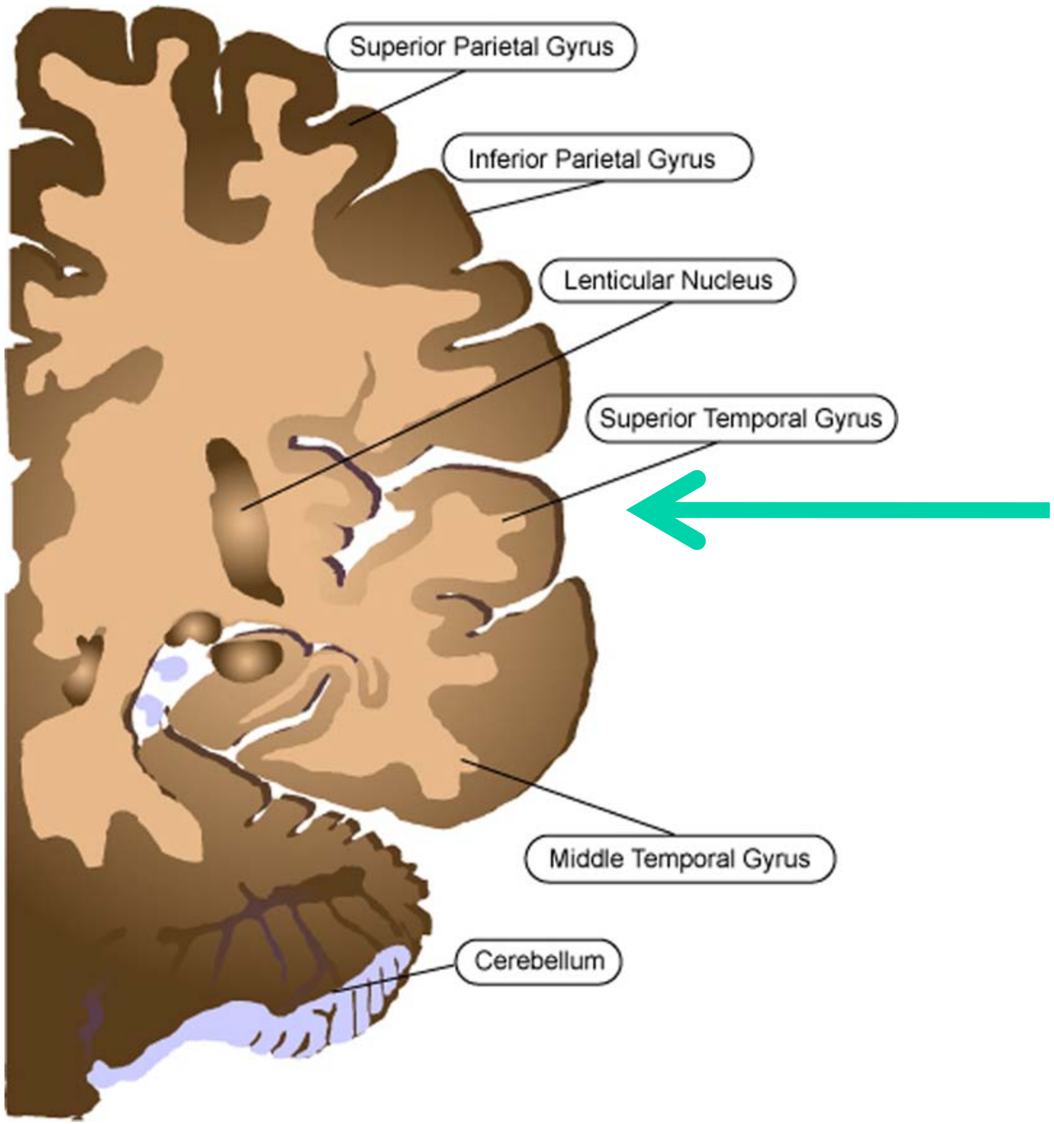
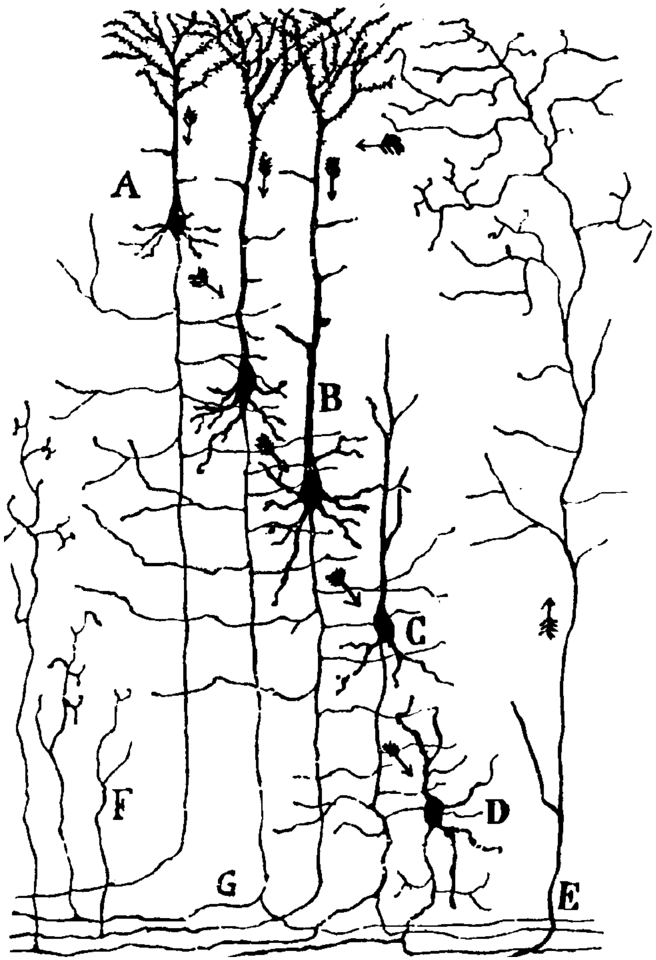


Figure by MIT OpenCourseWare.

Auditory cortex is located in the Superior Temporal

Laminated “cortical” structures



Ramon y Cajal

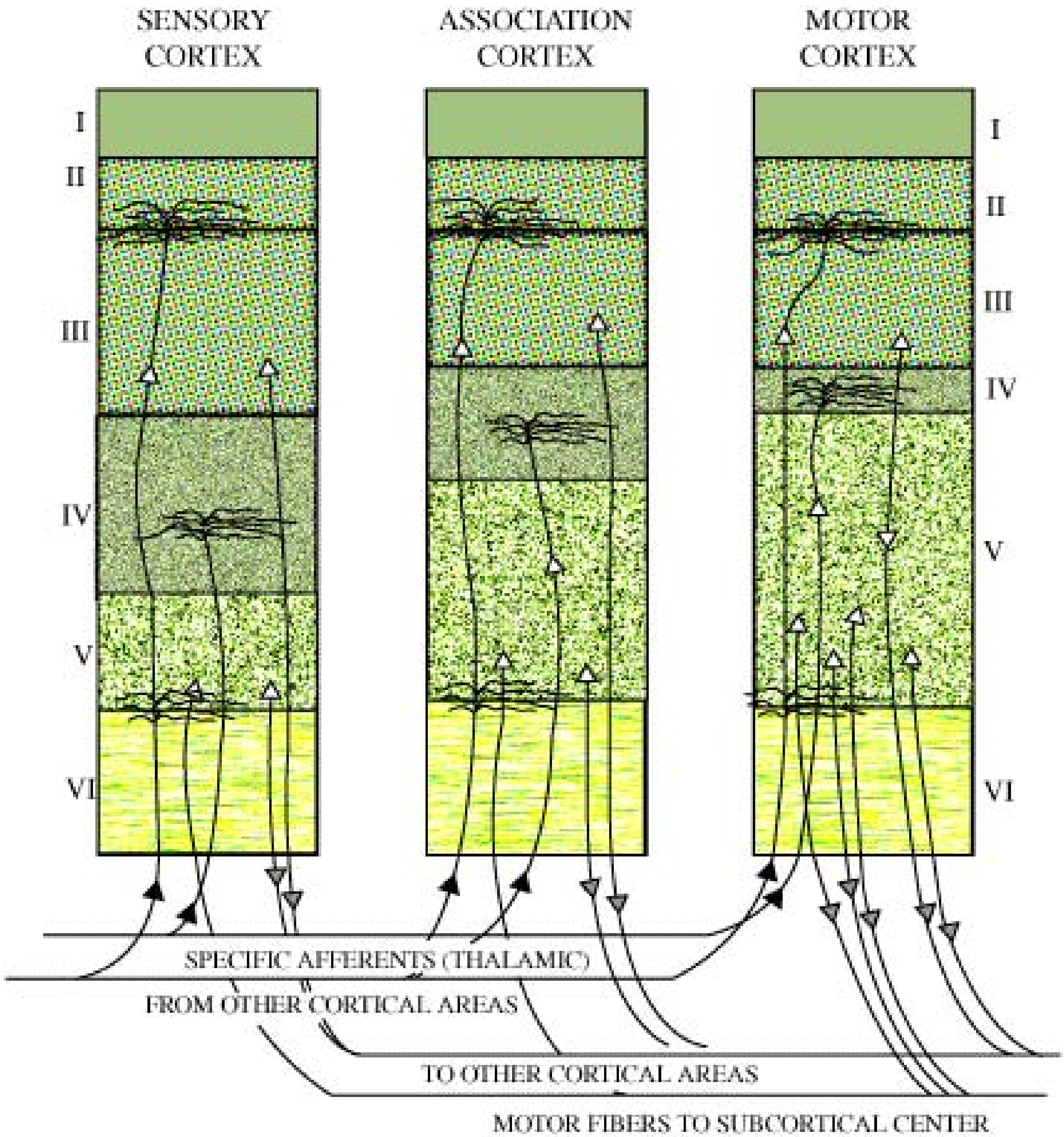


Figure by MIT OpenCourseWare, after Sutherland / Woodburne (1967).

Primary and secondary auditory cortex regions

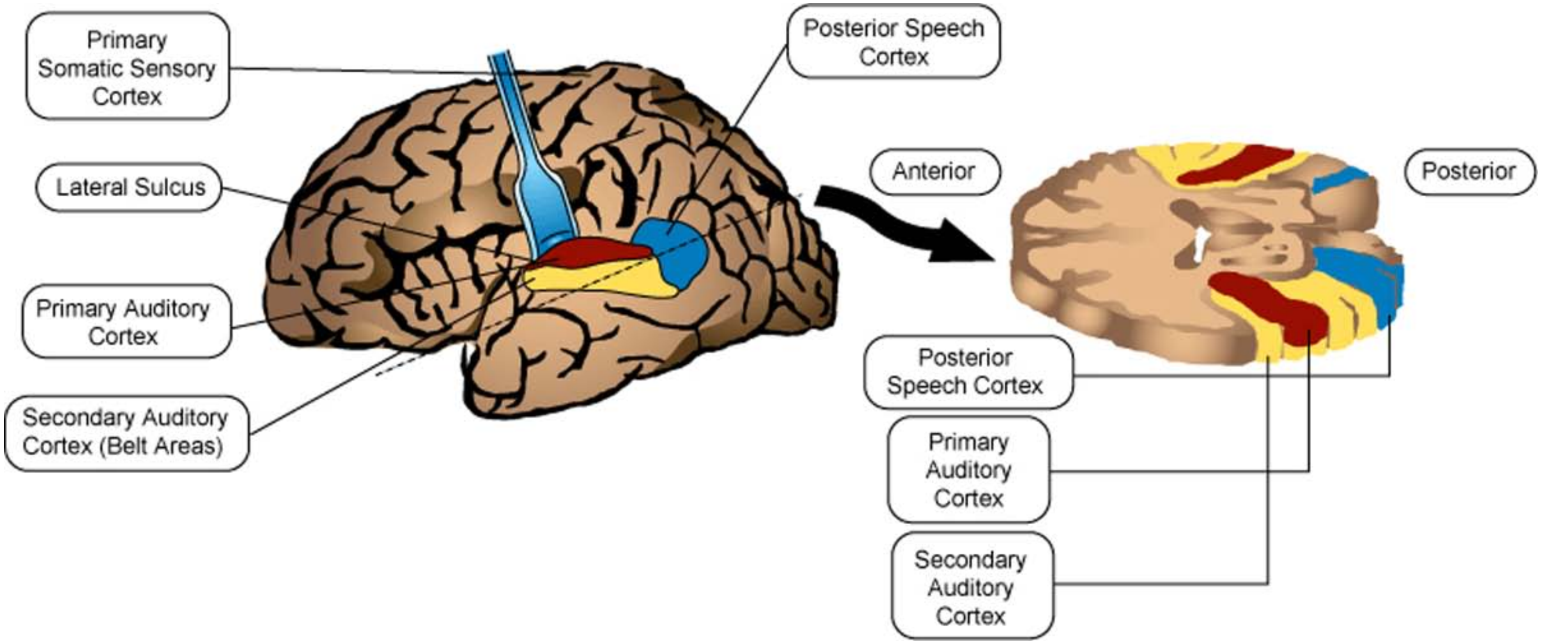


Figure by MIT OpenCourseWare.

Cochleotopic organization of auditory cortex (cartoon)

Two concepts best kept separate in one's mind:

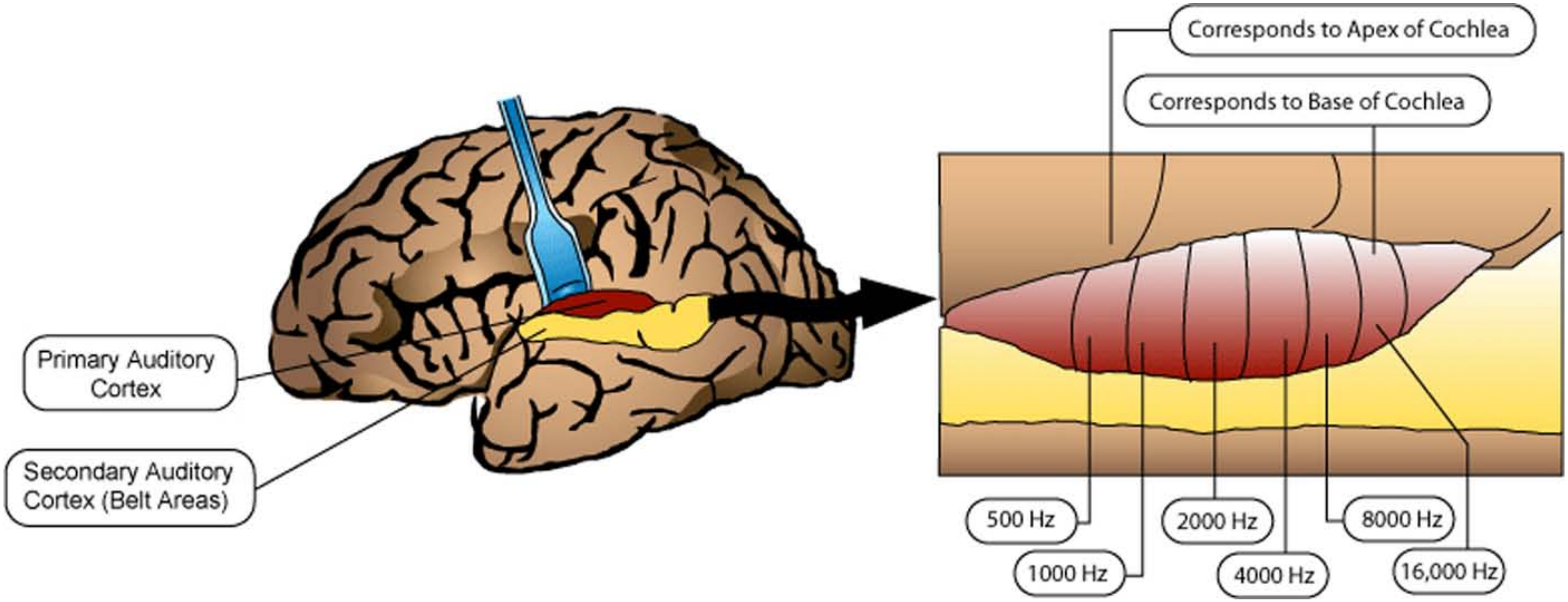


Figure by MIT OpenCourseWare.

[Purves et al])

Auditory cortex: cat

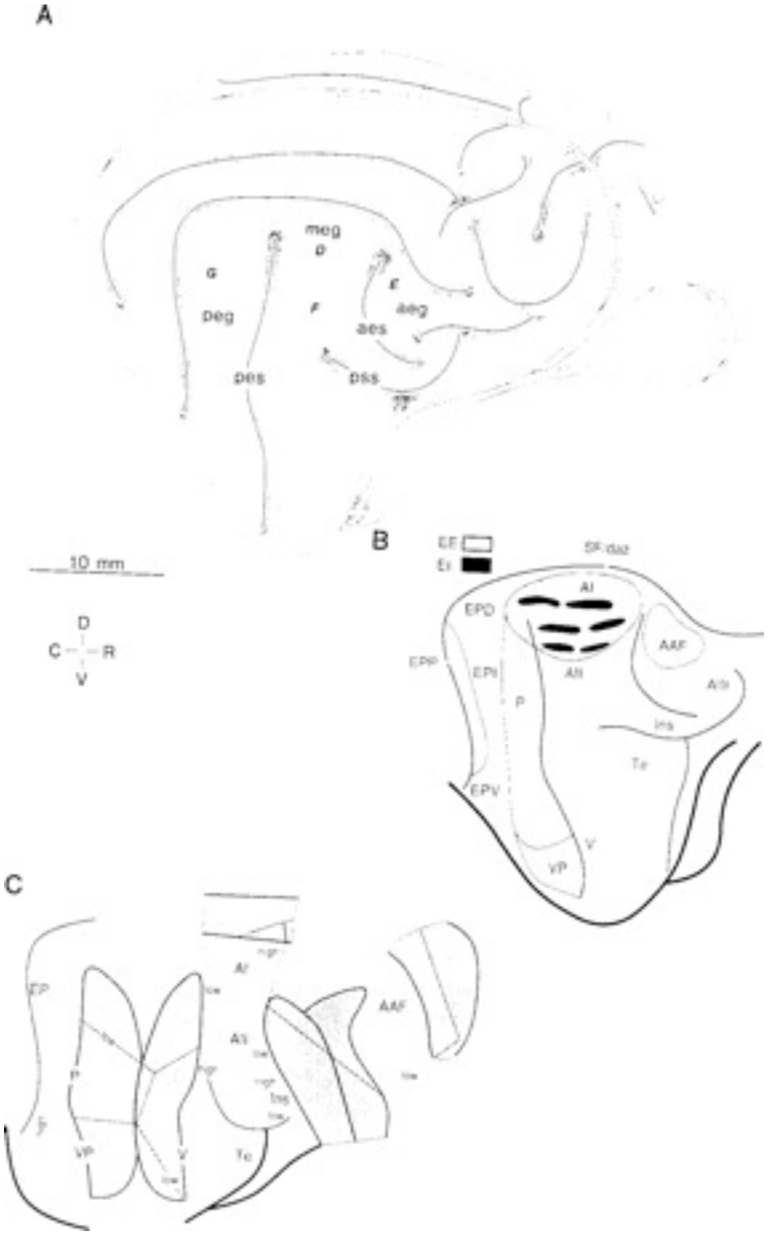
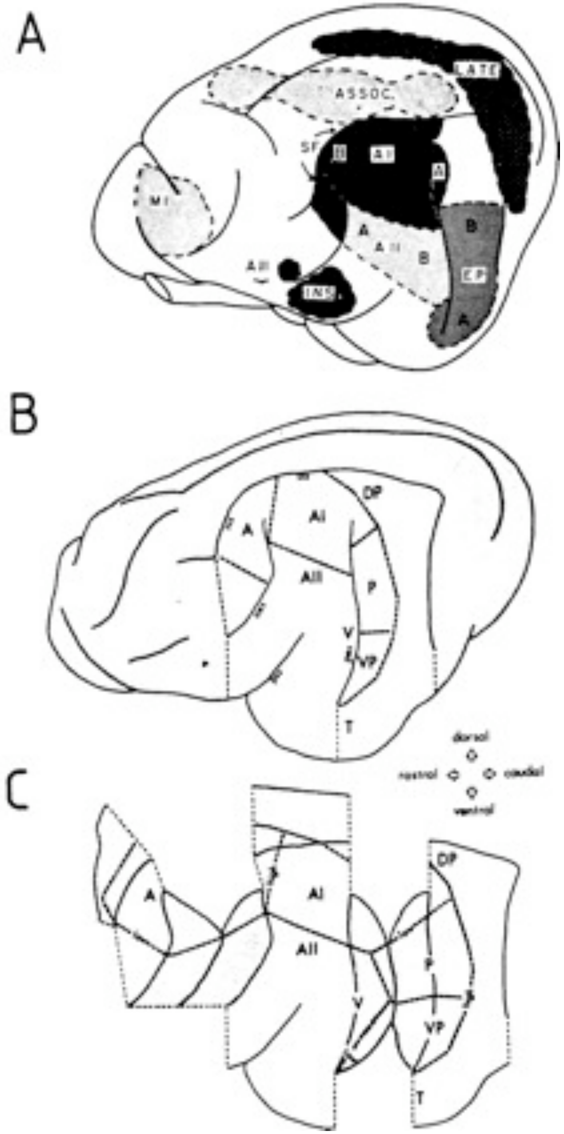


FIG. 23. Auditory cortical fields in cat. *A*: parcellation proposed by Woolsey (487) in 1960, showing cortical fields (AI, AII, EP, SF, INS, and AIII) described in text. In each field, orientation of cochlear representation is indicated by position of A (apex) and B (base). Auditory responses also recorded in association areas on suprasylvian and anterior lateral gyri (ASSOC.) and in pericruciate sensorimotor cortex (MI). Long-latency (LATE) responses are recorded in visual cortex. *B*, *C*: parcellation into four tonotopic (AI, A, P, and VP) and additional belt fields (AII, DP, V, and T) described by Reale and Imig (376). Because location of physiologically determined field boundaries in different animals varies with respect to sulci, positions shown are general approximations only. In *B*, field positions are shown relative to sulci on a lateral view of surface of left cerebral hemisphere; sss: suprasylvian sulcus; aes: anterior ectosylvian sulcus; pes: posterior ectosylvian sulcus; pss: pseudosylvian sulcus. In *C*, unfolded cortical surface forming gyral surfaces and sulcal banks in unshaded region in *B* is shown. Cortical surfaces forming sulcal banks are shaded, whereas those forming gyral surfaces are not. Tonotopic fields are delimited by heavy broken lines, and locations of highest and lowest best frequencies in these fields are indicated by *low* and *high*, respectively. [*A*: from Woolsey (487); *B*, *C*: from Imig and Reale (216).]



Images: Public domain

Auditory central pathways:

Fig. 1.11 (p. 38) in De Ribaupierre, F. "Acoustical Information Processing in the Auditory Thalamus and Cerebral Cortex." In *The Central Auditory System*. Edited by G. Ehret and R. Romand. New York, NY: Oxford University Press, 1997.
[Preview this image in Google Books]

Auditory central pathways: cortico-thalamic connections

Fig. 1.12 (p. 39) in De Ribaupierre, F. "Acoustical Information Processing in the Auditory Thalamus and Cerebral Cortex." In *The Central Auditory System*. Edited by G. Ehret and R. Romand. New York, NY: Oxford University Press, 1997.
[Preview this image in Google Books]

Auditory cortex: responses to high frequency pure tones

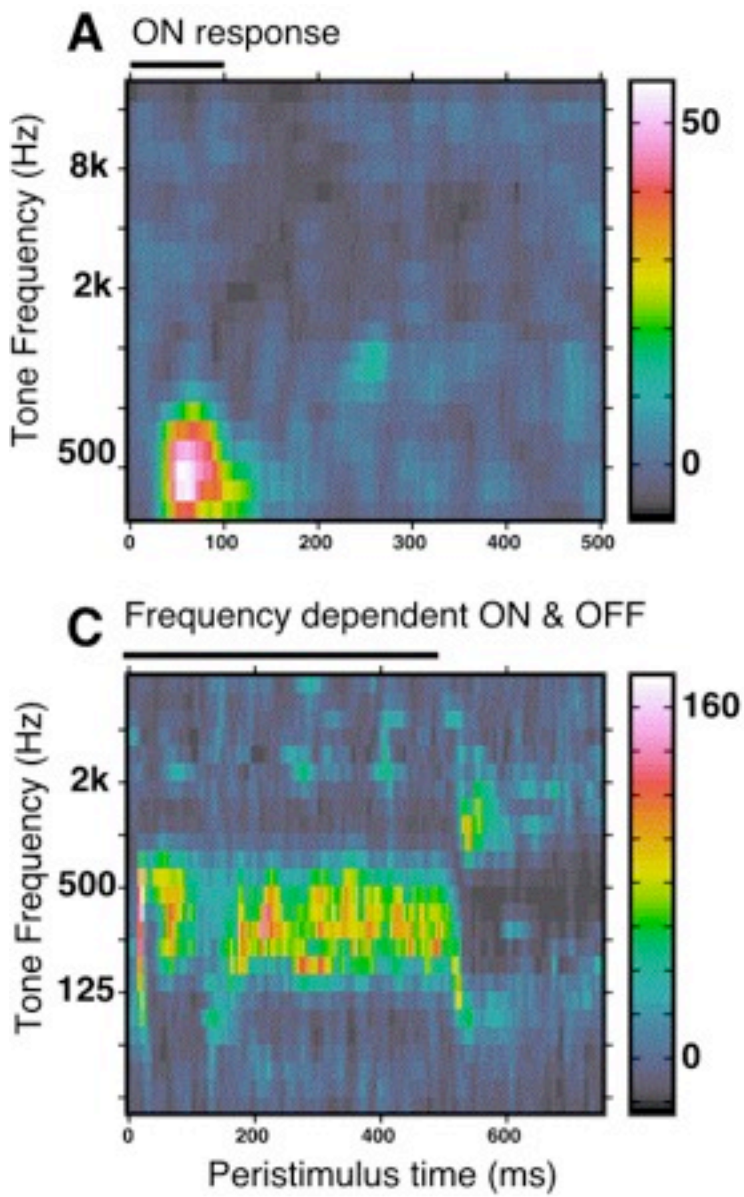
Figures removed due to copyright restrictions.
Fig. 2, 3, 4 and 9 in Phillips, D. P., et al. "Level-dependent Representation of Stimulus Frequency in Cat Primary Auditory Cortex." *Exp Brain Res* 102 (1994): 210-226. DOI: 10.1007/BF00227510.

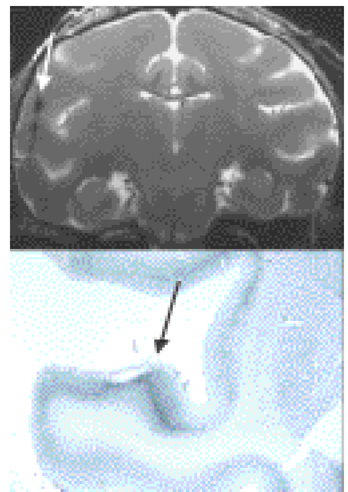
Pitch-related temporal patterns in field potentials in awake monkey cortex

Figure. Averaged cortical field potentials (current source density analysis, lower lamina 3, site BF=5 kHz) in response to 50 ms click trains F0=100-500 Hz. Ripples up to 300-400 Hz show synchronized component of the ensemble-response, From Steinschneider (1999).

Image removed due to copyright restrictions.
See Fig. 9 right, in Steinschneider, M., et al. "Click Train Encoding in Primary Auditory Cortex of the Awake Monkey: Evidence for Two Mechanisms Subserving Pitch Perception." *J Acoust Soc Am* 104, no. 5 (1998): 2935-2955. DOI: 10.1121/1.423877.

Pure tone temporal response profiles in auditory cortex (A1)





Source: Tramo, Mark J. "Neural Representations of Acoustic Information in Relation to Voice Perception." Havard University PhD Thesis, 1999.

Where everything takes place: from cochlea to cortex, and beyond

10,000k	Primary auditory cortex (Auditory forebrain)
	Auditory thalamus
500k	Inferior colliculus (Auditory midbrain)
	Lateral lemniscus
	Auditory brainstem
30k	Auditory nerve (VIII)
3k	Cochlea

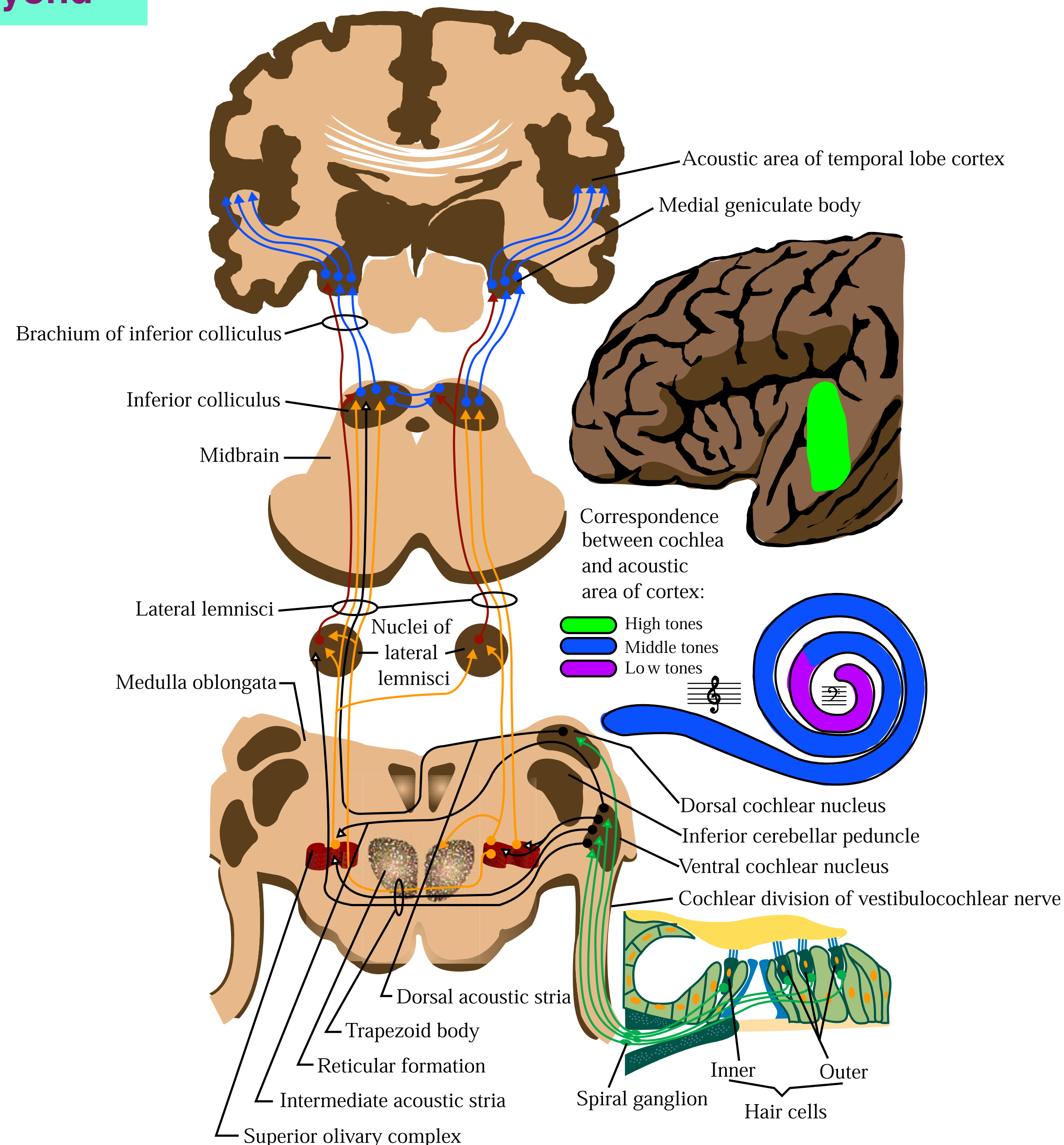


Figure by MIT OpenCourseWare.

Tonotopy, cochleotopy & frequency maps: Common pitfalls

Orderly spatial arrangements of frequency-tuned neurons ("auditory frequency maps") exist at every auditory station.

However, these maps are coarse relative to perceptual discriminations, especially for low frequencies (< 2 kHz) and for moderate to high sound levels (> 60 dB SPL).

I have yet to see evidence in the literature for neuronal tuning finer than about 1/2 octave for low frequency tones at high levels (barely good enough to resolve the 2nd harmonic).

In auditory cortex the ordering of frequency tunings is only seen at very low sound levels -- tonotopy breaks down at moderate to high levels (> 60 dB SPL).

Tonotopy: seen at all auditory stations

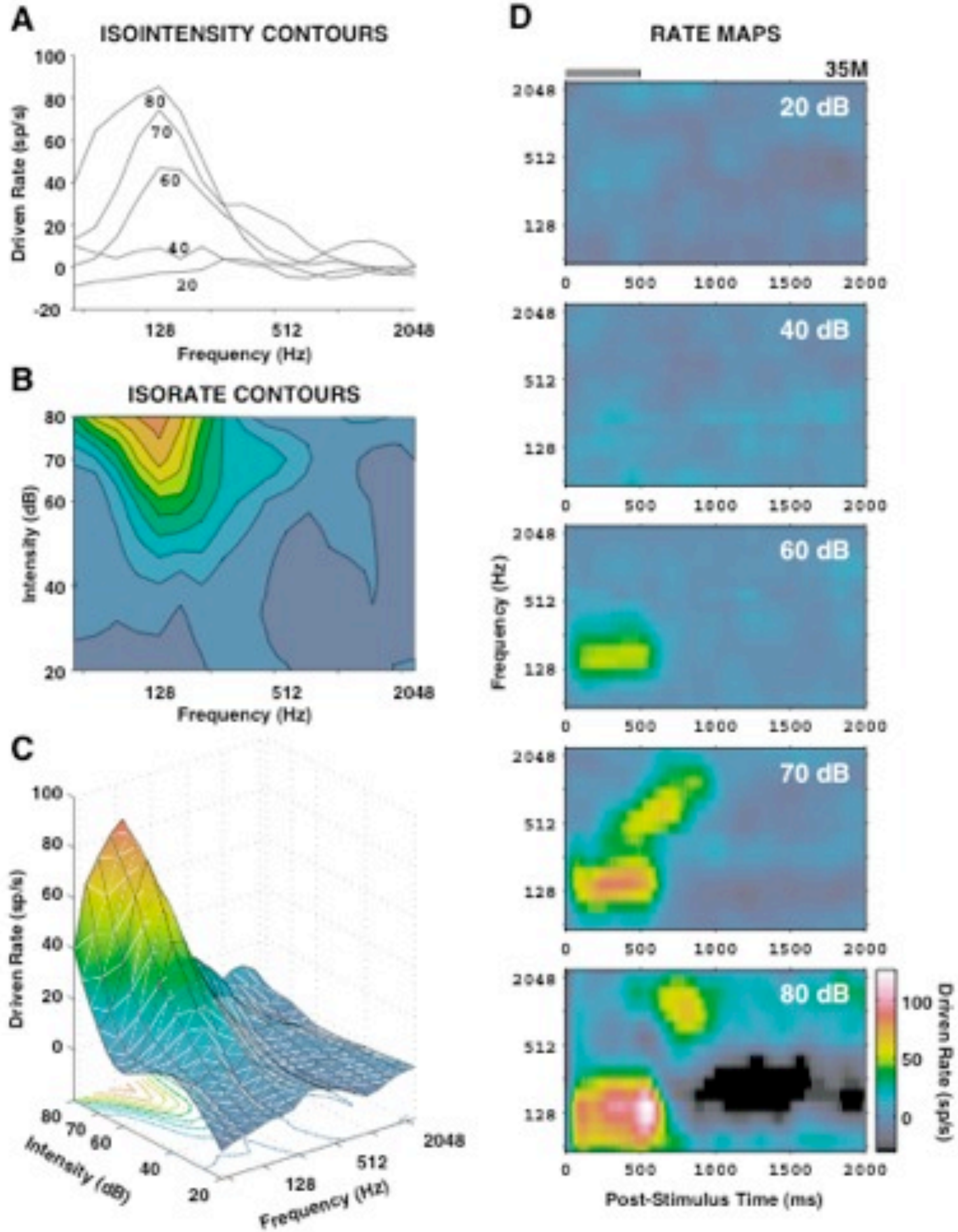
- Simple tonotopic order only seen at levels near neural thresholds; this order breaks down at mod-high levels
- At every auditory station, tuning of most units broadens at higher intensities (especially for tones < 1 KHz; exceptions to this rule usually involve high-BF units)
- Q values (BW/BF) increase with BF; however frequency discrimination declines with BF
- Does not solve the problem of pitch of complex tones
  - Additional mechanisms are needed
- Tonotopy likely reflects mappings of most direct connections to sensory surfaces rather than carrying the information for frequency coding per se

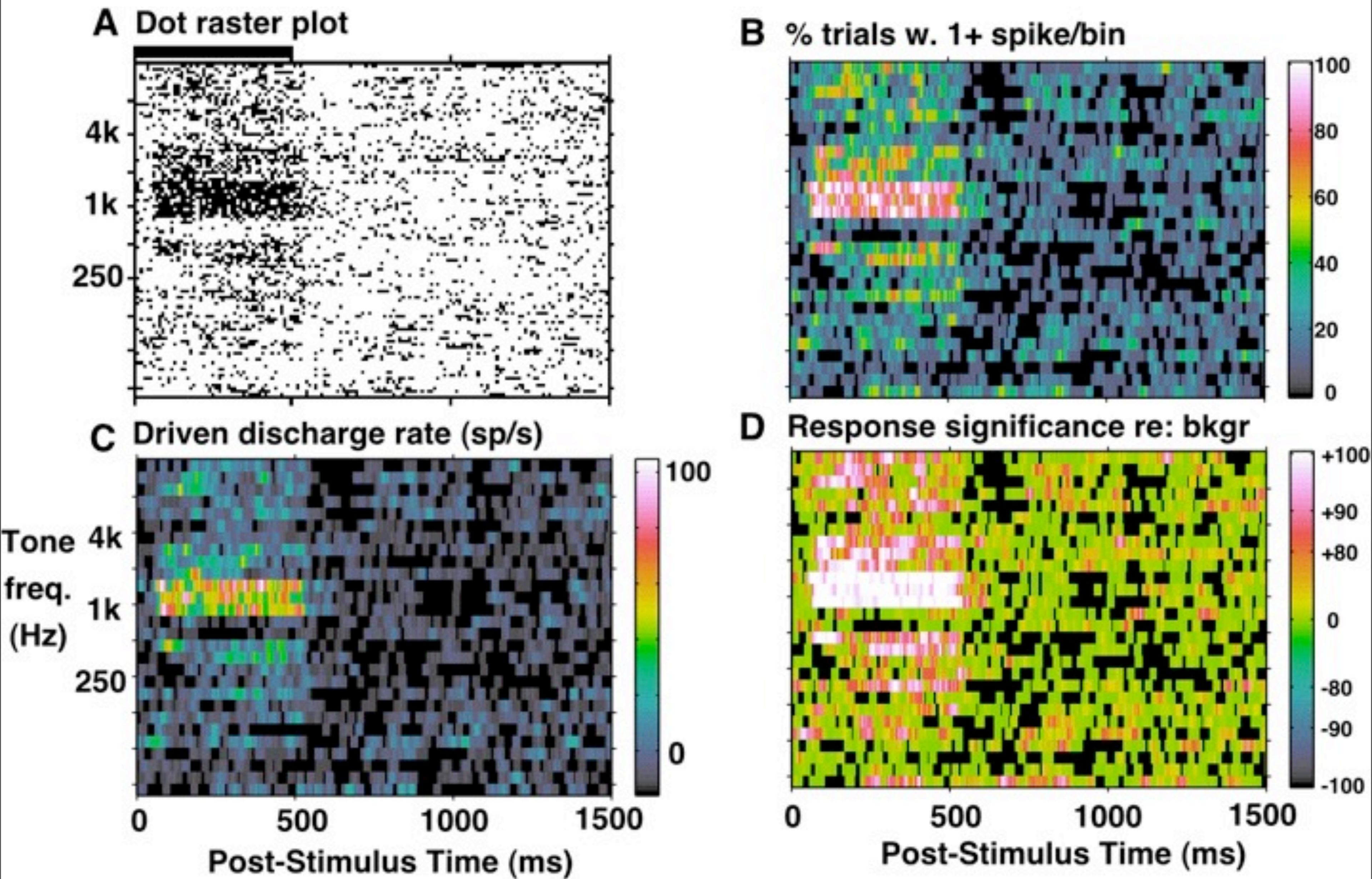
Figure removed due to copyright restrictions.
Fig. 9 in Phillips, D. P., et al. "Level-dependent Representation of Stimulus Frequency in Cat Primary Auditory Cortex."
Exp Brain Res 102 (1994): 210-226. DOI: 10.1007/BF00227510.

Phillips et al 1994

Narrowly-tuned units in auditory cortex (high BF)

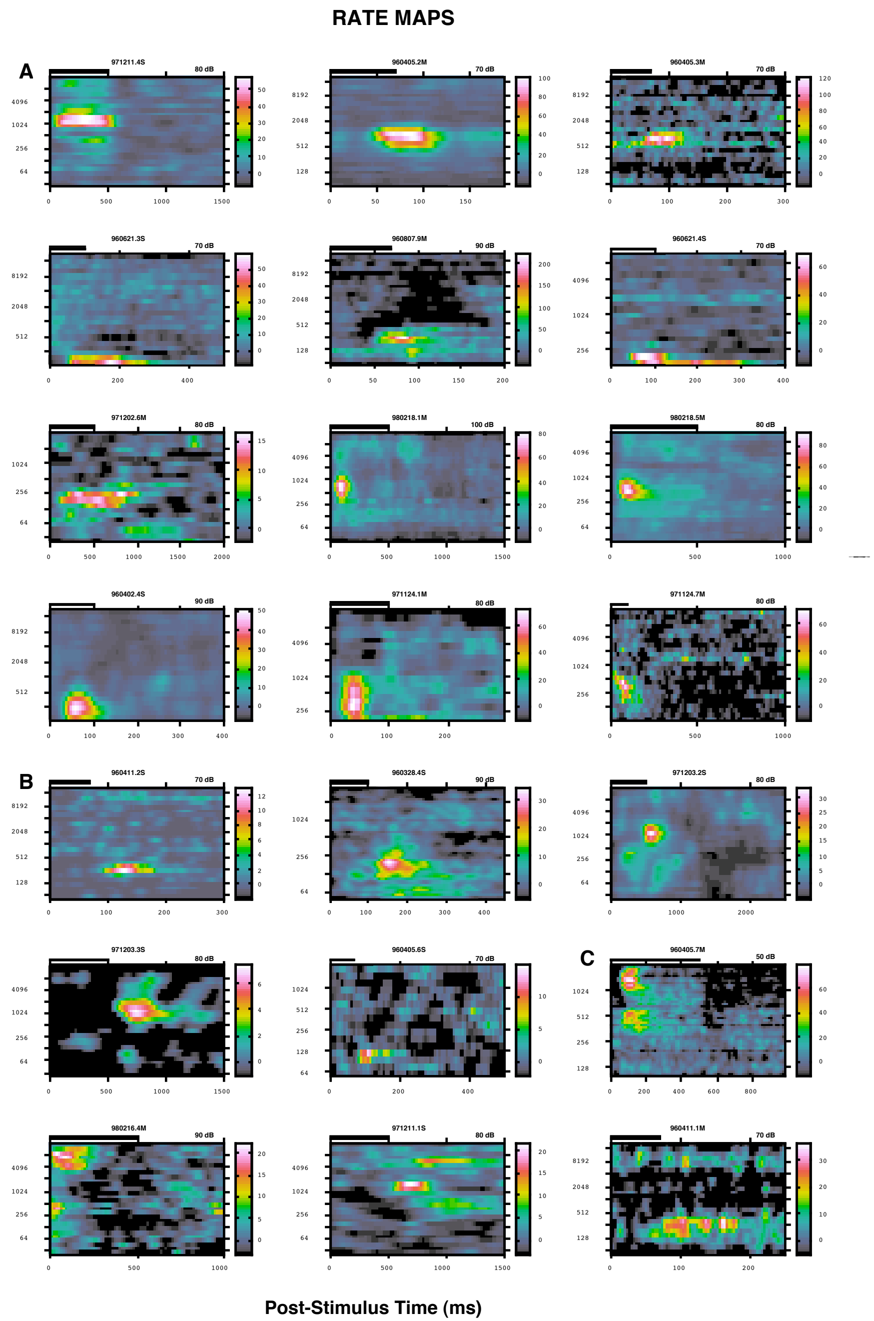
Image removed due to copyright restrictions.
Set of six graphs (latency, intensity and spike count vs. tone frequency) from Phillips, 1989.


A
ISOINTENSITY CONTOURS
Driven Rate (sp/s)
Frequency (Hz)
B
ISORATE CONTOURS
Intensity (dB)
Frequency (Hz)
C
Driven Rate (sp/s)
Intensity (dB)
Frequency (Hz)
D
RATE MAPS
35M
20 dB
40 dB
60 dB
70 dB
80 dB
Frequency (Hz)
Post-Stimulus Time (ms)
Driven Rate (sp/s)



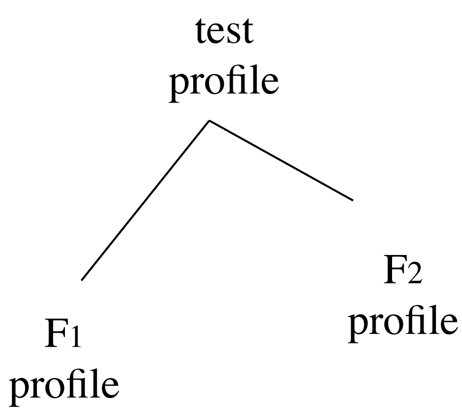
Source: Tramo, Mark J. "Neural Representations of Acoustic Information in Relation to Voice Perception." Havard University PhD Thesis, 1999.

Temporal response profiles

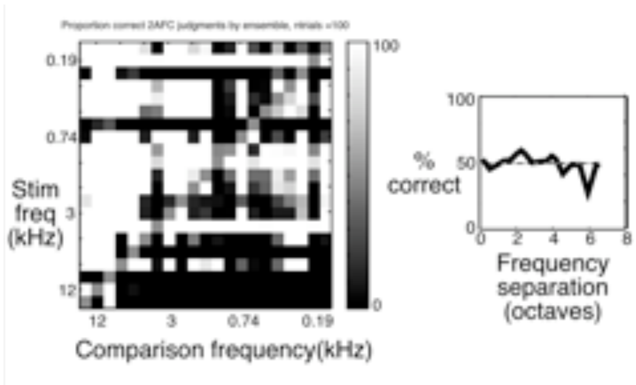


Courtesy of Prof. Mark J. Tramo, M.D., Ph.D. Used with permission.
Source: Tramo, Mark J. "Neural Representations of Acoustic Information in Relation to Voice Perception." Havard University PhD Thesis, 1999.

Rate-frequency profiles for 15 cortical ON units



Decision analysis



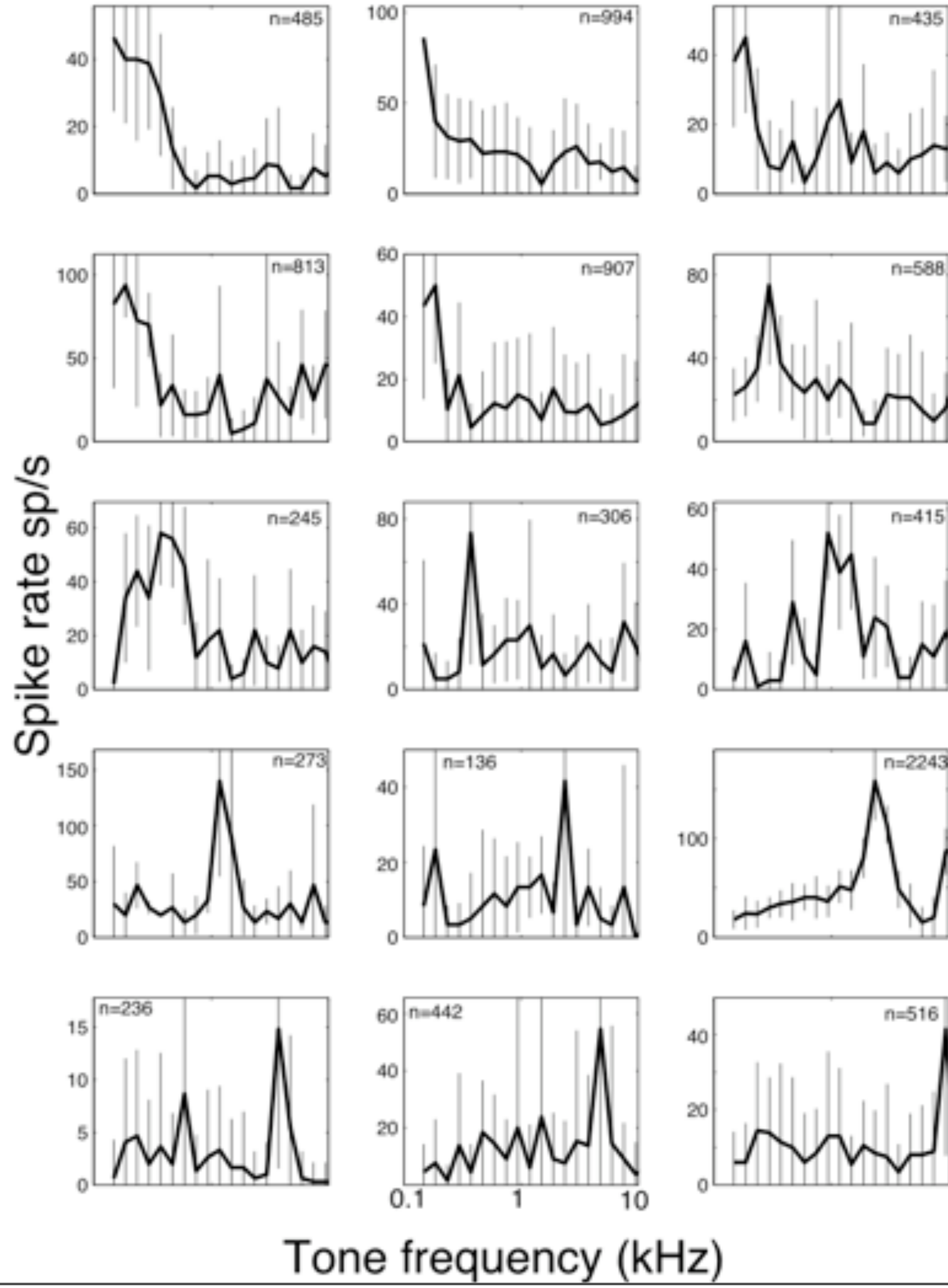


Image removed due to copyright restrictions.
See Fig. 7 in Schwarz, D. W., and R. W. Tomlinson.
"Spectral response patterns of auditory cortex neurons to harmonic complex tones in alert monkey (Macaca mulatta)." *J Neurophysiol* 64, no. 1 (1990): 282-298.

Schwarz & Tomlinson 1990

- The results of lesion studies motivated by interest in music and the brain have led to major revisions in fundamental hypotheses about the functional role of primary auditory cortex (A1) in frequency processing and pure-tone pitch perception

- The results of single- and multi-unit neuron recordings in A1 raise questions about the functional relevance of tonotopy and “sharp-tuning” to pitch perception

Bendor & Wang(2005) F0-tuned units in auditory cortex

See Fig. 1 in Bendor and Wang, "The neuronal representation of pitch in primate auditory cortex." *Nature* 436 (2005): 1161-1165.

Bendor & Wang (2005)

F0-tuned neurons:

first evidence of "true" F0 sensitive neurons

coarsely tuned (1 octave)

not clear what the SPLs are

nonmonotonic responders

High degree of level dependence begs the question of how a rate-based representation using these units can account for level-invariance of the pitch percept (same problem as Phillips et al, 1994)

Some of the difficulties: rate-place profiles

- Saturation of firing rates at higher levels (> 80 dB SPL)
- Units are generally coarsely tuned (ctx neural bandwidths 0.5-2 oct)
- Disconnect between freq. discrim. and neural Q values
  - (Reccanzone, however correlation with cortical territory/# neurons)
- High response variability; low firing rates
- May be difficult to account for jnd's < 1%, esp. at higher levels (Siebert's classical analysis was carried at lower SPLs)
- No mechanisms for complex tones are evident
- Components spaced < 300 Hz apart not resolved in either cat auditory nerve or macaque auditory ctx (Steinschneider)
- No low-BF harmonic combination units seen

How do higher auditory stations represent and process sounds?

- What is the fate of neural timing information?
- How does the auditory CNS make use of it?
- Where do representations responsible for fine pitch distinctions reside?

What are the central neural codes & computations?

Tramo, Mark Jude **BIOLOGY AND MUSIC:**
Enhanced: Music of the Hemispheres
Science 2001 291: 54-56

Music & Cortex

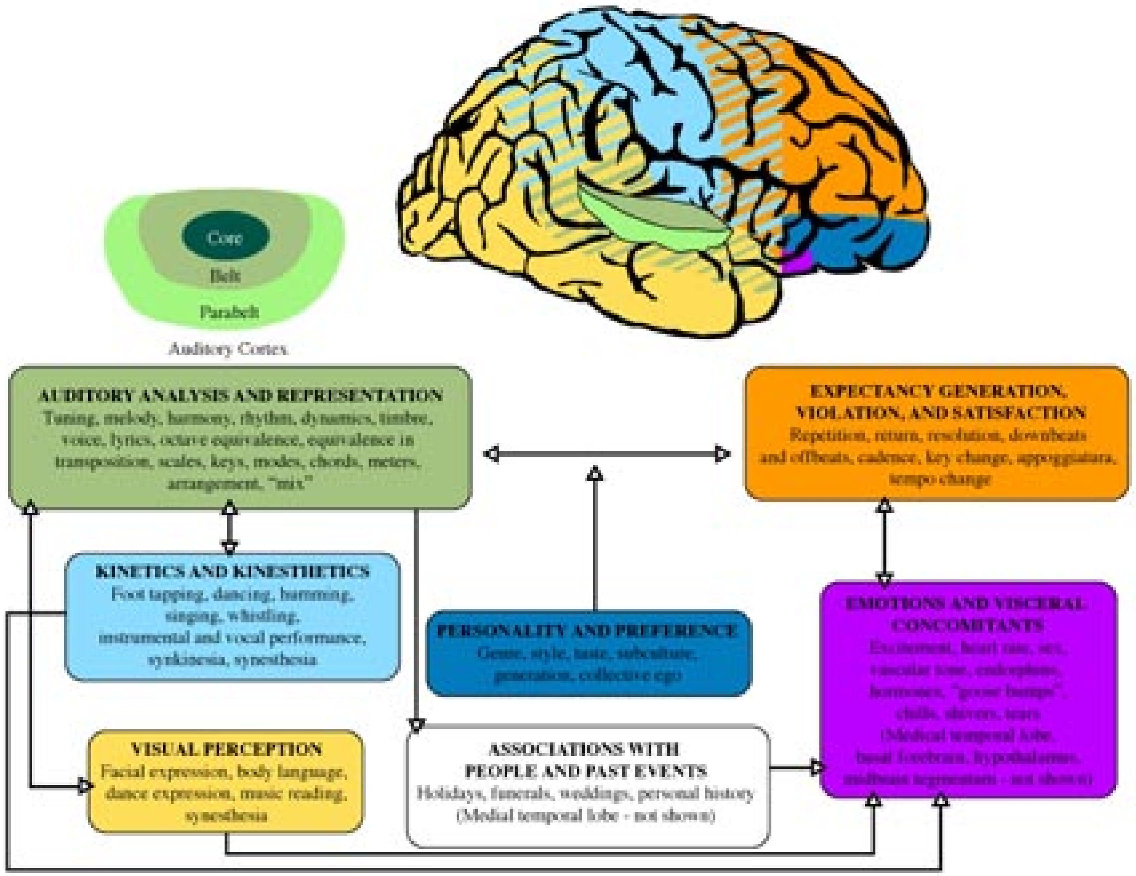
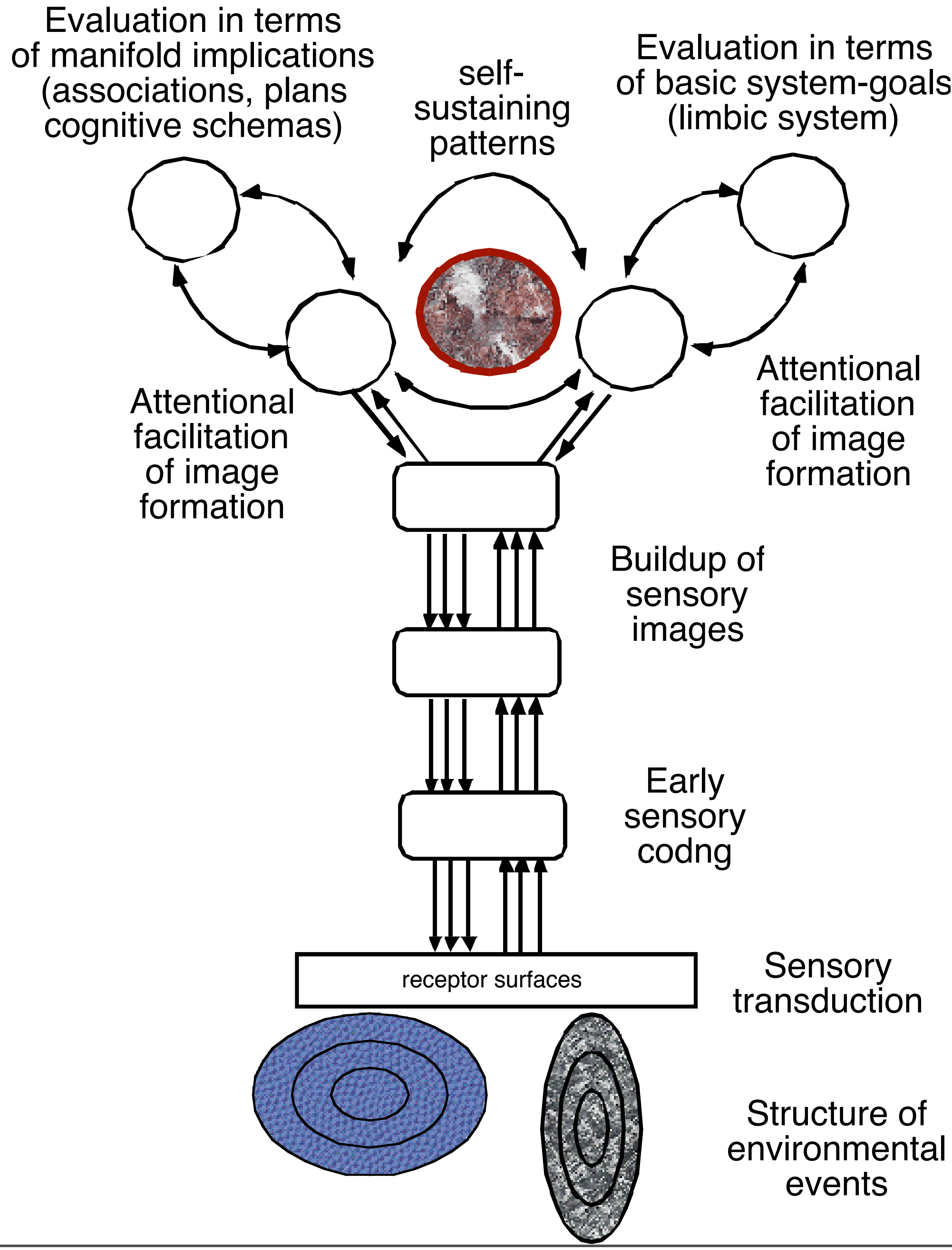


Figure by MIT OpenCourseWare. After Tramo, M. *Science* 291, no. 5501 (2001): 54-56.

Functional organization of the perceptual side



Two figures removed due to copyright restrictions.
Fig 2.8, input projections to the cochlear nucleus; and Fig 7.8, pathways from auditory cortex to cochlea.
In *The Central Auditory System*. Edited by G. Ehret and R. Romand. New York, NY: Oxford University Press, 1997.

Some generalities about the auditory system

- Rough cochleotopy is found at all levels, but not necessarily in all neural populations
- Highly ordered tonotopic maps exist only at low tone levels, near neural thresholds
- As one ascends the afferent pathway:
- Numbers of neurons at each level increases
- Fine timing information exists in great superabundance in lower stations, but becomes successively sparser
- Firing rates (spontaneous & driven) decline
- Inhibition increases; % nonmontonic rate-level fns incr.
- Diversity and complexity of response increases
- History-dependence and contextual effects increase
- Some modulation tuning that suc. declines in periodicity

  Typical BMFs: AN: 200-300 Hz; IC: 50-100 Hz; Ctx (< 16 Hz)
- No clear "pitch detectors" (Schwarz & Tomlinson, 1991)
- No narrow (BW < 0.3 octaves) "frequency channels" for BFs < 2 kHz

Basic problems to be solved

• "Hyperacuity problem"

• Account for the precision of pitch discriminations given the relatively coarse tunings of auditory neurons (at all levels), especially lower-frequency ones (BFs < 2 kHz)

• "Dynamic range problem"

• Account for the ability of listeners to discriminate small fractional changes ($\Delta I/I$) in intensity over a large dynamic range, and especially at high SPLs, where the vast majority of firing rates are saturated.

• "Level-invariance problem"

• Account for the invariance (and precision) of auditory percepts over large dynamic ranges given the profound changes in neural response patterns that occur over those ranges (rate saturation, rate non-montonicities).

•Pitch equivalence

•Account for the ability to precisely match pitches of pure and complex tones (pitch equivalence, metamery) given differences in spectra and under conditions where stimulus intensities are roved 20 dB or more

•Relative nature of pitch & transpositional invariance

• Account for the ability to precisely match pitches an octave apart (and/or to recognize patterns of pitch sequences) in the absence of an ability to identify absolute frequencies/periodicities. Account for ability to recognize transposed melodies as similar.

HST.725 Music Perception and Cognition
Spring 2009