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

Neural representation of musical pitch

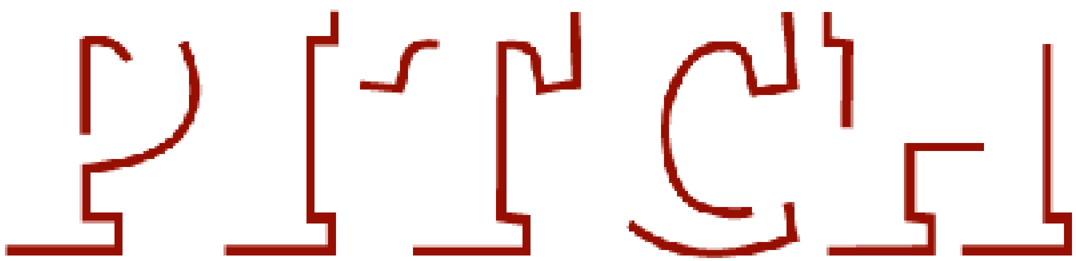


Figure by MIT OpenCourseWare.

www.cariani.com

Pitch: the basis of musical tonality

- Operational definition of pitch
- Pitch of pure tones
- Pitch of harmonic complexes at the fundamental
- Pitch of the missing fundamental
- Pitch of unresolved harmonics
- Repetition pitch
- Pitch salience
- Relative vs. absolute pitch
- Pitch circularity
- Pitch: how many dimensions do we need?
- Models of pitch: Place vs. temporal theories
- Envelope vs. fine structure (Schouten & de Boer)
  - Residue theories (incomplete filtering of unresolved harmonics)
  - Temporal autocorrelation theories (Licklider)
- Analytical (Helmholtz) vs. Gestalt (Stumpf) perspectives
  - Spectral pattern analysis/completion vs. dominant periodicity

Dimensions of auditory objects

Auditory qualities and their organization

Objects: Quasi-stationary assemblages of qualities

Pitch

Duration

Loudness

Timbre

Location

Spatial Dimensions

FUSION/SEPARATION

Common onset & harmonic structure => fusion

Different F0s, locations, onset => separation

POLYPHONY

Dimensions of event perception

Unitary events & their organization

Events: abrupt perceptual discontinuities

TEMPORAL EVENT STRUCTURE

Timing & order (metric, sequence)

FUSION/SEPARATION

Common onset, offset => fusion

Diff. meters, pitch, timbre => separation

STREAMS, POLYRHYTHMS

A few words about Loudness

- **Loudness is the perceptual attribute that covaries with the intensity of sounds (loudness is the subjective attribute, intensity is the physical, acoustical property)**

- **We mentioned that the auditory system has a huge dynamic range, over a factor of 100,000 between the sound pressure level of the softest and the loudest sounds.**

- **Loudness is important in music for several reasons**
  - **Listening level (louder music is more salient, captures attention)**
  - **Onsets and accents (loudness contrast accents notes)**
  - **Dynamics (changes in loudness communicate tension, relaxation)**
  - **Safety issues (listening to music at high levels (>100 dB SPL) for prolonged periods of time will damage your ears and impair your ability to hear music**

Typical sound levels in music

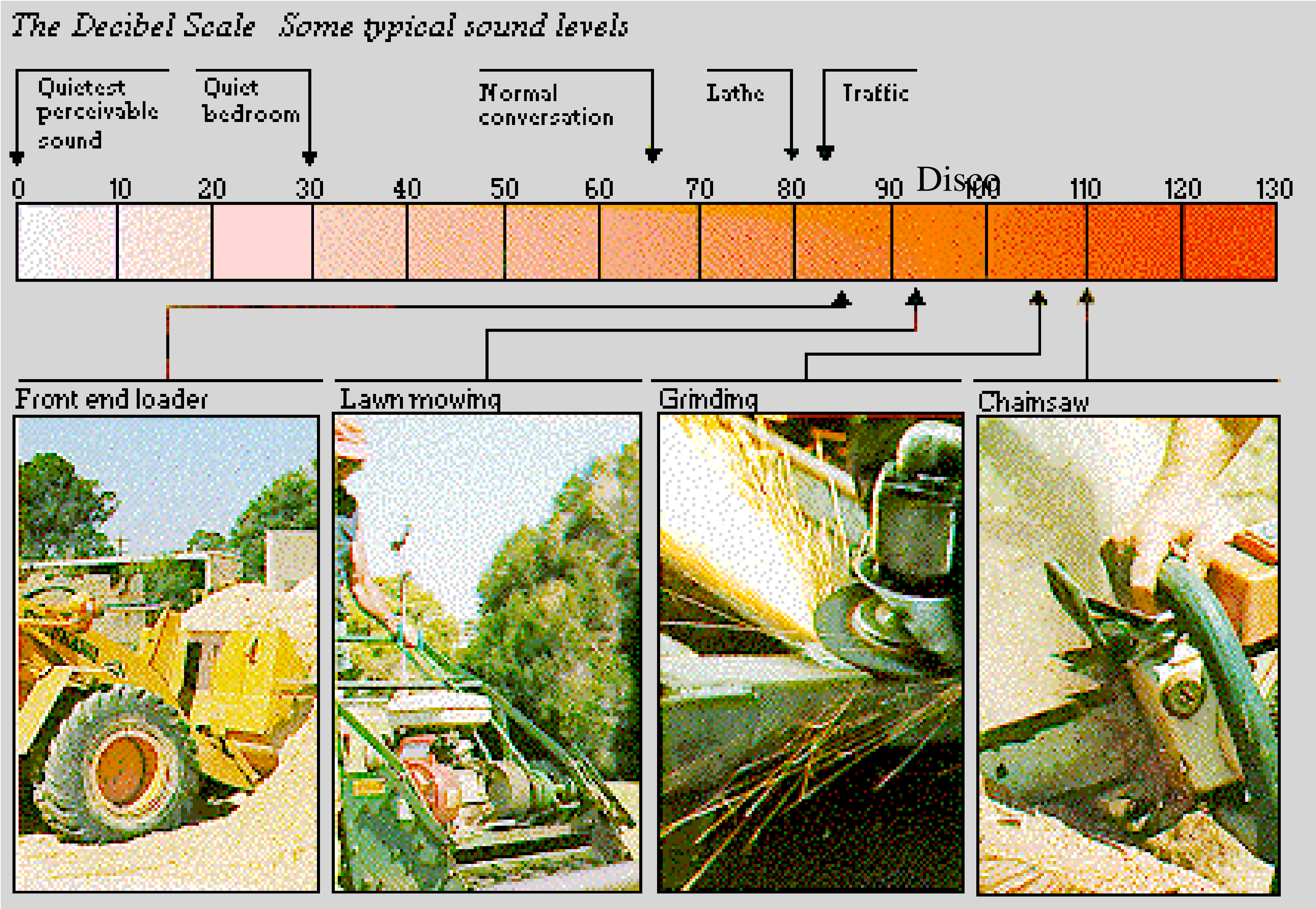
On origins of music dynamics notation
http://www.wikipedia.org/wiki/Pianissimo

Text removed due to copyright restrictions. See the Wikipedia article.

- Pain > 130 dB SPL
- Loud rock concert 120 dB SPL
- Loud disco 110 dB SPL
- *fff* 100 dB SPL
- *f* *(forte, strong)* 80 dB SPL
- *p* *(piano, soft)* 60 dB SPL
- *ppp* 40 dB SPL

- Lower limit
- Theshold of hearing 0 dB SPL

Typical sound pressure levels in everyday life



Courtsey of WorkSafe, Department of Consumer and Employment Protection, Western Australia (http://www.safetyline.wa.gov.au).

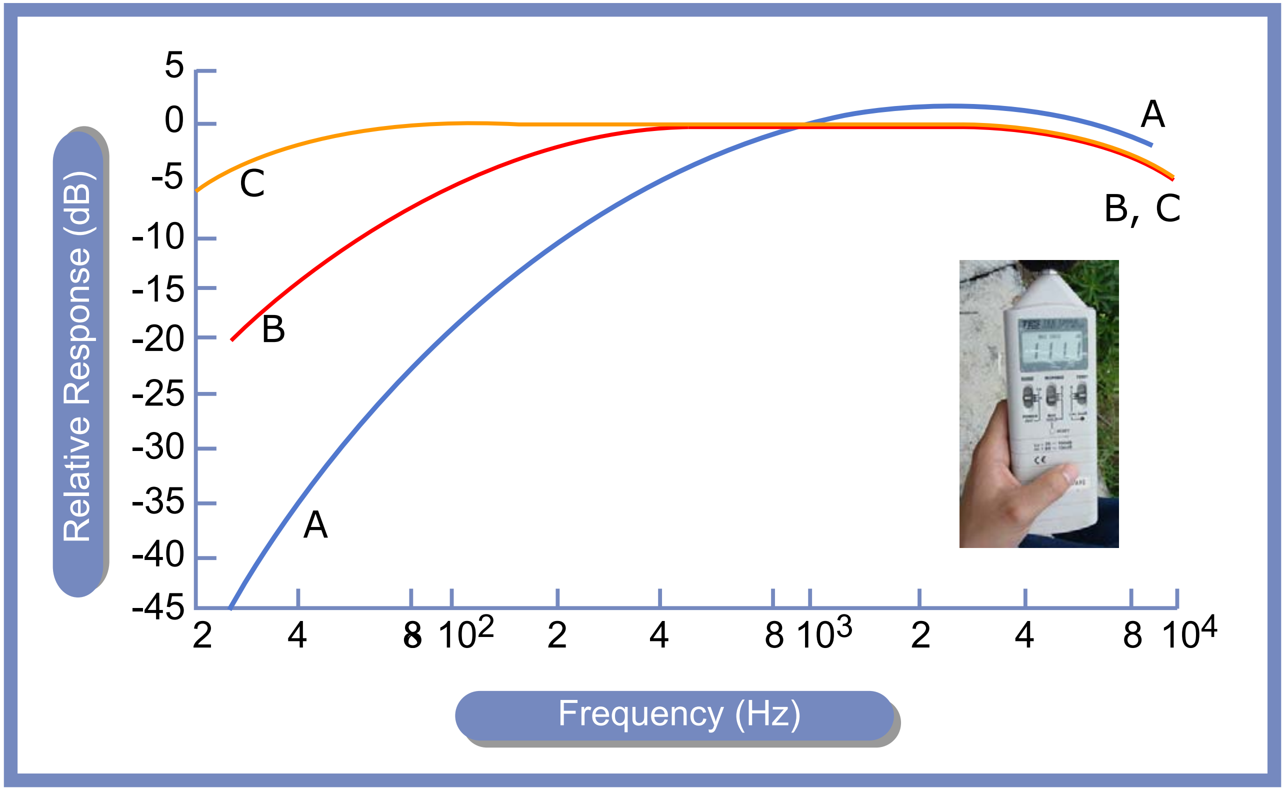
http://www.safetyline.wa.gov.au/institute/level2/course18/lecture54/l54_03.asp

Sound level meters and frequency weightings

A: based on human equal loudness contours @ ~40 dB SL

Fletcher-Munson curves (recently revised)

C: flat-weighting



Graph by MIT OpenCourseWare. SPL meter photo courtesy of EpicFireworks on Flickr.

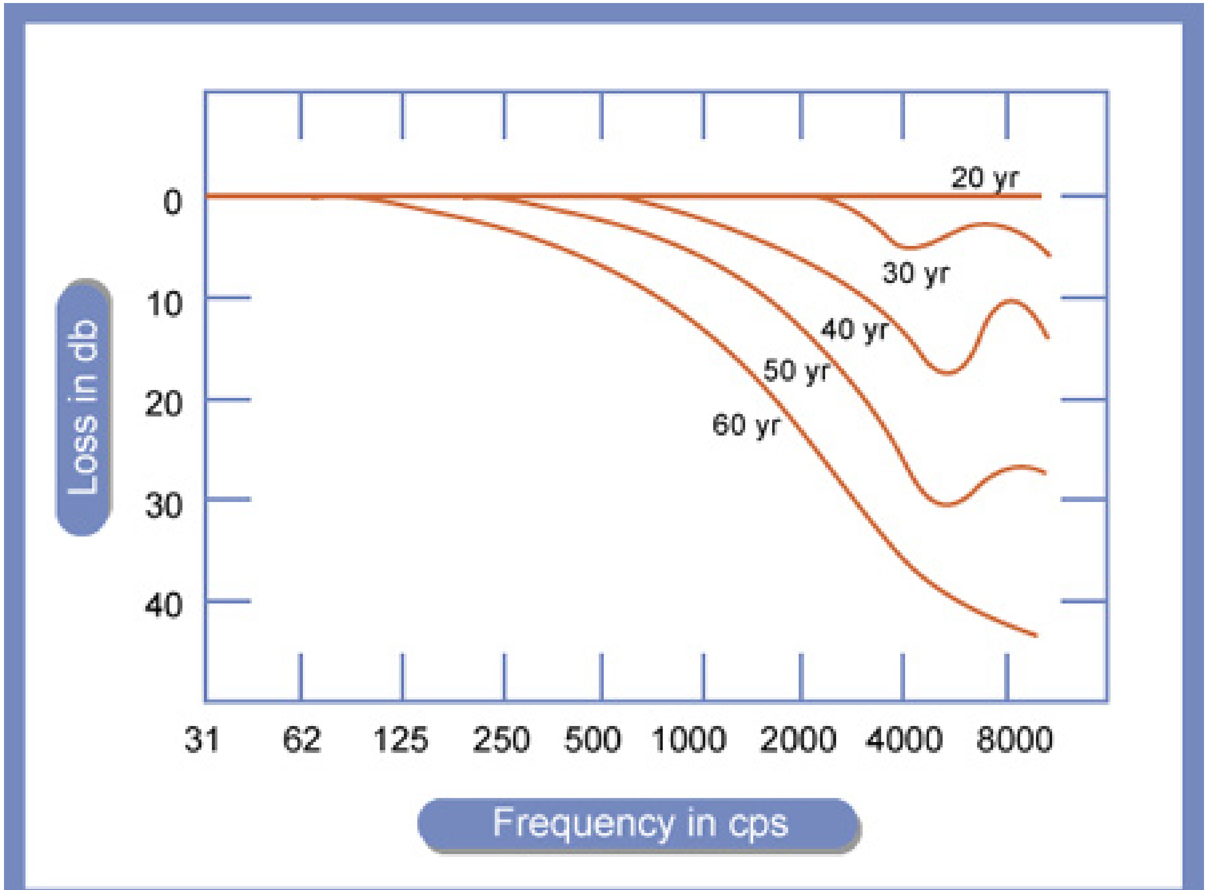
Dynamic range of some musical instruments

Images removed due to copyright restrictions.

Graphs of relative intensity vs. pitch for different instruments: violin, double bass, flute, B-flat clarinet, trumpet, french horn.
Figure 8.5 in Pierce, J. R. *The Science of Musical Sound*. Revised ed. New York, NY: W.H. Freeman & Co., 1992. ISBN: 9780716760054.

Noise floor: ~ 45 dB SPL
Conversation: 60 dB SPL
Symphony: 80-90 dB SPL
Disco: 100 dB SPL

Hearing loss with age (overexposure to loud sounds accelerates this process)



Progressive loss of sensitivity at high frequencies with increasing age. The audiogram at 20 years of age is taken as a basis of comparison.

(From Morgan, 1943, after Bunch, 1929.)

Figure by MIT OpenCourseWare.

Steps to prevent hearing damage


Photo courtesy of KarenD on Flickr.

Use earplugs (reduce levels by 20-30 dB)

Take measures in recurring situations where you experience ringing in your ears (concerts, discos)

Impulsive loud sounds are worst (gunshots, hammering)

Be wary of cranking up the level in cars, especially when the windows are down (if it sounds terribly loud when you're stopped at a light, this should tell you something)

With personal sound players (MP3s, iPods, walkman),
- always set the listening level in quiet
- don't crank up the level in noisy situations
- use the volume limiter feature (set this in quiet)
- if you listen in noisy situations (mowing the lawn), then by all means use noise-cancellation headphones

Pitch in music

- Tonal music is based in large part on pitch relations
- Sequences of pitches constitute melodies
- Relations between combinations of pitches constitute harmonies
- Sets of pitches make up musical scales, which are the perceptual atoms of musical tonality
- Musical pitch is relative pitch (transpositional invariance)
- We will discuss absolute pitch later in the course

Operational definition of "pitch"

- **Pitch is that auditory quality that varies with the periodicity and/or frequency of sounds.**
- **(i.e. not loudness, duration, location, or timbre)**

- **Operationally, pitch is defined as the frequency of a pure tone to which a sound is matched.**

- **Since pure tone pitch changes very slightly (0-4%) with large changes in sound pressure level (40 dB), the level of the reference tone also has to be specified.**

- **For musical sounds (complex tones), this much celebrated dependence of pitch on level is quite minimal (< 1% over 40 dB). The more harmonics, the smaller the effect.**

Pitch metamery (perceptual equivalence)

Sounds with different frequency spectra can produce the same pitch

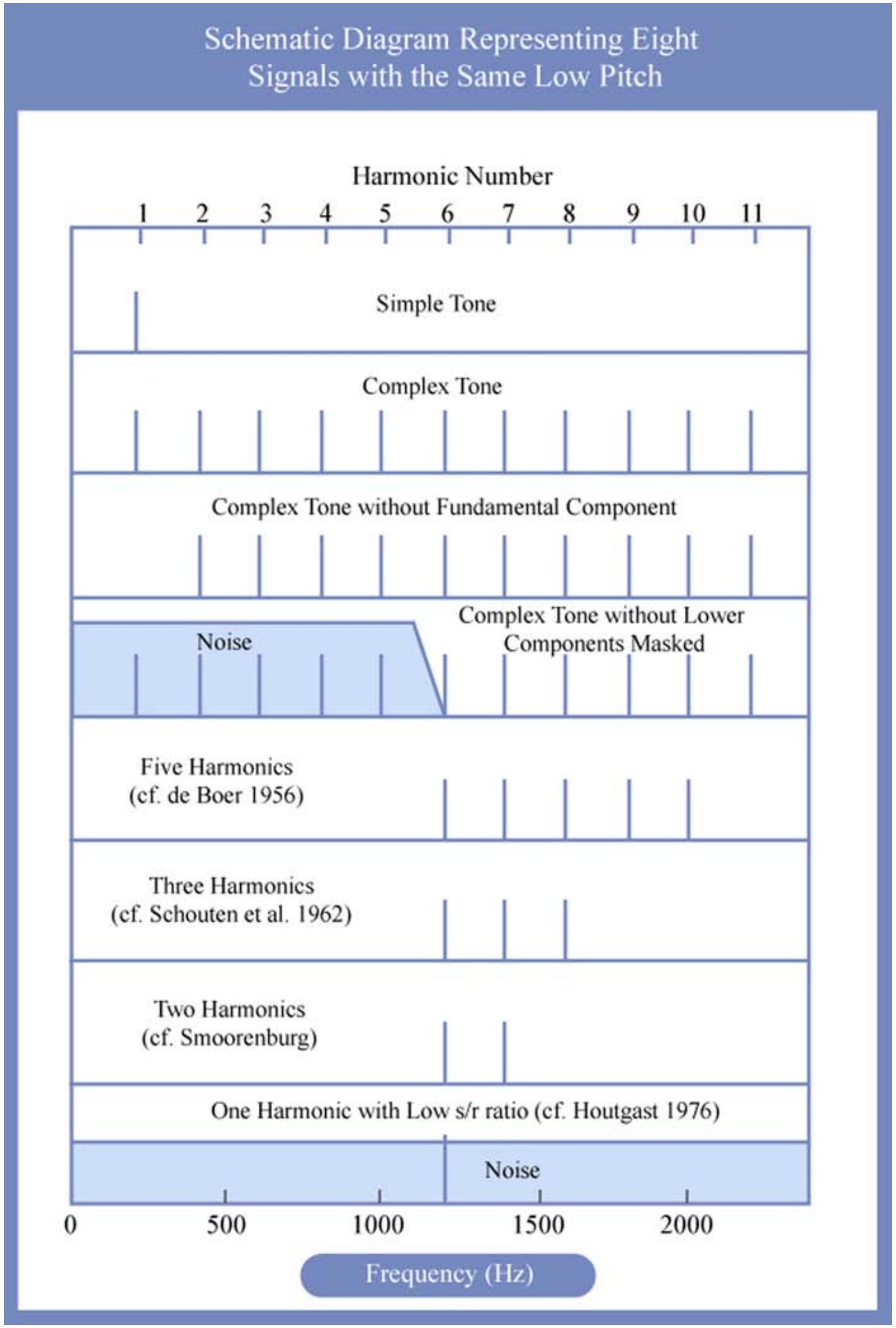


Figure by MIT OpenCourseWare.

Range of pitches of pure & complex tones

- **Pure tone pitches**
  - Range of hearing (30-20,000 Hz)
  - Range in tonal music (100-4000 Hz)
  - Pitches of individual partials in a complex, "analytical" pitch
- **Most (tonal) musical instruments produce harmonic complexes that evoke pitches at their fundamentals (F0’s)**
  - Called virtual pitch, periodicity pitch, low pitch
  - Range of F0’s in tonal music (30-4000 Hz)
  - Range of missing fundamental (30-1200 Hz)

Music spectrograms

Pitch is not simply frequency

Music pitches are not pure tones --
They are mostly harmonic complex tones

The pitch that is heard for a harmonic complex tone
corresponds to the fundamental frequency of
the tone (with very few exceptions)

“Virtual” pitch: F0-pitch as pattern completion

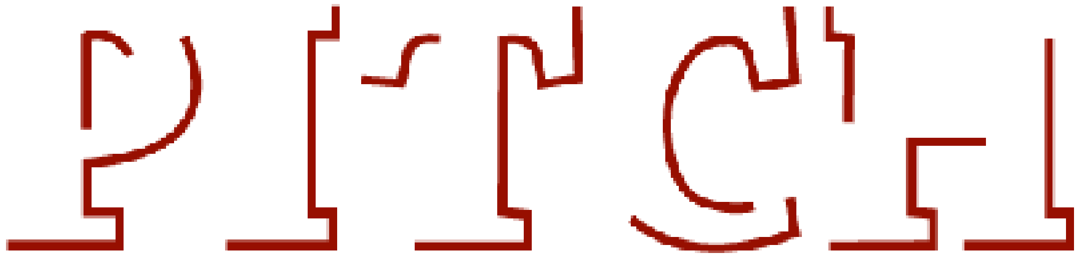


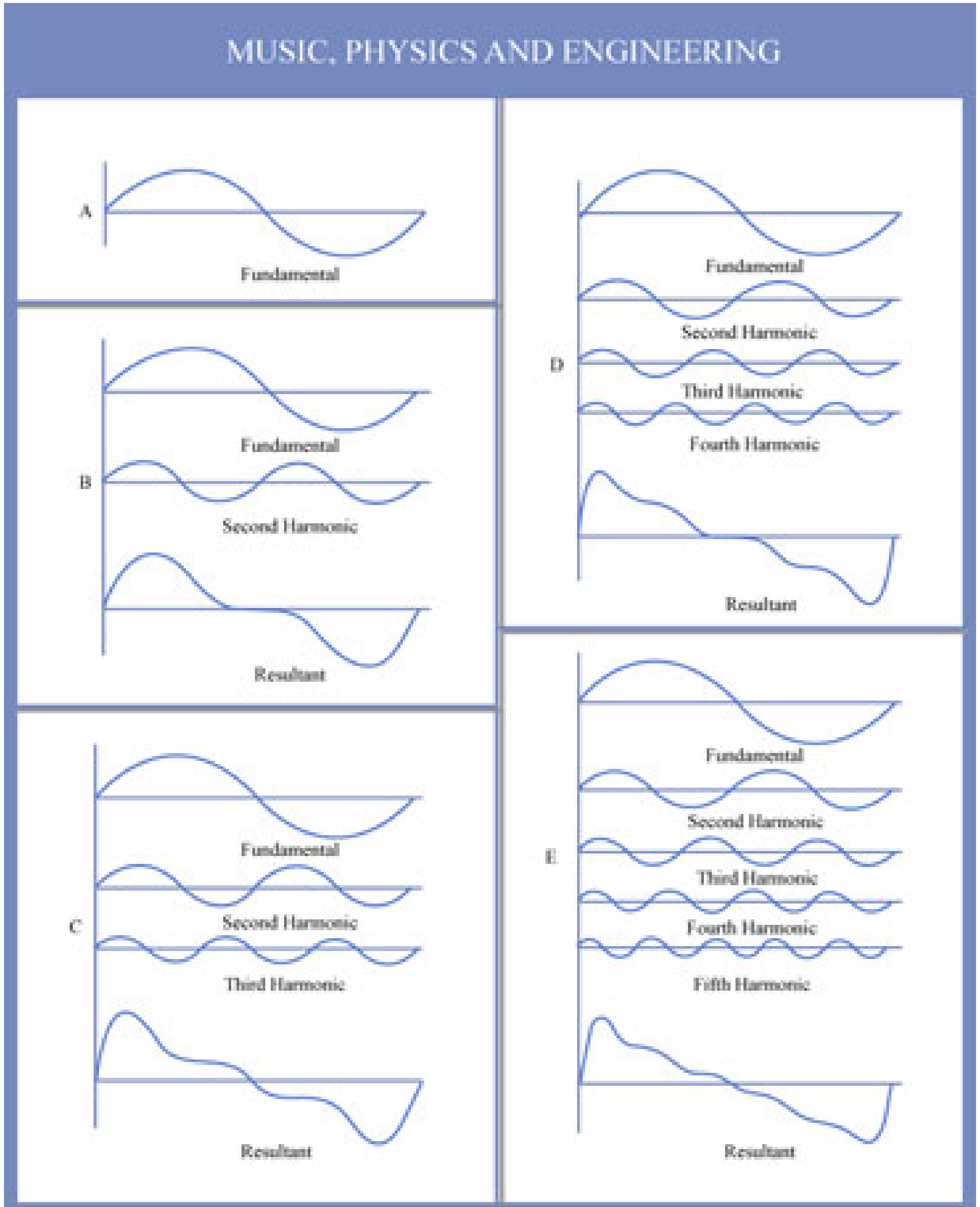
Figure by MIT OpenCourseWare.

Figure 1. Visual analogy of virtual pitch. The visual system perceives contours which are not physically present (Terhardt, 1974).

“Missing fundamental“ analogy to illusory contour

Fundamentals and harmonics

- Periodic sounds (30-20kHz) produce pitch sensations.
- Periodic sounds consist of repeating time patterns.
- The fundamental period (F0) is the duration of the repeated pattern.
- The fundamental frequency is the repetition frequency of the pattern.
- In the Fourier domain, the frequency components of a periodic sound are all members of a harmonic series (n = 1*F0, 2*F0, 3*F0...).
- The fundamental frequency is therefore the greatest common divisor of all of the component frequencies.
- The fundamental is also therefore a subharmonic of all component frequencies.



A. Fundamental sine wave.
B. Fundamental, second harmonic, and the resultant.
C. Fundamental, second and third harmonics, and the resultant.
D. Fundamental, second, third, and fourth harmonics, and the resultant.
E. Fundamental, second, third, fourth, and fifth harmonics, and the resultant.

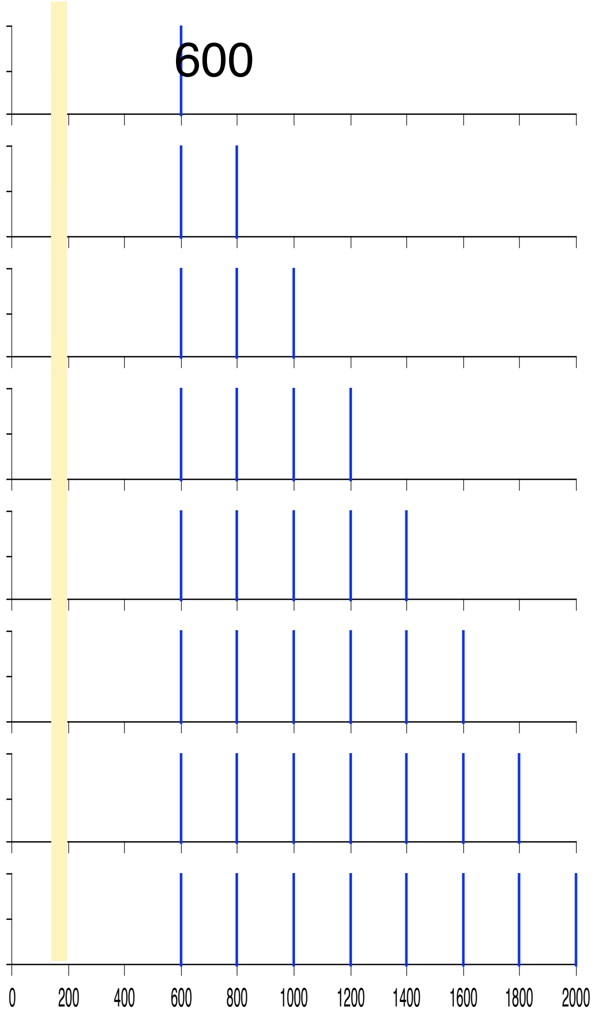
Figure by MIT OpenCourseWare.

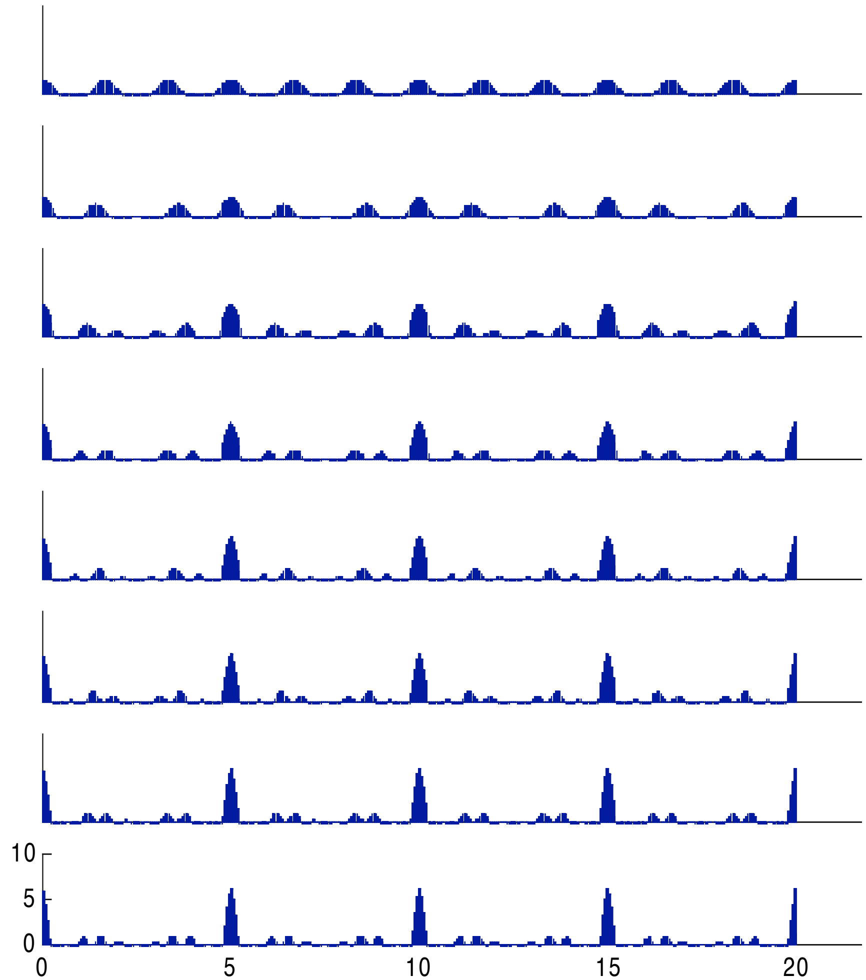
Emergent pitch

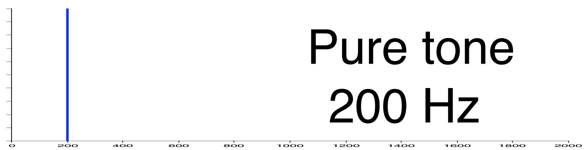
Missing F0

Line spectra

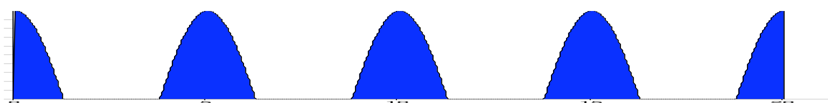
Autocorrelation (positive part)







Pure tone
200 Hz



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

Auditory system:
Frequency analyzer vs. Periodicity analyzer

Conceptual models of resonance:

Helmholz resonator
- band-pass filter -- one frequency
- frequency decomposition
- simple oscillator models

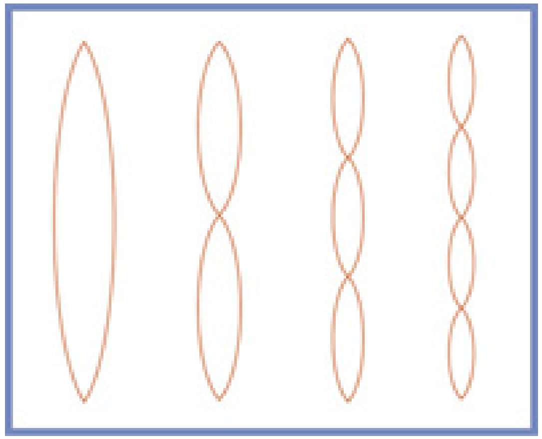


Figure by MIT OpenCourseWare.

String
- many harmonics
- comb filter

Complex oscillator - delay loop
- matched filter, complex pattern generator

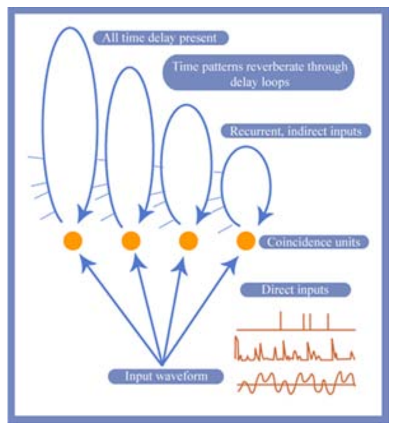
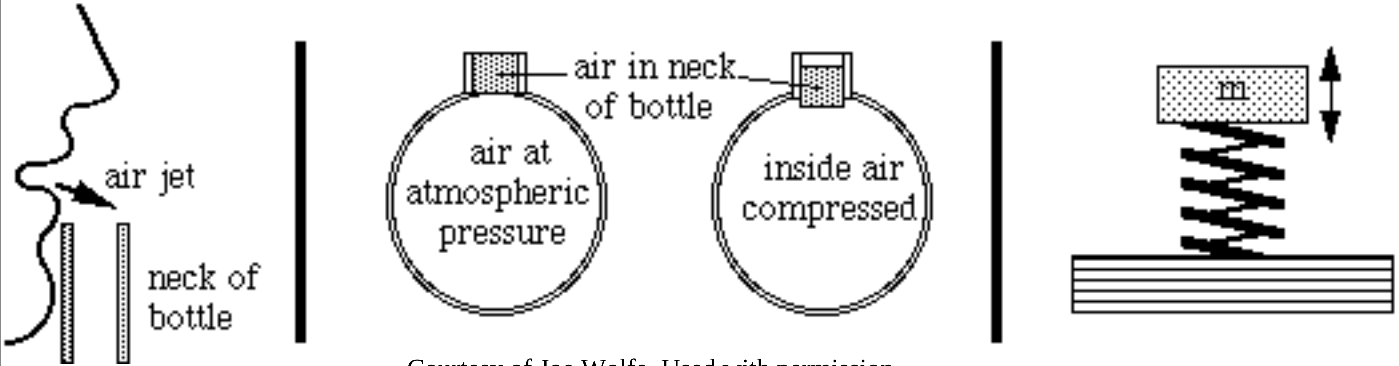


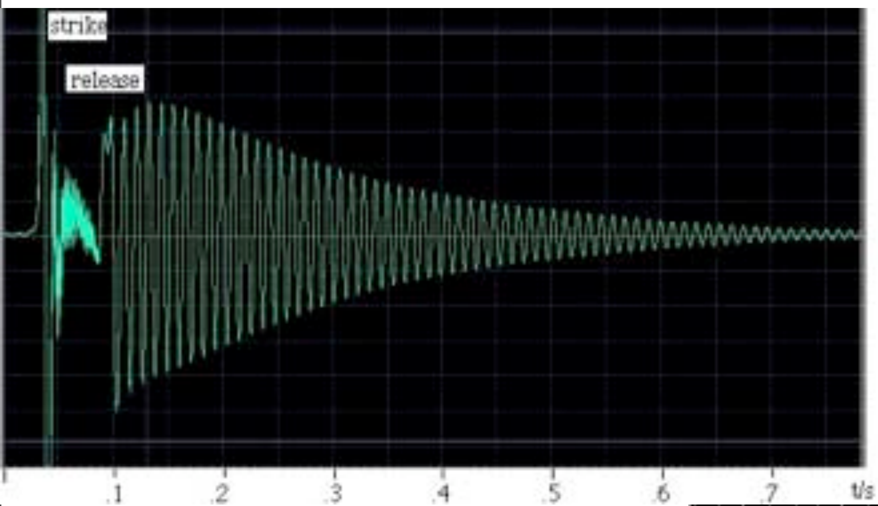
Figure by MIT OpenCourseWare.

Helmholtz resonator

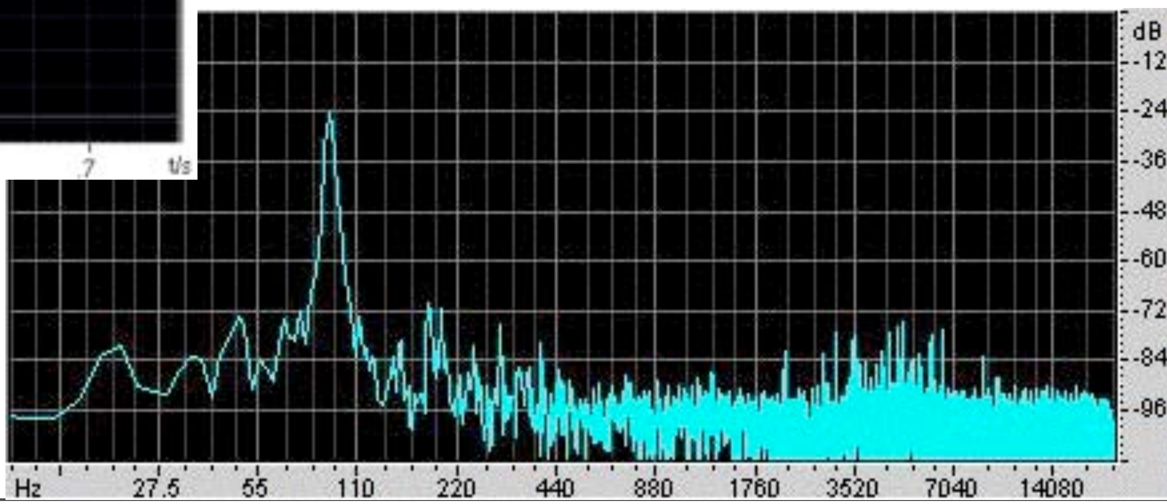
http://www.phys.unsw.edu.au/jw/Helmholtz.html



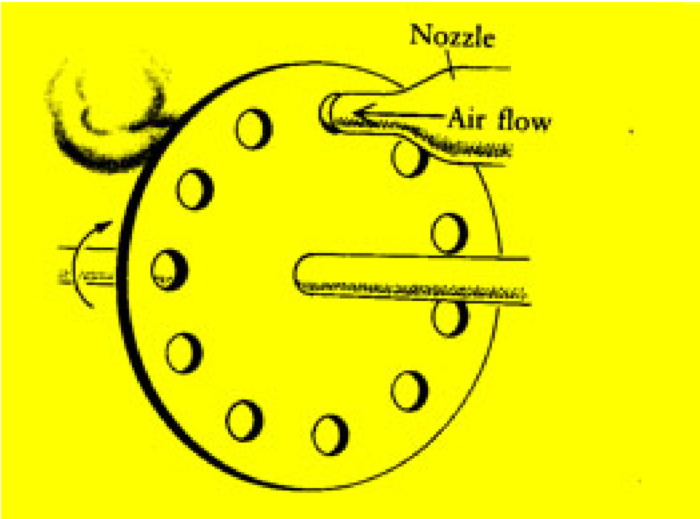
Courtesy of Joe Wolfe. Used with permission.
http://www.phys.unsw.edu.au/music



Courtesy of Joe Wolfe. Used with permissions.
http://www.phys.unsw.edu.au/music



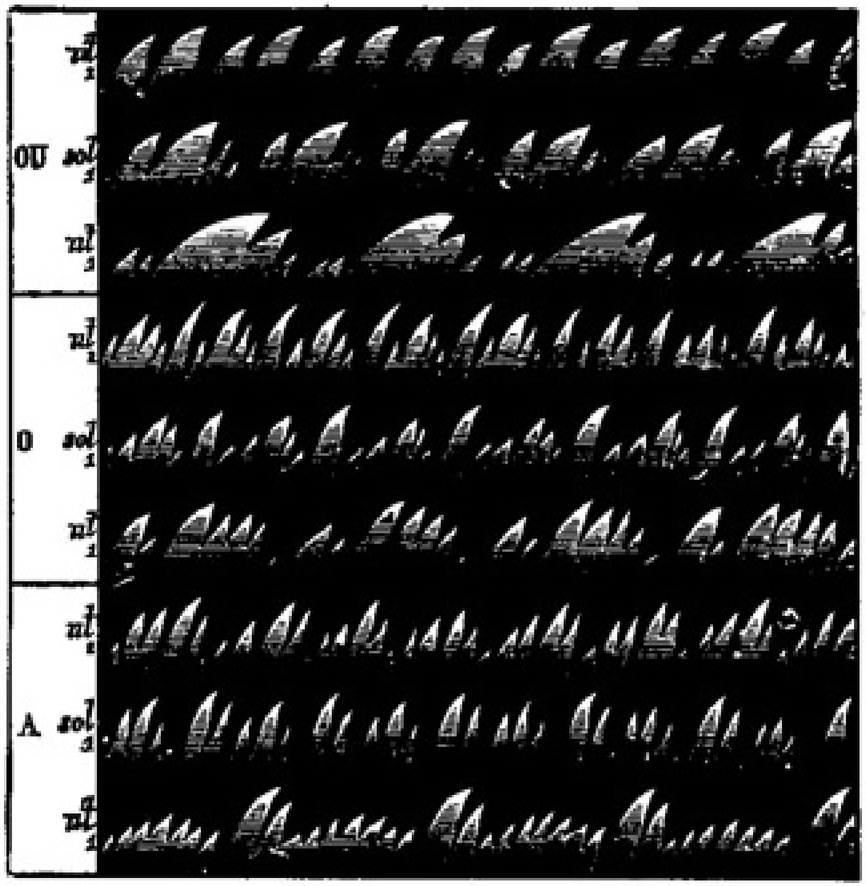
Sirens



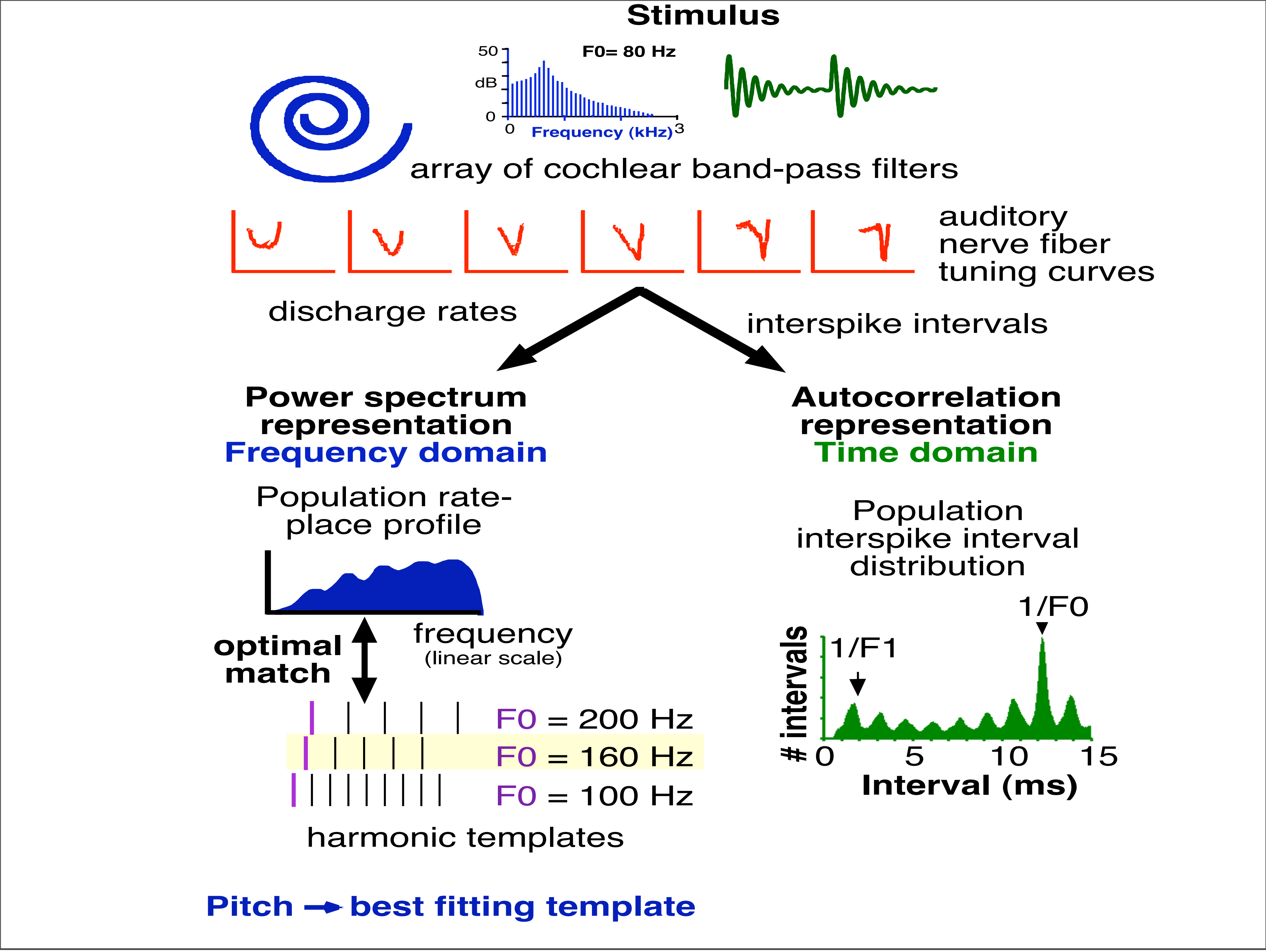
A simple siren is produced by forcing compressed air through equally spaced holes on a rotating disk. This produces a periodic vibration whose frequency equals the rate of holes passing by the air nozzle.

De la Tour's siren

Early sound analysis of vowels



In this 19th century apparatus developed by Koenig, waveforms were visualized by viewing a flame reflected on a rotating mirrored drum. Vowel sounds resulted in the same flame pattern regardless of their pitch level.


Stimulus
50
dB
0
0
Frequency (kHz)
3
F0= 80 Hz
array of cochlear band-pass filters
auditory nerve fiber tuning curves
discharge rates
interspike intervals
Power spectrum representation
Frequency domain
Population rate-place profile
frequency
(linear scale)
optimal match
F0 = 200 Hz
F0 = 160 Hz
F0 = 100 Hz
harmonic templates
Pitch → best fitting template
Autocorrelation representation
Time domain
Population interspike interval distribution
1/F0
1/F1
intervals
0
5
10
15
Interval (ms)

Correlograms

Images removed due to copyright restrictions.
See Figures 6.16A-D and 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.

Licklider virtual pitch masking demonstration (iTunes/Spectrograph)

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

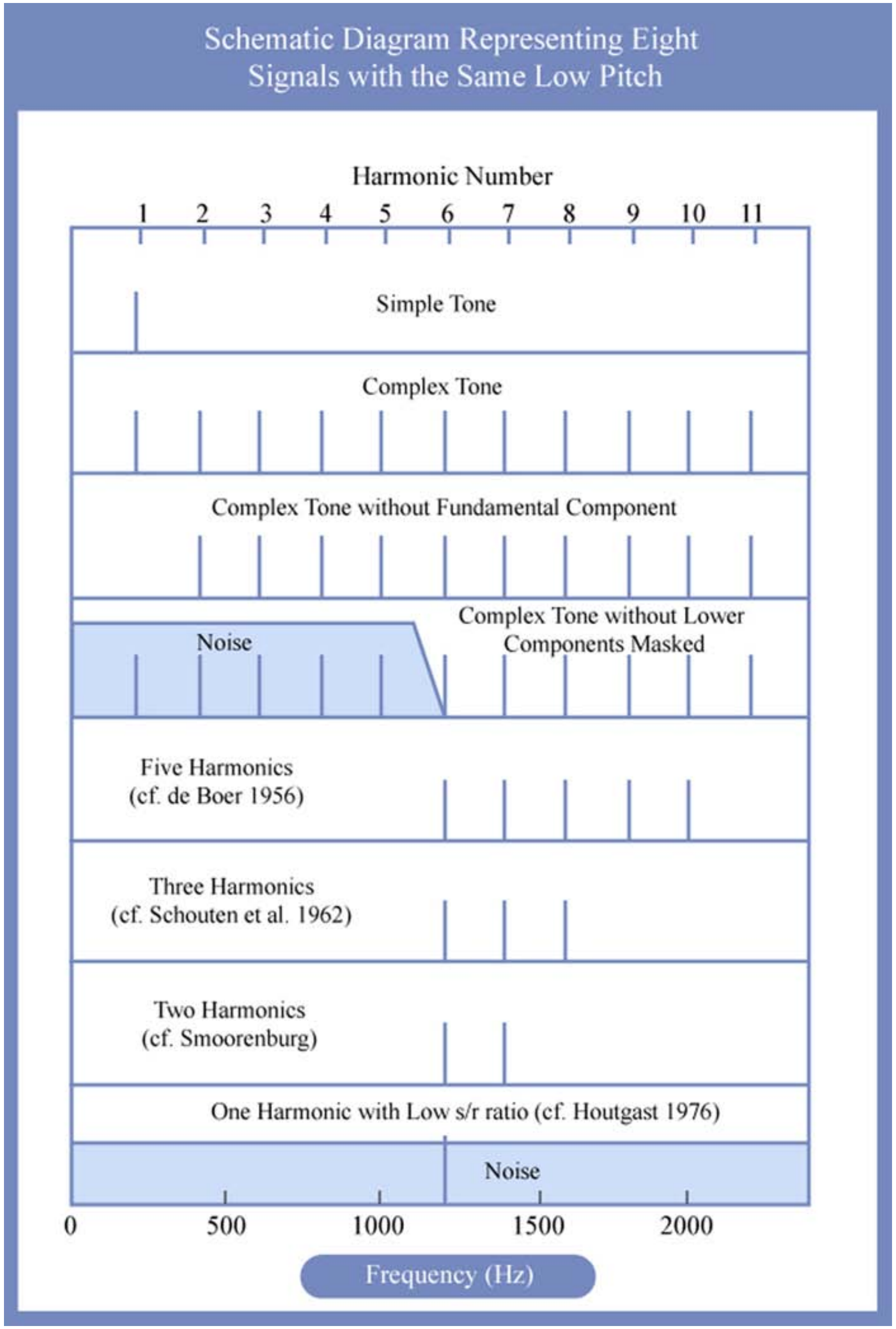


Figure by MIT OpenCourseWare.

As harmonic numbers increase, the missing fundamental gets weaker.

Pitch basics

- **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)

Pure tone pitch discrimination becomes markedly worse above 2 kHz

Weber fractions for frequency (Δf/f) increase 1-2 orders of magnitude between 2 kHz and 10 kHz

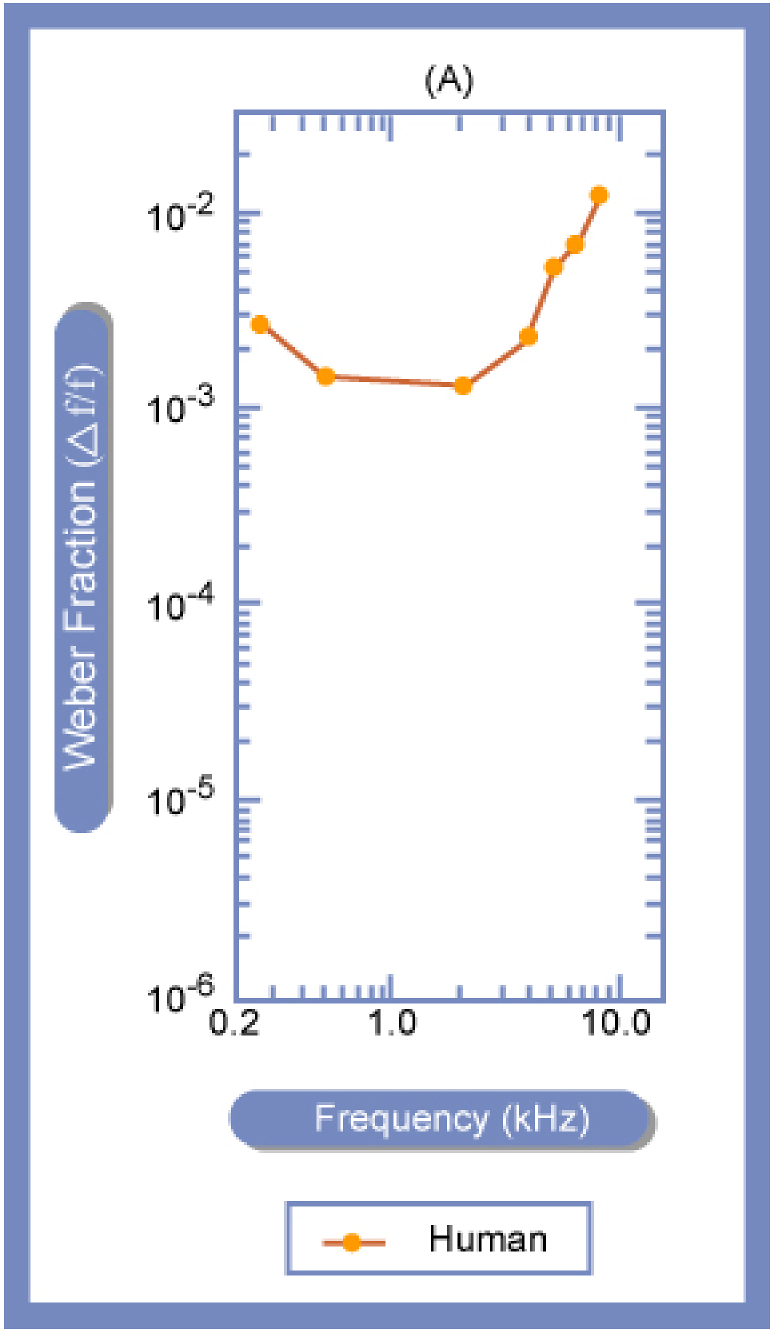


Figure by MIT OpenCourseWare.

JND's

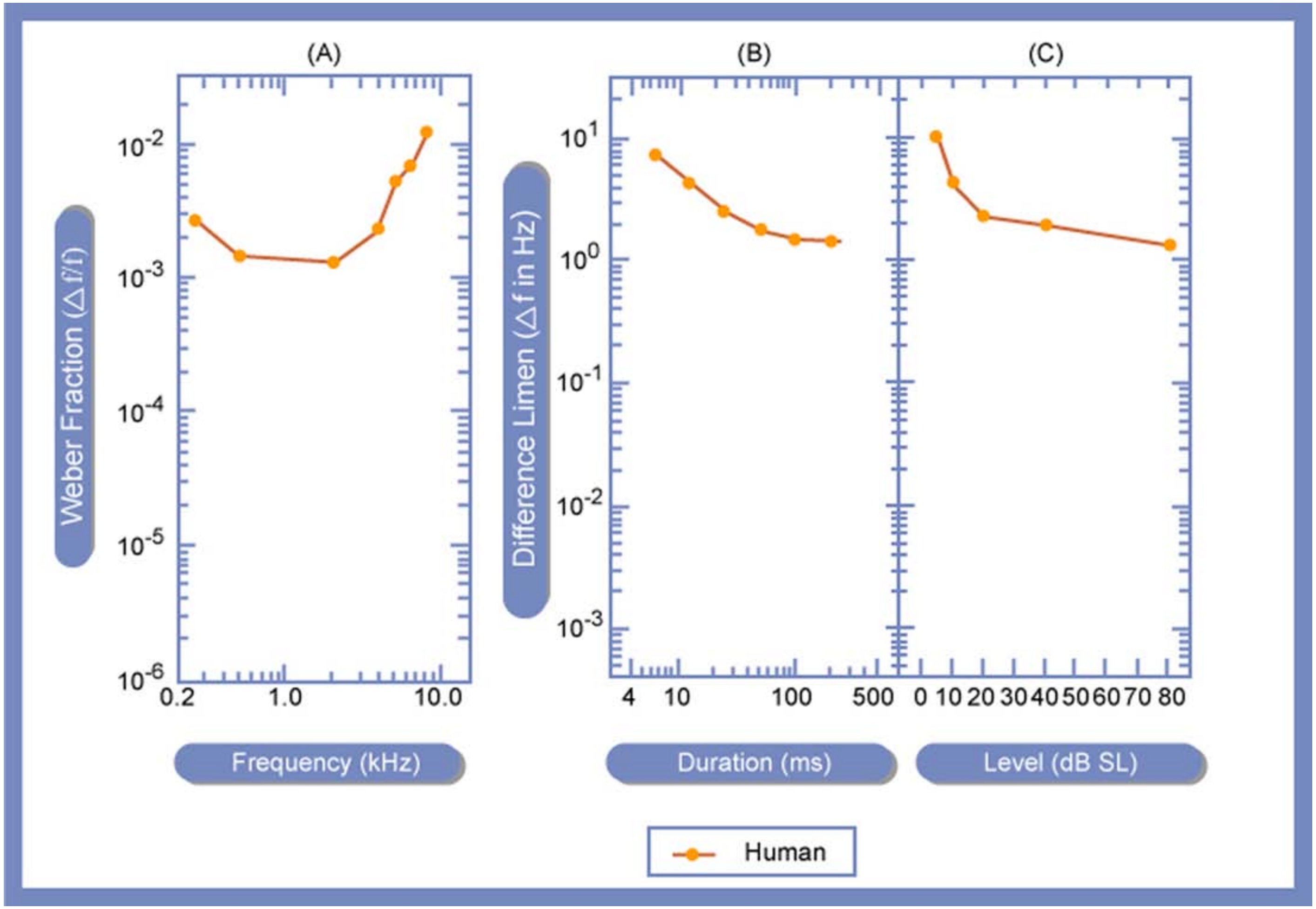


Figure by MIT OpenCourseWare.

Pure tone pitch discrimination improves

at longer tone durations

and

at higher sound pressure levels

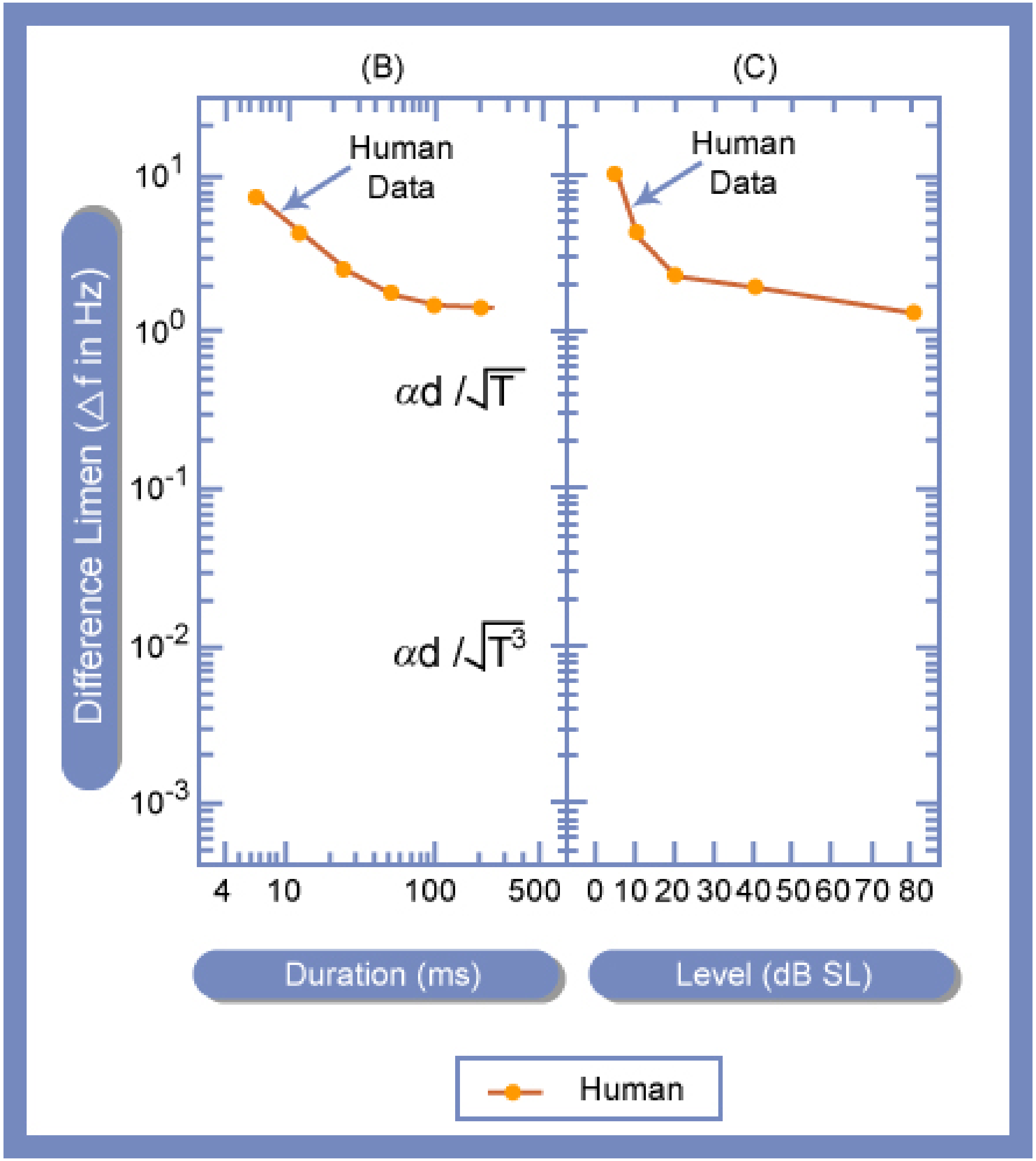


Figure by MIT OpenCourseWare.

"Pitchedness" as a function of sound duration

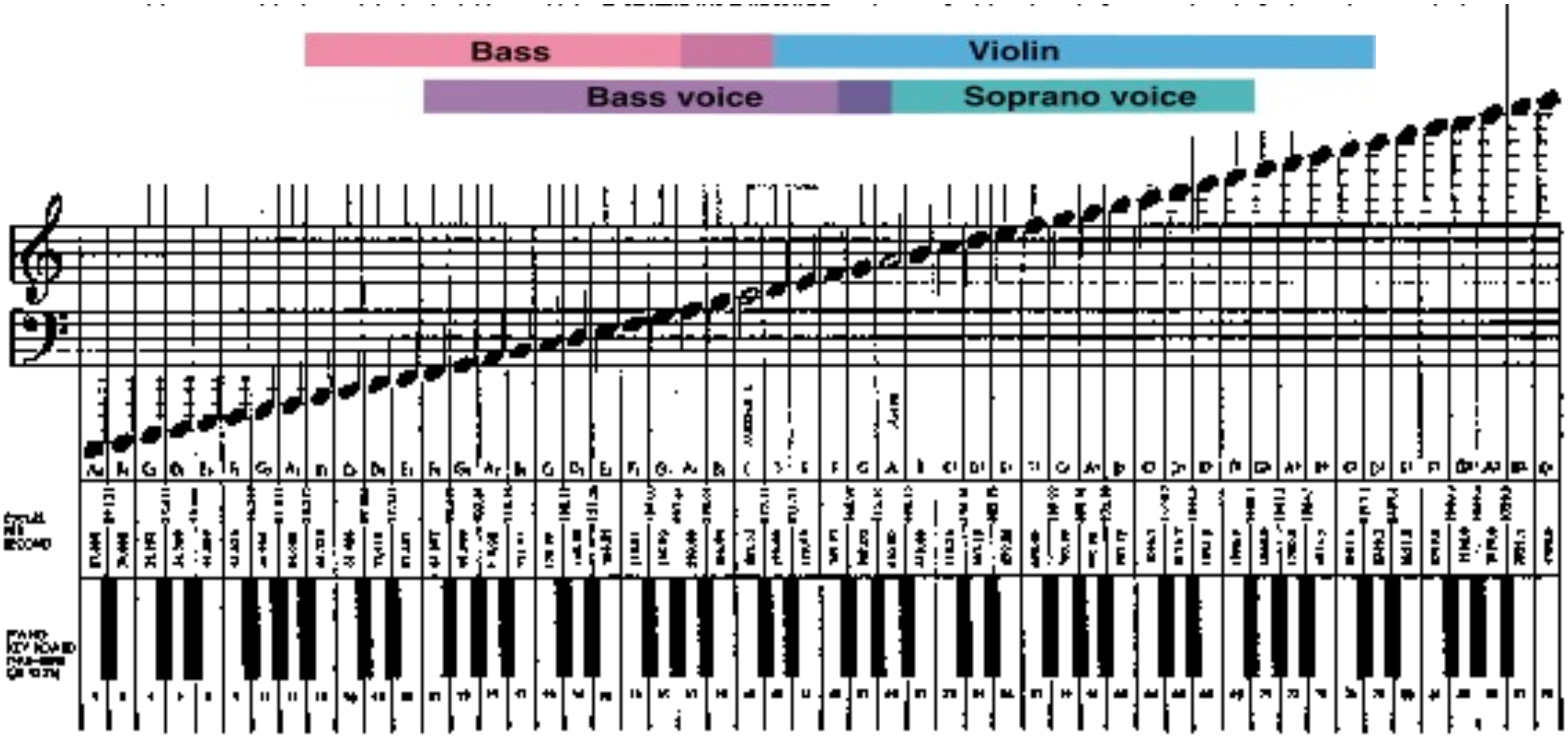
Graph removed due to copyright restrictions.
Figure 36, comparing "Tone pitch" and "click pitch" response. In Licklider, J. C. R. "Basic Correlates of the Auditory Stimulus." *Handbook of experimental psychology*. Edited by S. S. Stevens. Oxford, UK: Wiley, 1951. pp. 985-1039.

Frequency ranges of (tonal) musical instruments

8k
6
5
4
3
2

> 6 kHz

2.5-4 kHz



27 Hz

110 Hz

262 Hz

440 Hz

880 Hz

4 kHz

The neural coding problem in audition:

How does the brain represent and process acoustic patterns, such that we hear what we hear?

In particular, how does it represent periodic sounds, such that we hear pitches at the fundamentals of musical sounds?

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

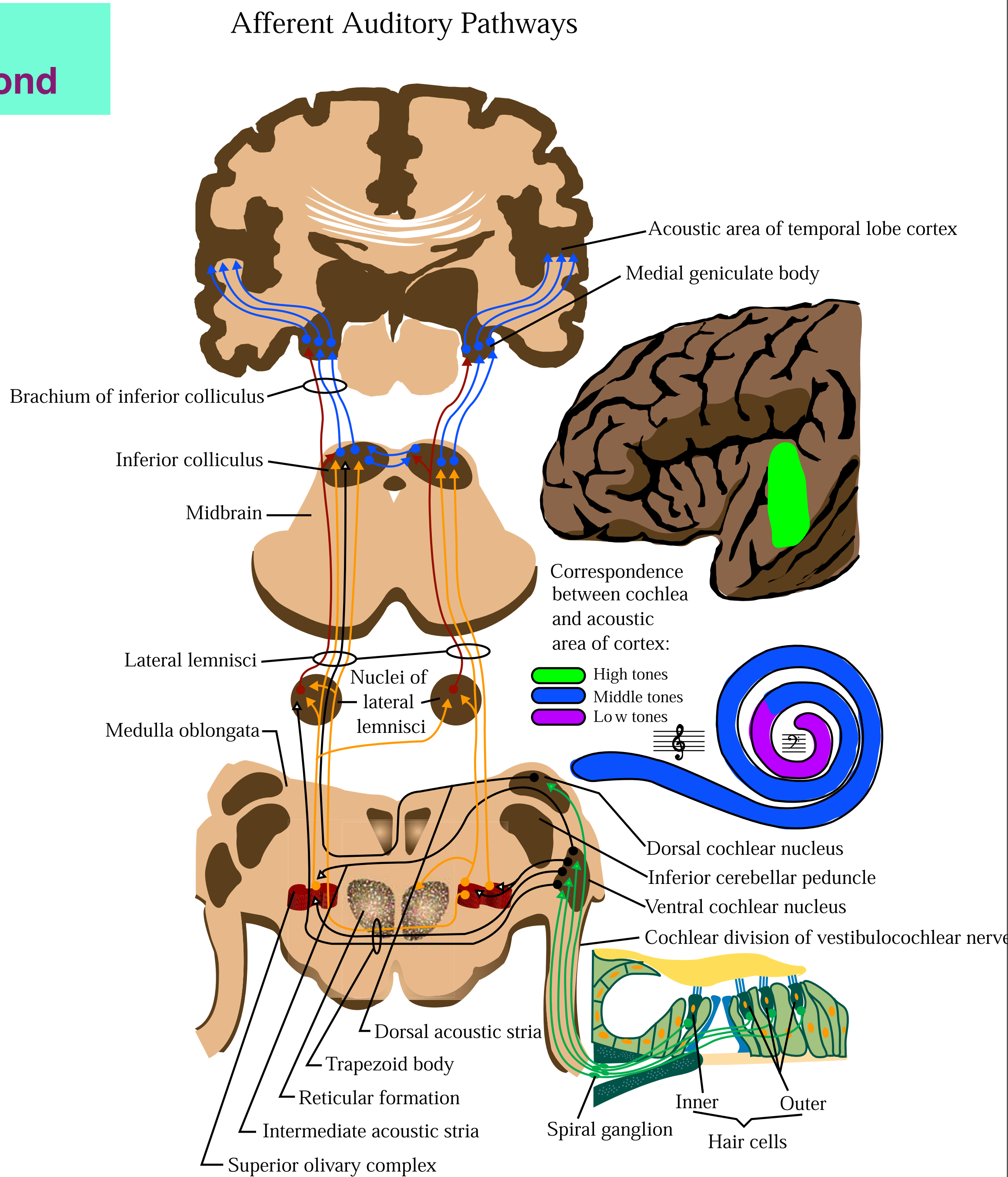
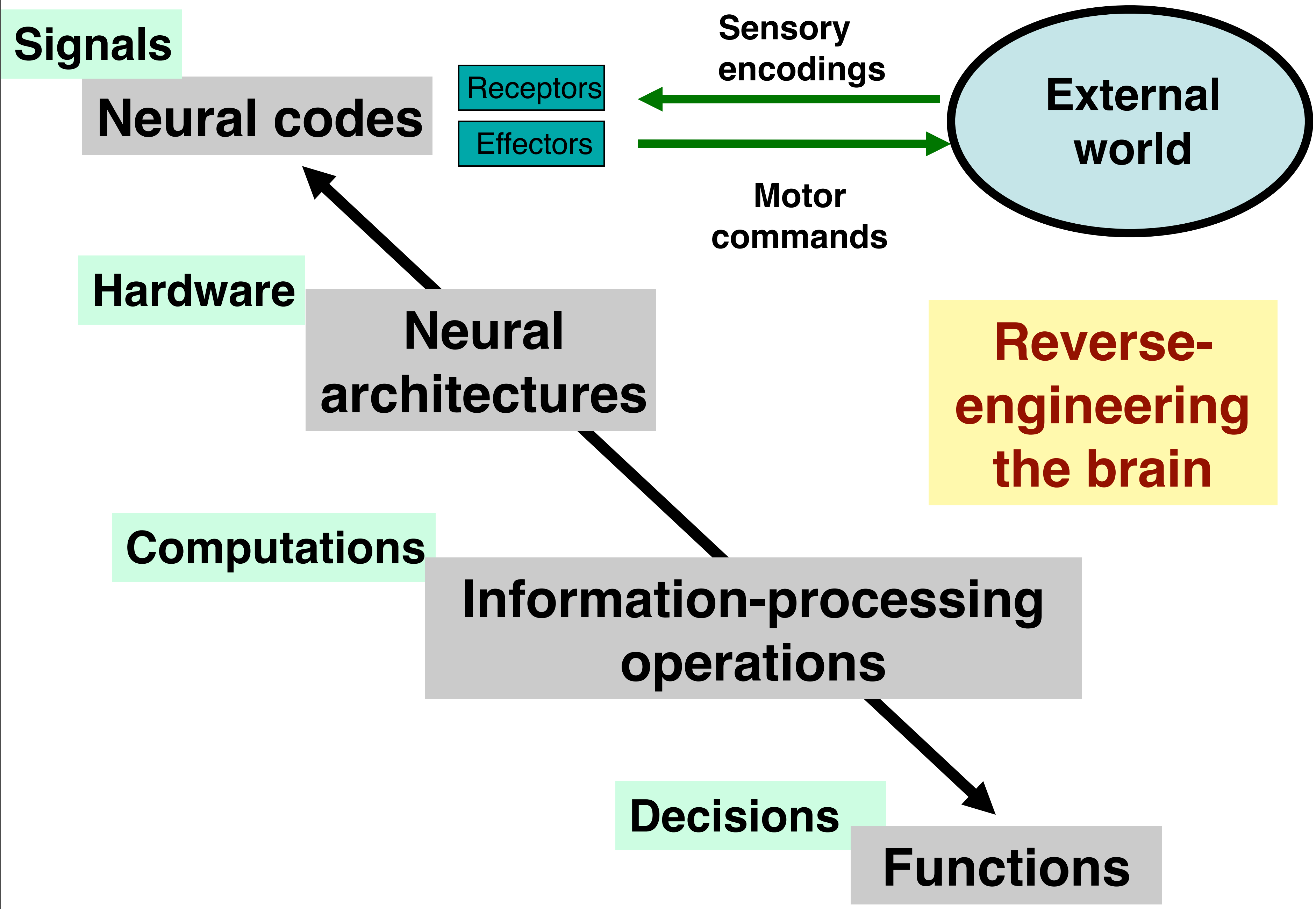


Figure by MIT OpenCourseWare.


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

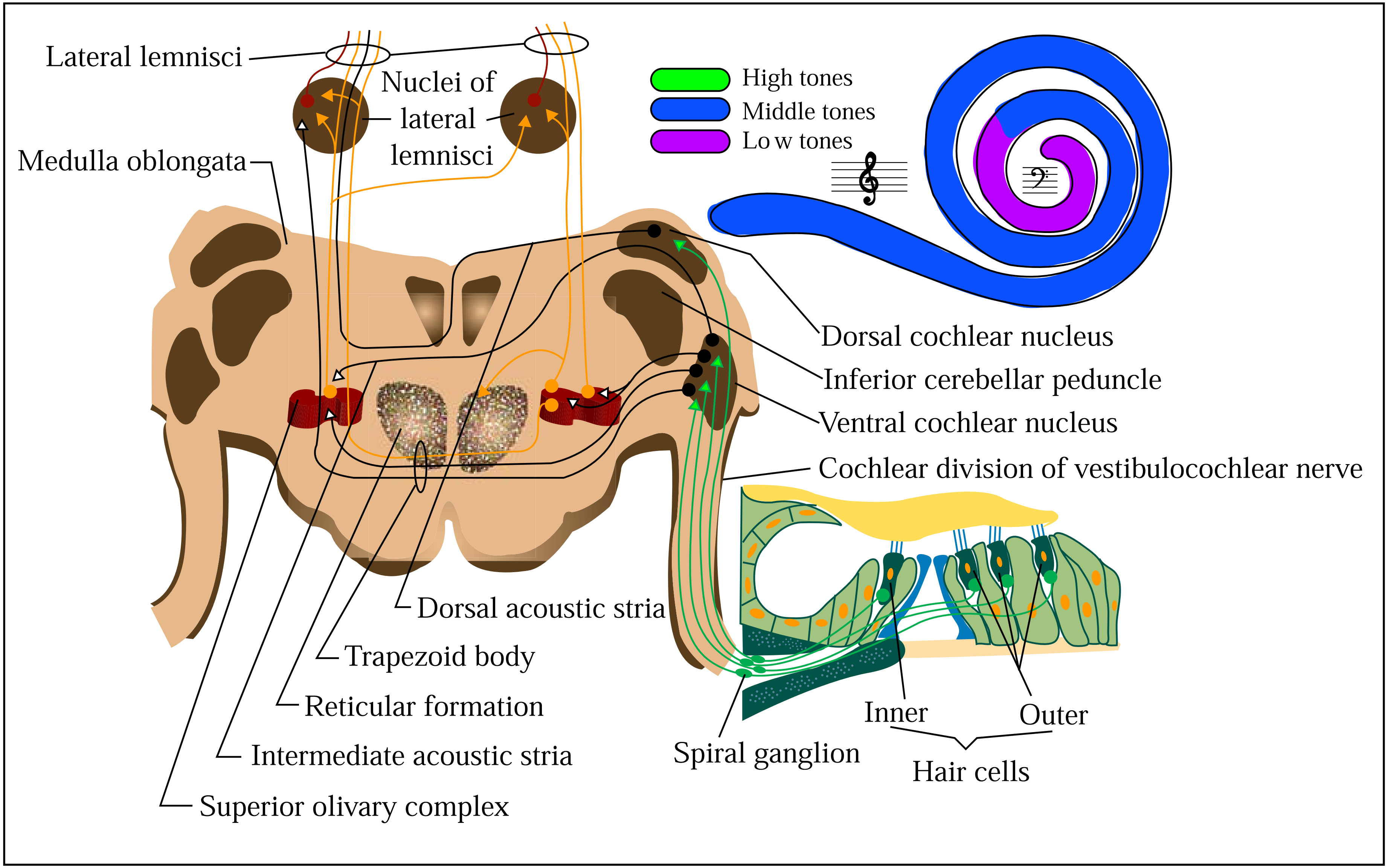
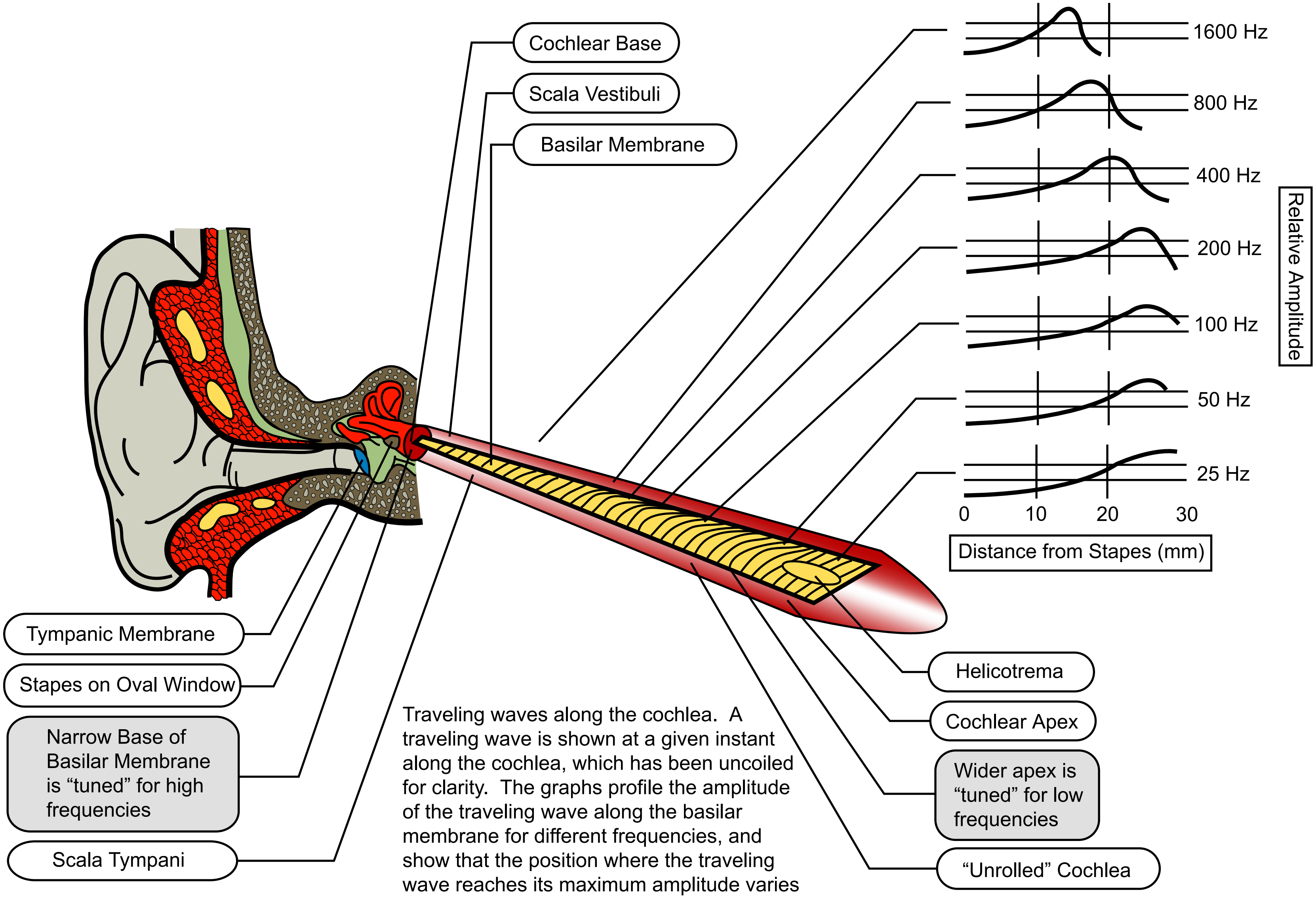


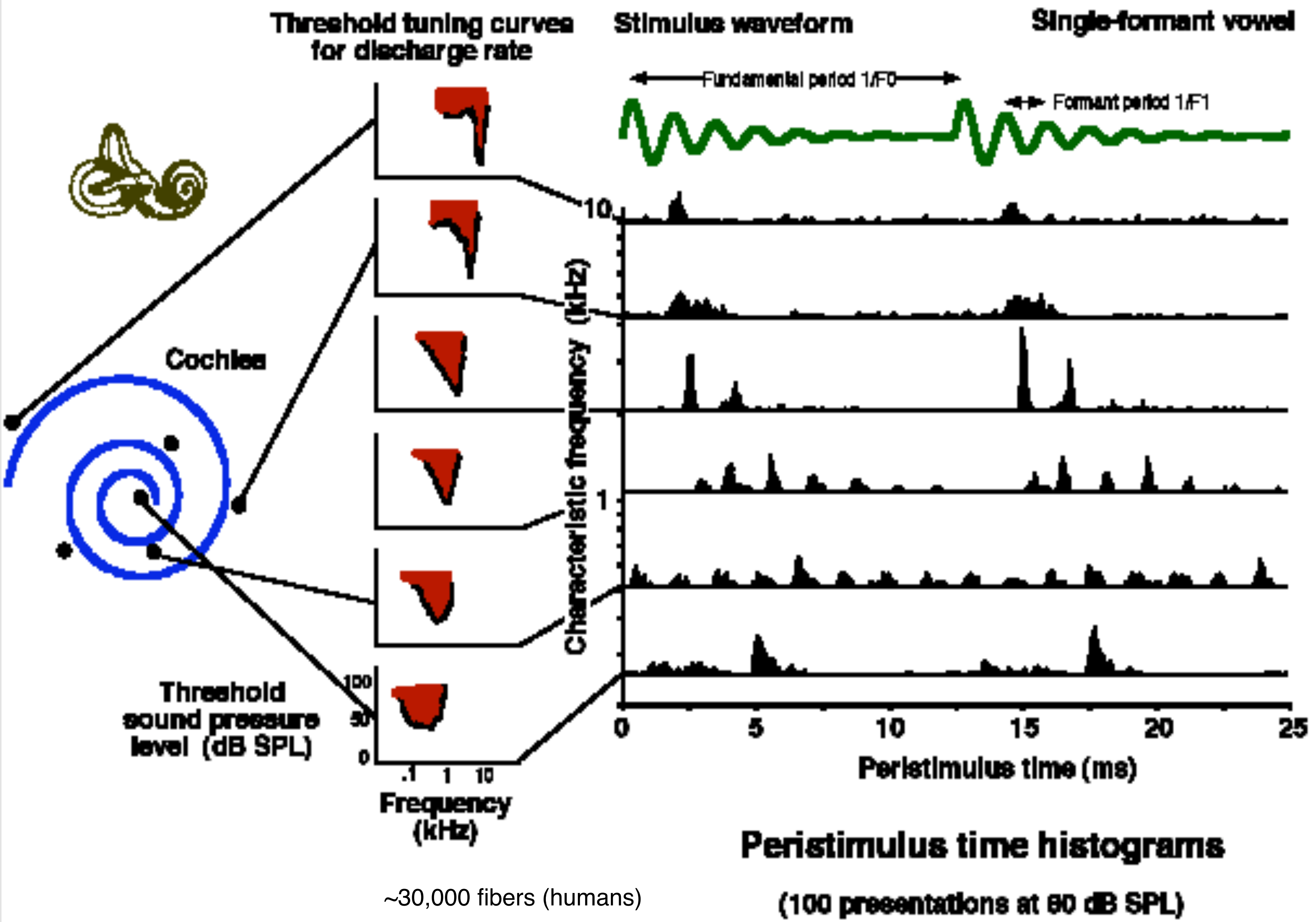
Figure by MIT OpenCourseWare.

Ear and cochlea



Traveling waves along the cochlea. A traveling wave is shown at a given instant along the cochlea, which has been uncoiled for clarity. The graphs profile the amplitude of the traveling wave along the basilar membrane for different frequencies, and show that the position where the traveling wave reaches its maximum amplitude varies directly with the frequency of stimulation. (Figures adapted from Dallos, 1992 and von Bekesy, 1960)

Figure by MIT OpenCourseWare.



~30,000 fibers (humans)

Neural frequency tuning
PLACE PRINCIPLE

Disconnect between cochlear tuning
& pitch discrimination for freqs < 4 khz

CF = characteristic frequency

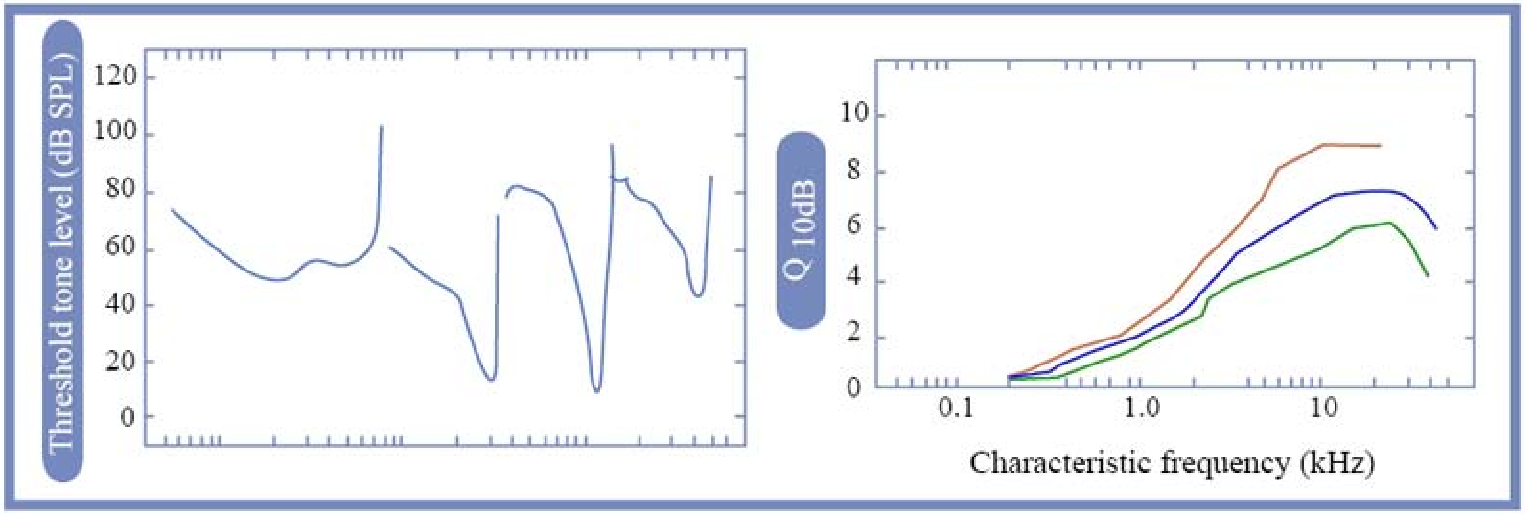
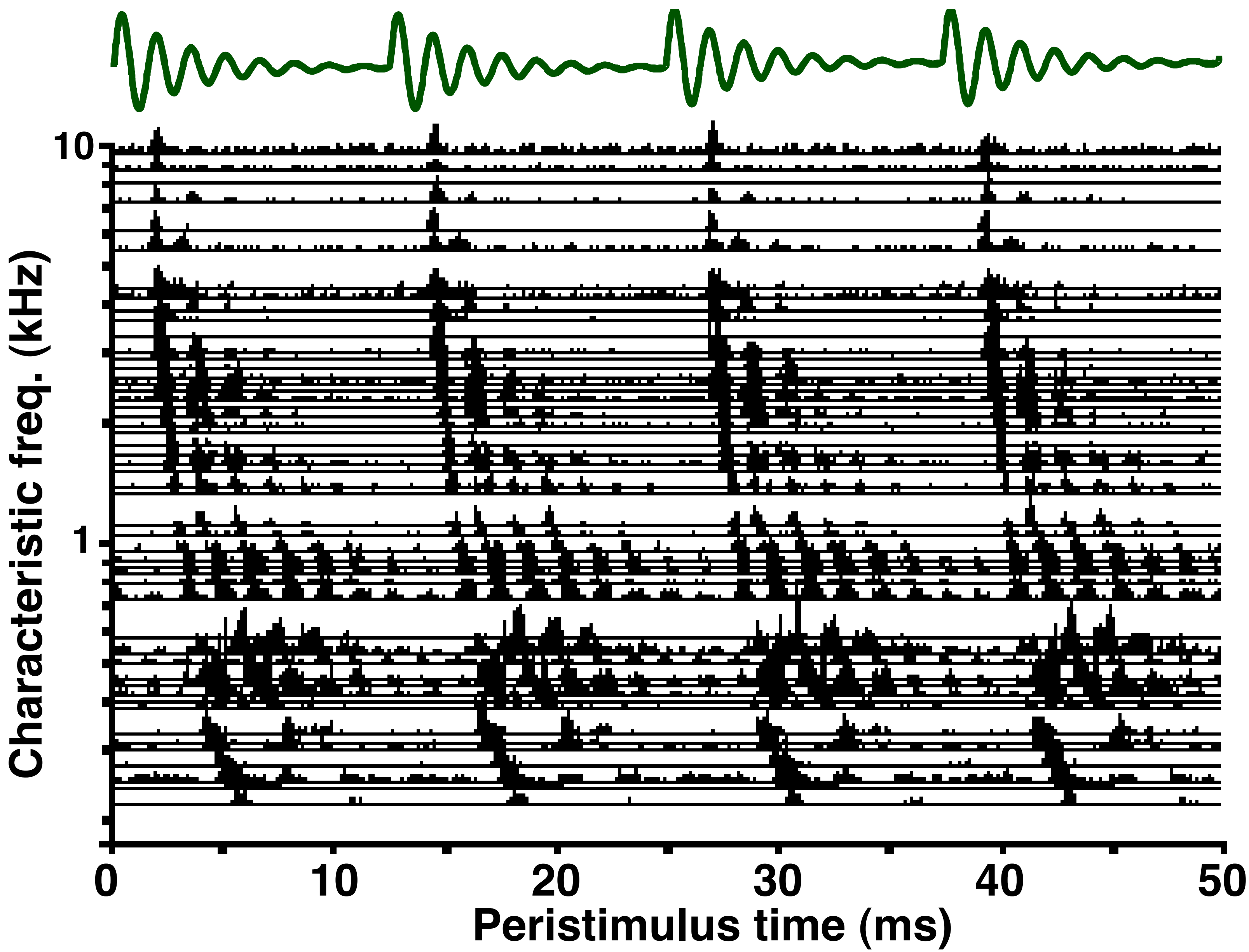


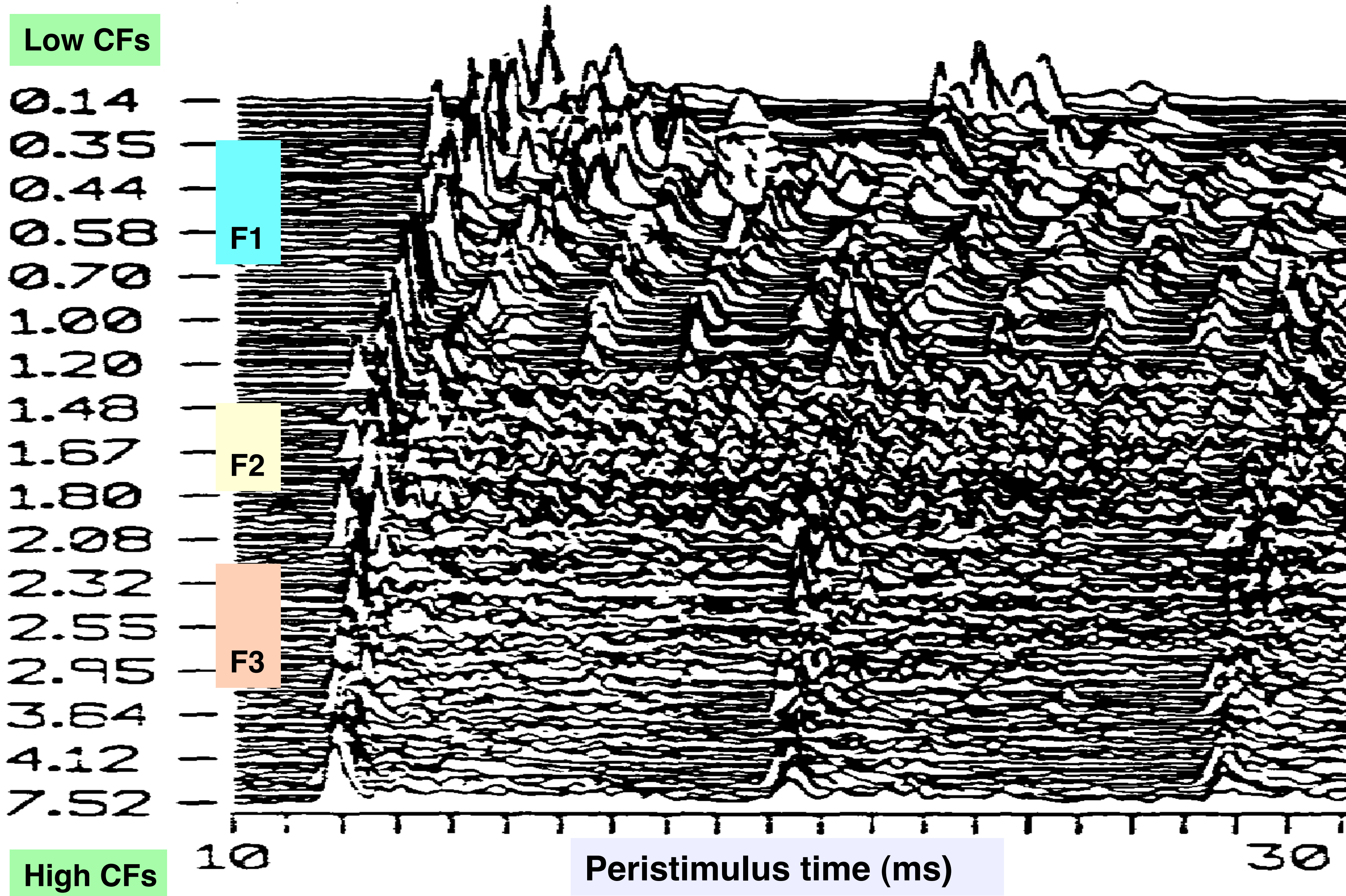
Figure by MIT OpenCourseWare.

Temporal coding in the auditory nerve

Cat, 100x @ 60 dB SPL



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)



Phase-locking in auditory nerve fibers

250 Hz tone

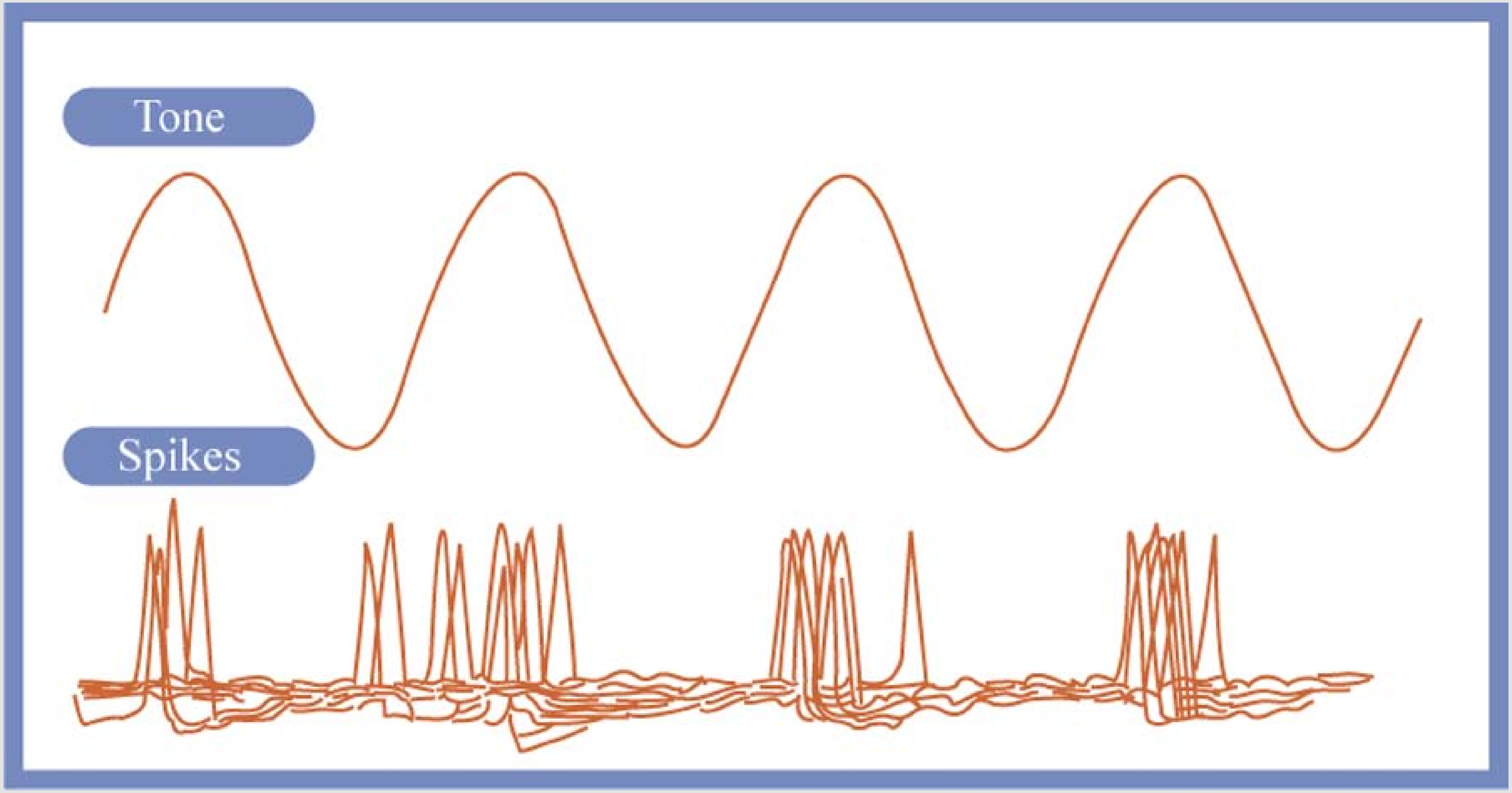
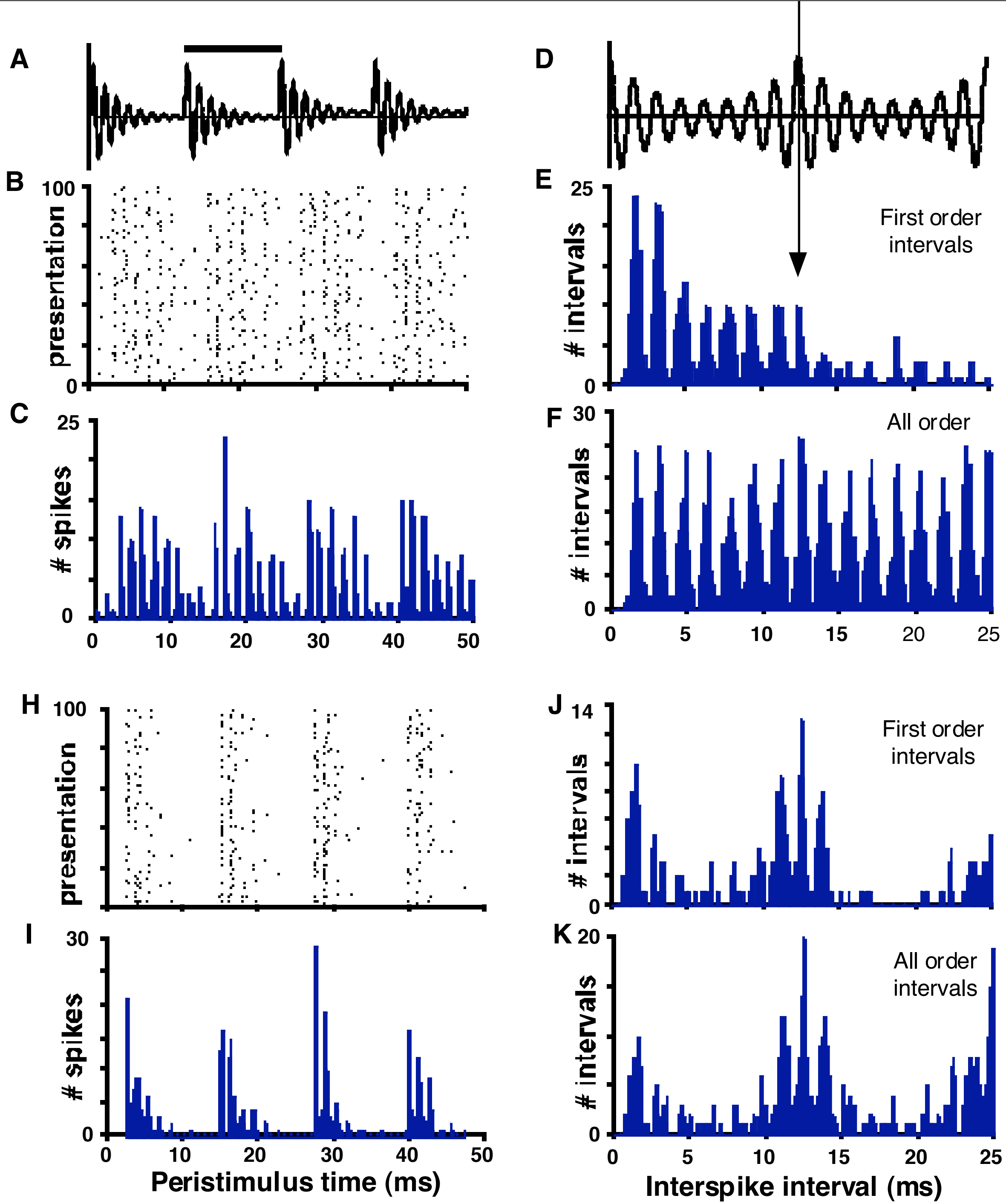


Figure by MIT OpenCourseWare.

See Javel E, McGee JA, Horst W, Farley GR, "Temporal mechanisms in auditory stimulus coding." In: G. M. Edelman, W. E. Gall and W. M. Cowan, ed, Auditory Function: Neurobiological Bases of Hearing, Wiley: New York 1988; p. 518.

Auditory nerve fiber responses to a periodic sound



Interspike intervals in the auditory nerve encode stimulus periodicities

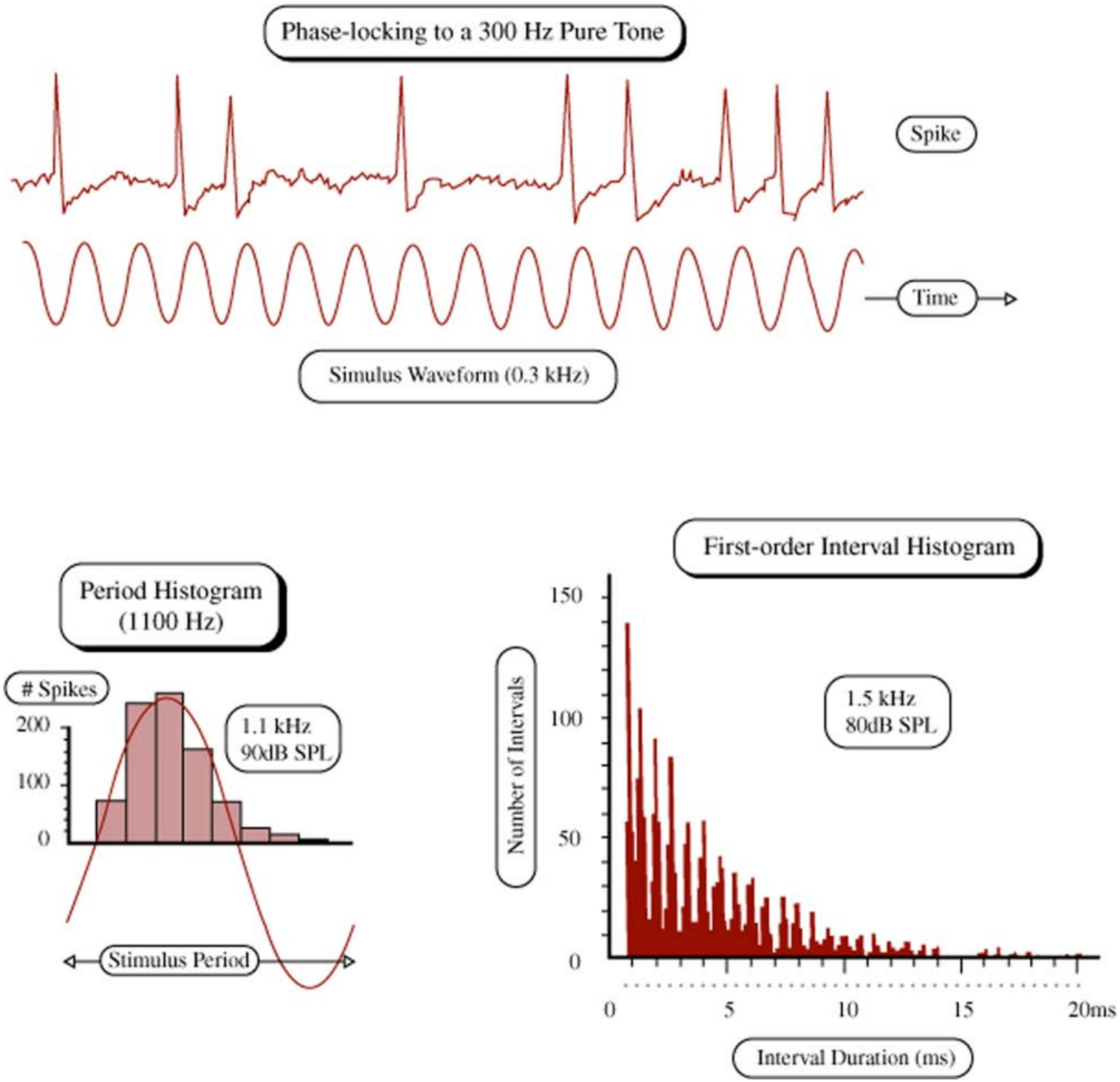


Figure by MIT OpenCourseWare.

Global temporal pitch representation (Collaborator B. Delgutte)

- **All-order** interspike intervals
- **Population-wide** distribution:
  - –All auditory nerve fibers
  - –(all CFs, all SRs)

Predictions

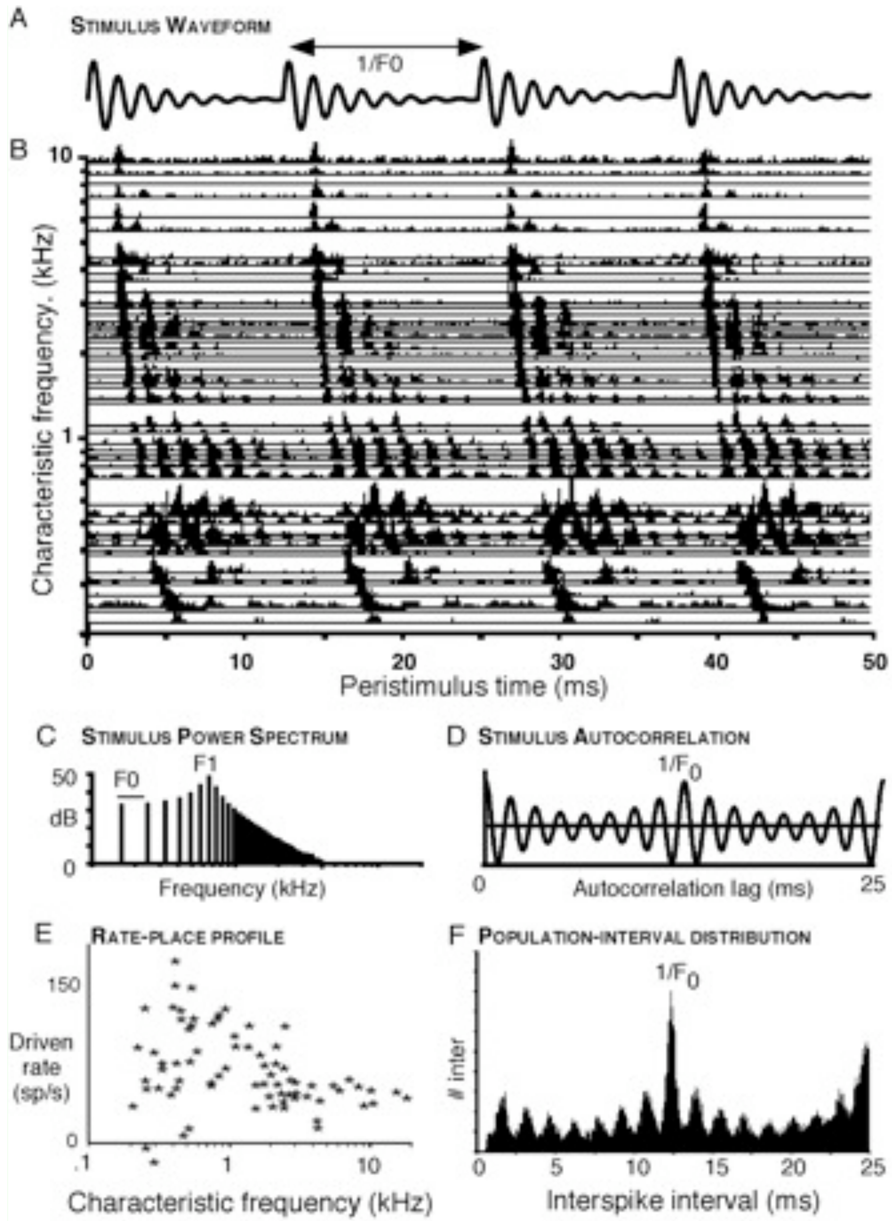
Pitch (frequency) =

the predominant interval or interval pattern

Pitch strength (salience) =

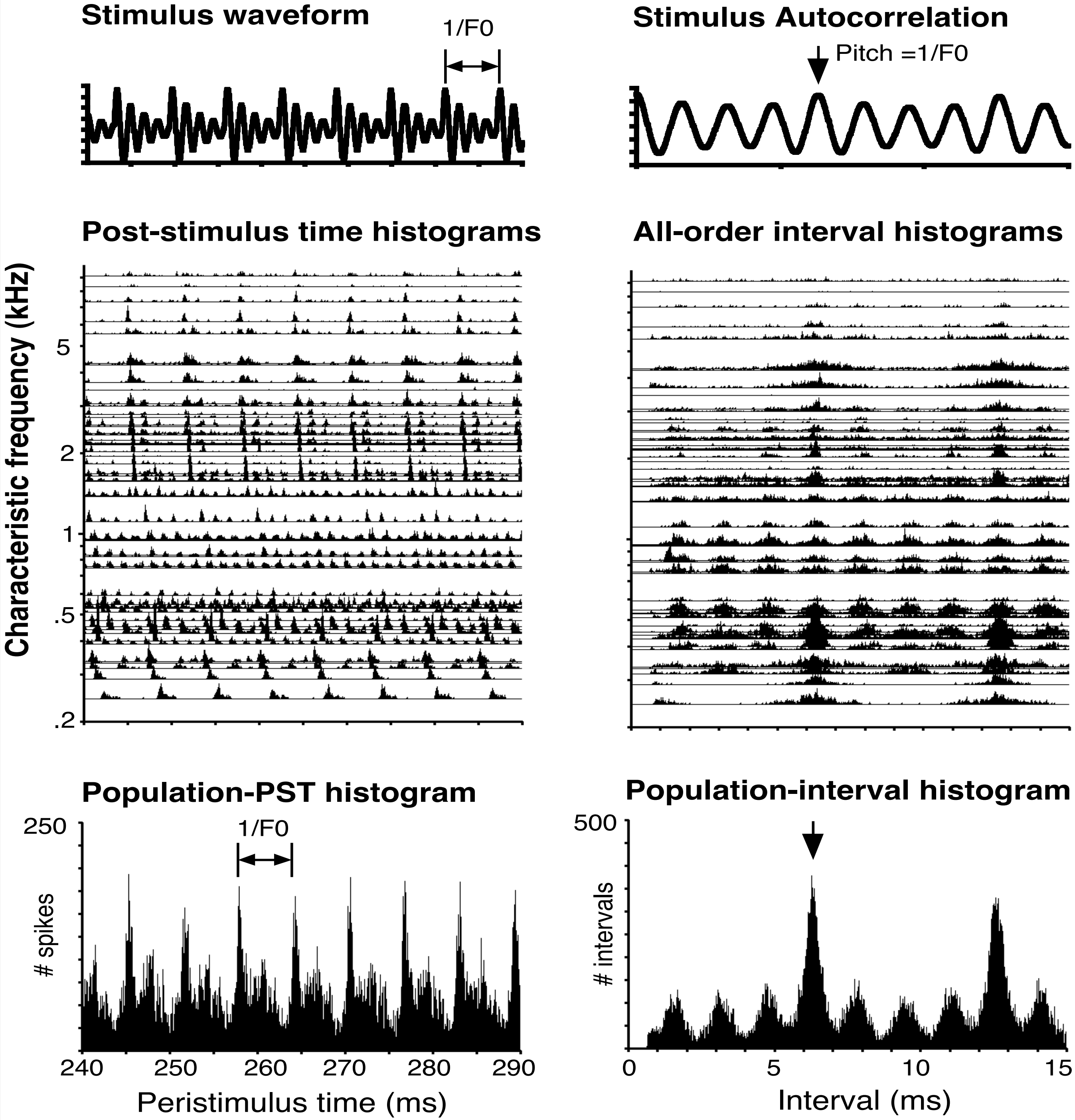
the relative fraction of pitch-related intervals in the whole distribution

Detectability: A pitch can be heard iff its salience

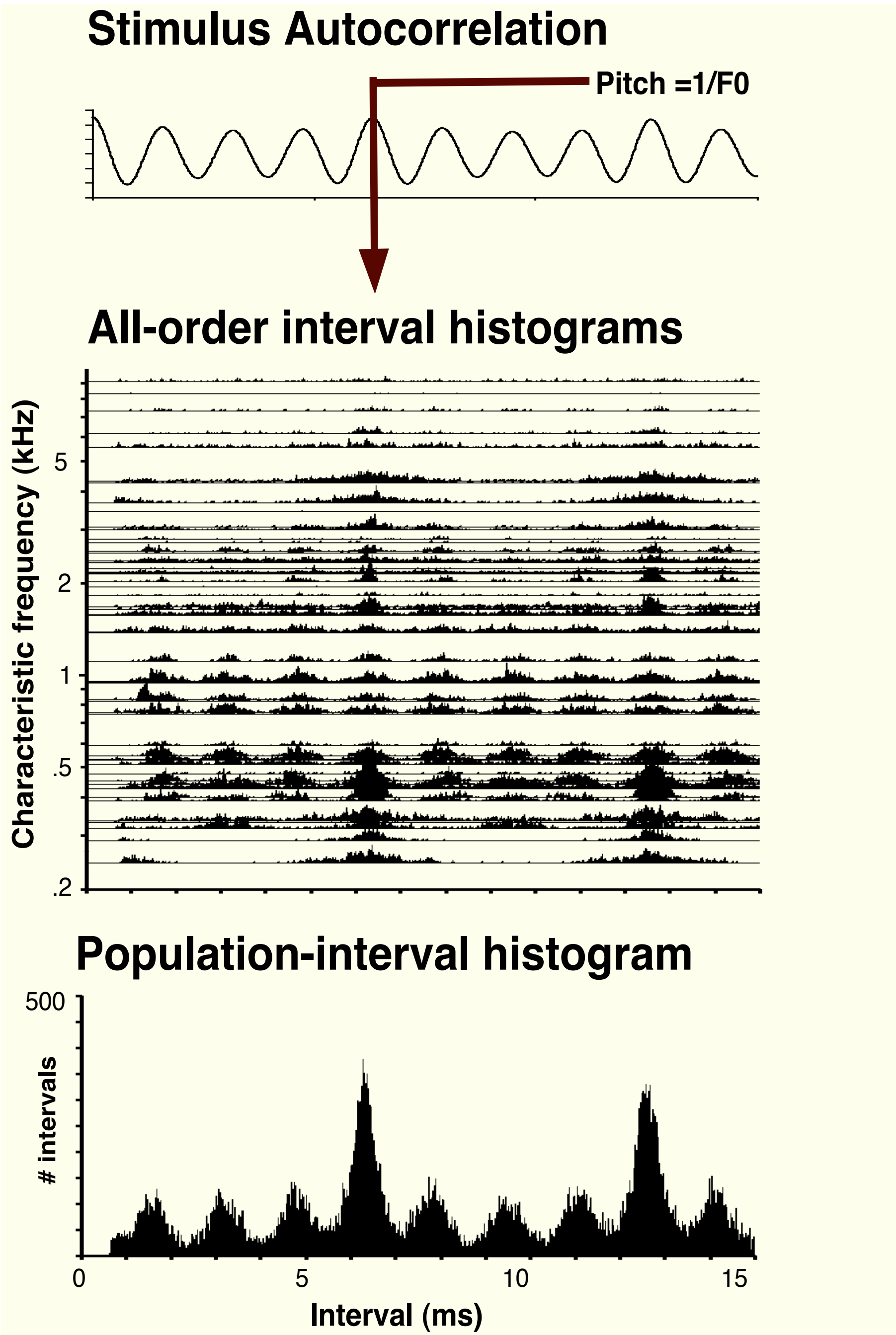


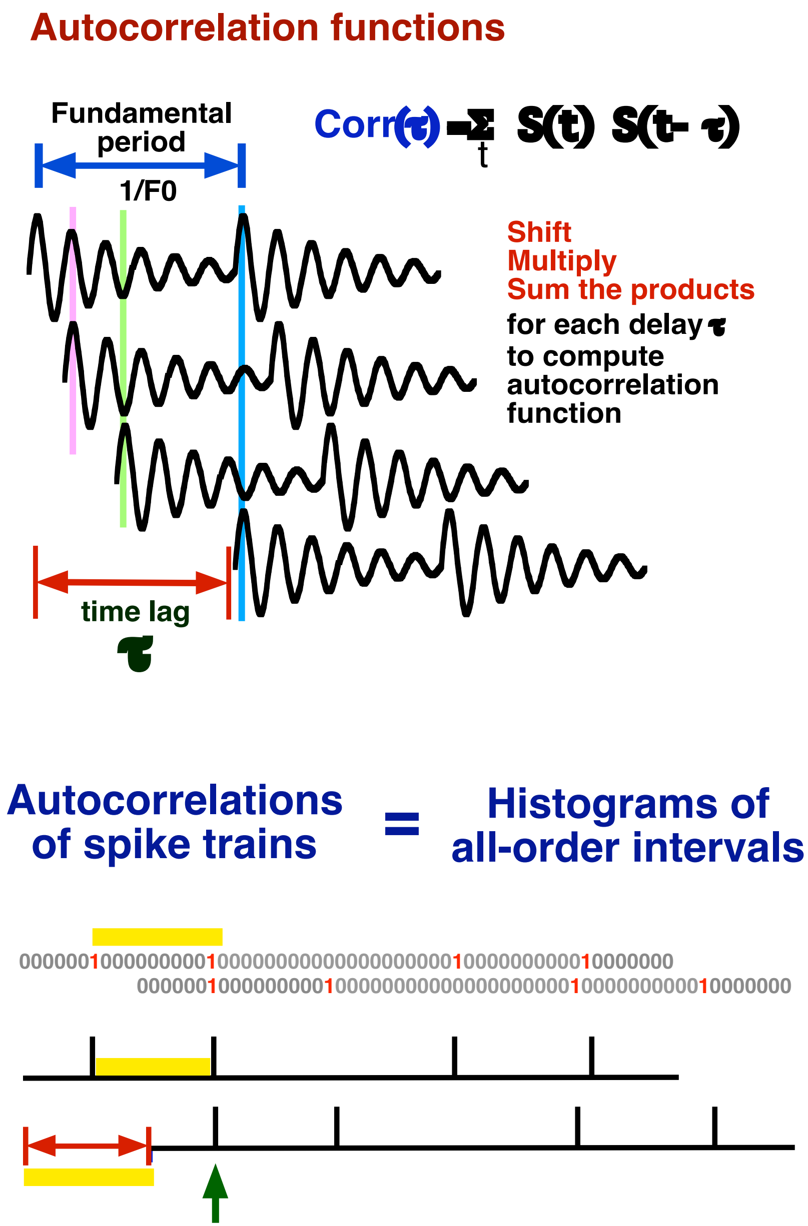
Construction of the population-interval distribution of the auditory nerve

(Cariani & Delgutte, J. Neurophysiol. 1996)



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].
Courtesy of the American Physiological Association. Used with permission.





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]. Courtesy of the American Physiological Association. Used with permission.

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

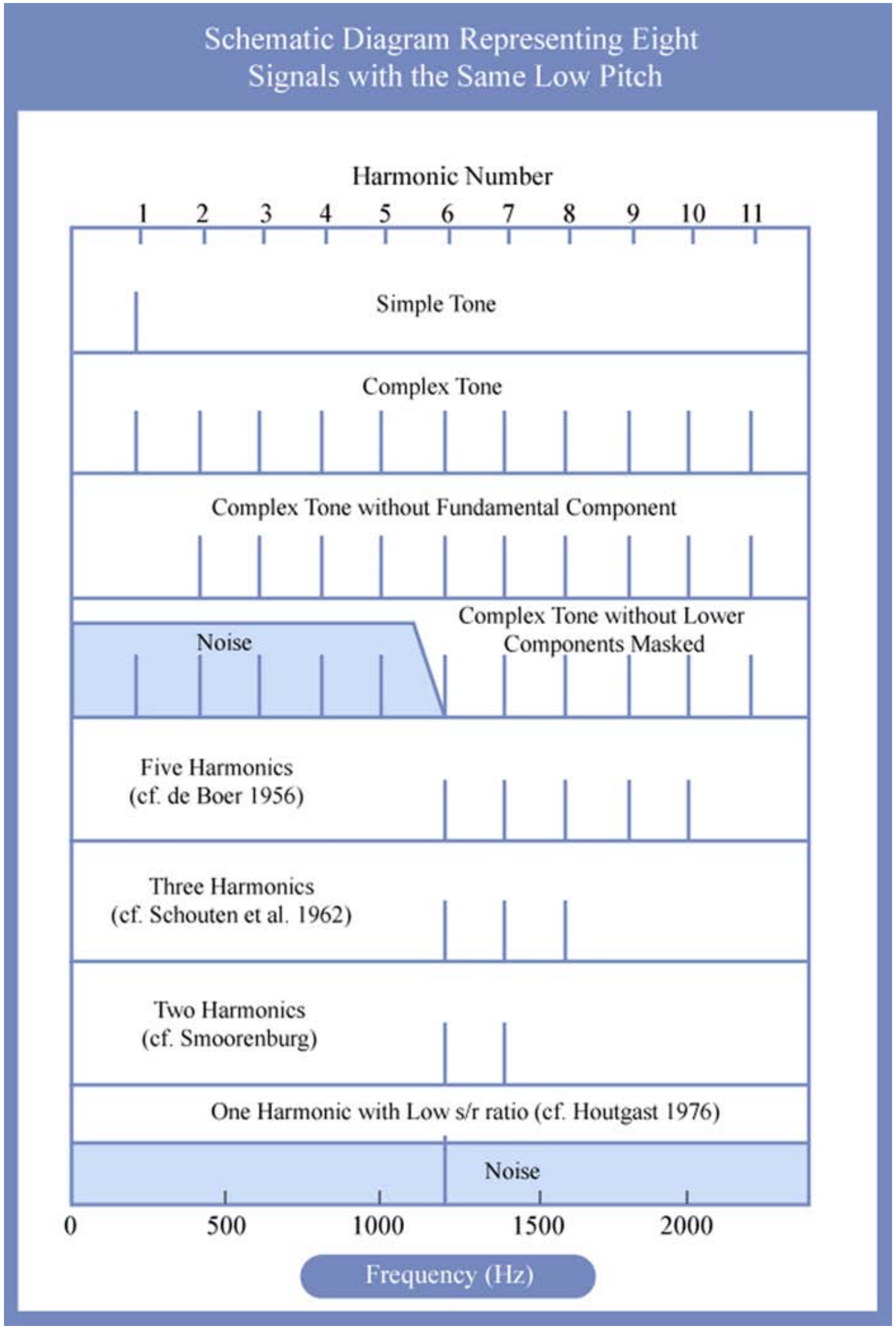
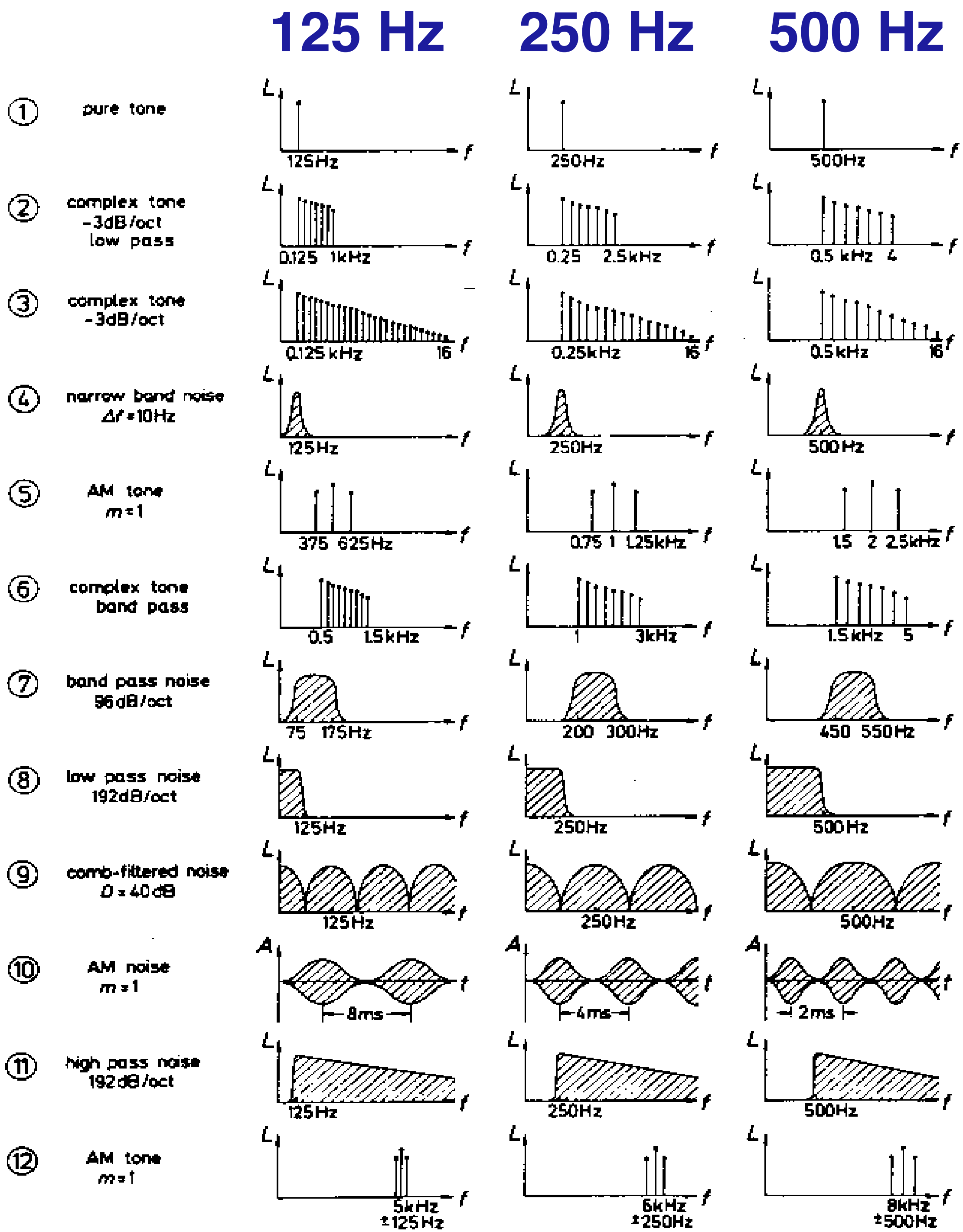


Figure by MIT OpenCourseWare.

Many different sounds produce the same pitch

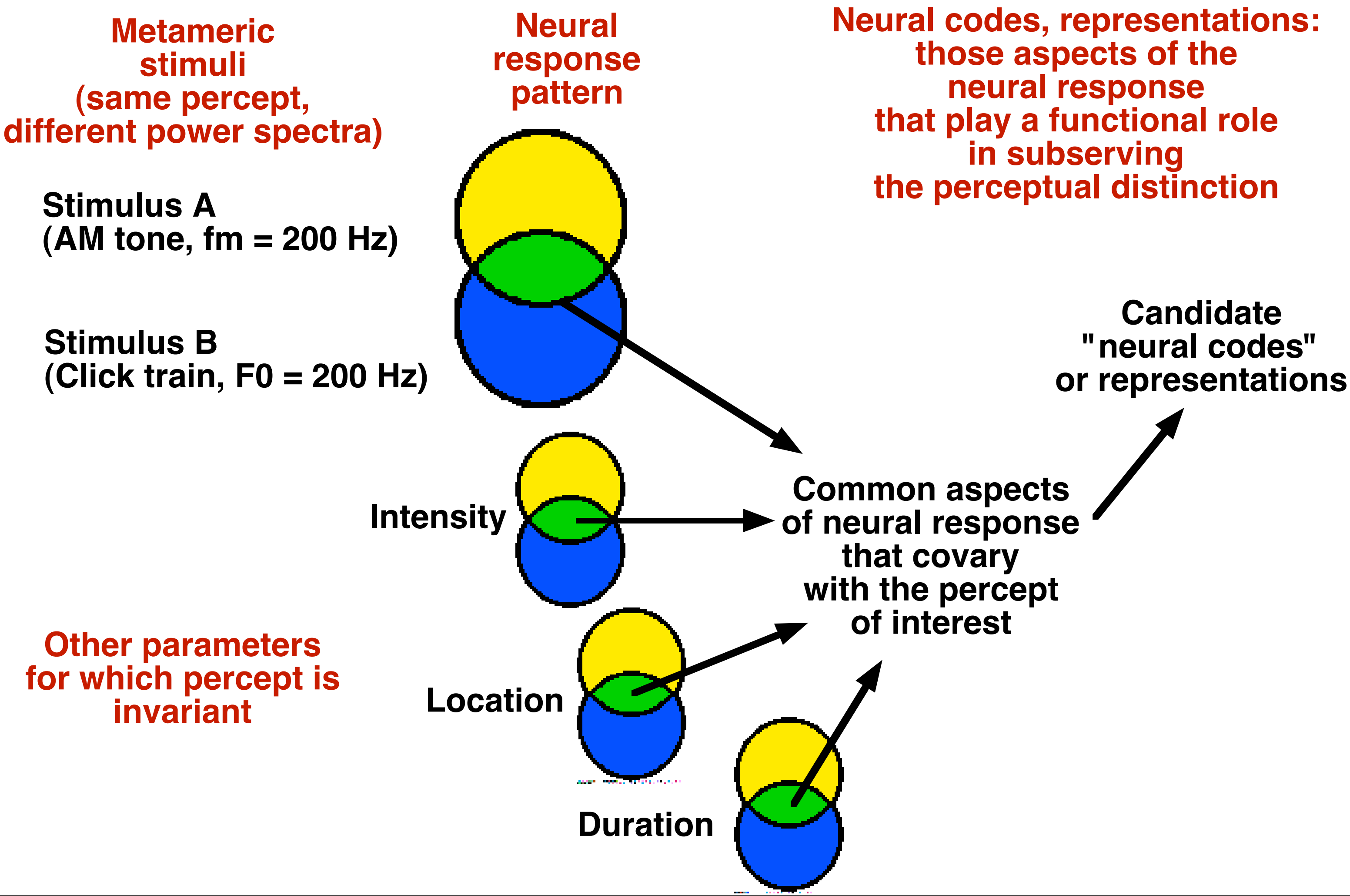
pitch metamery

Fastl, H. & Stoll, G. Scaling of pitch strength, Hearing Research 1(1979): 293-301



Use the structure of perception to find neural codes:

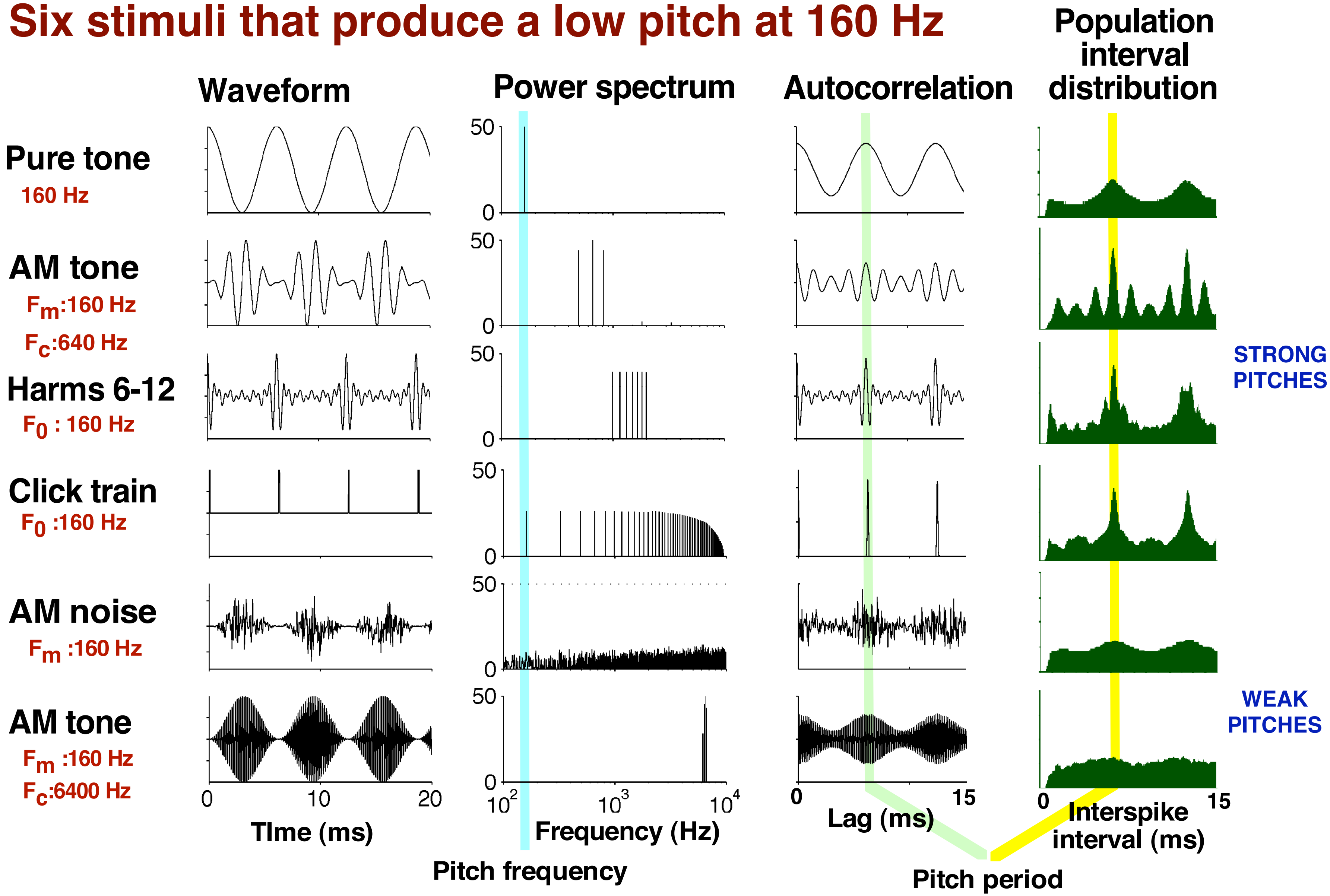
1. Use stimuli that produce equivalent percepts
2. Look for commonalities in neural response
3. Eliminate those aspects that are not invariant



Pitch equivalence classes

(keep the percept constant, identify neural response invariances)

Six stimuli that produce a low pitch at 160 Hz



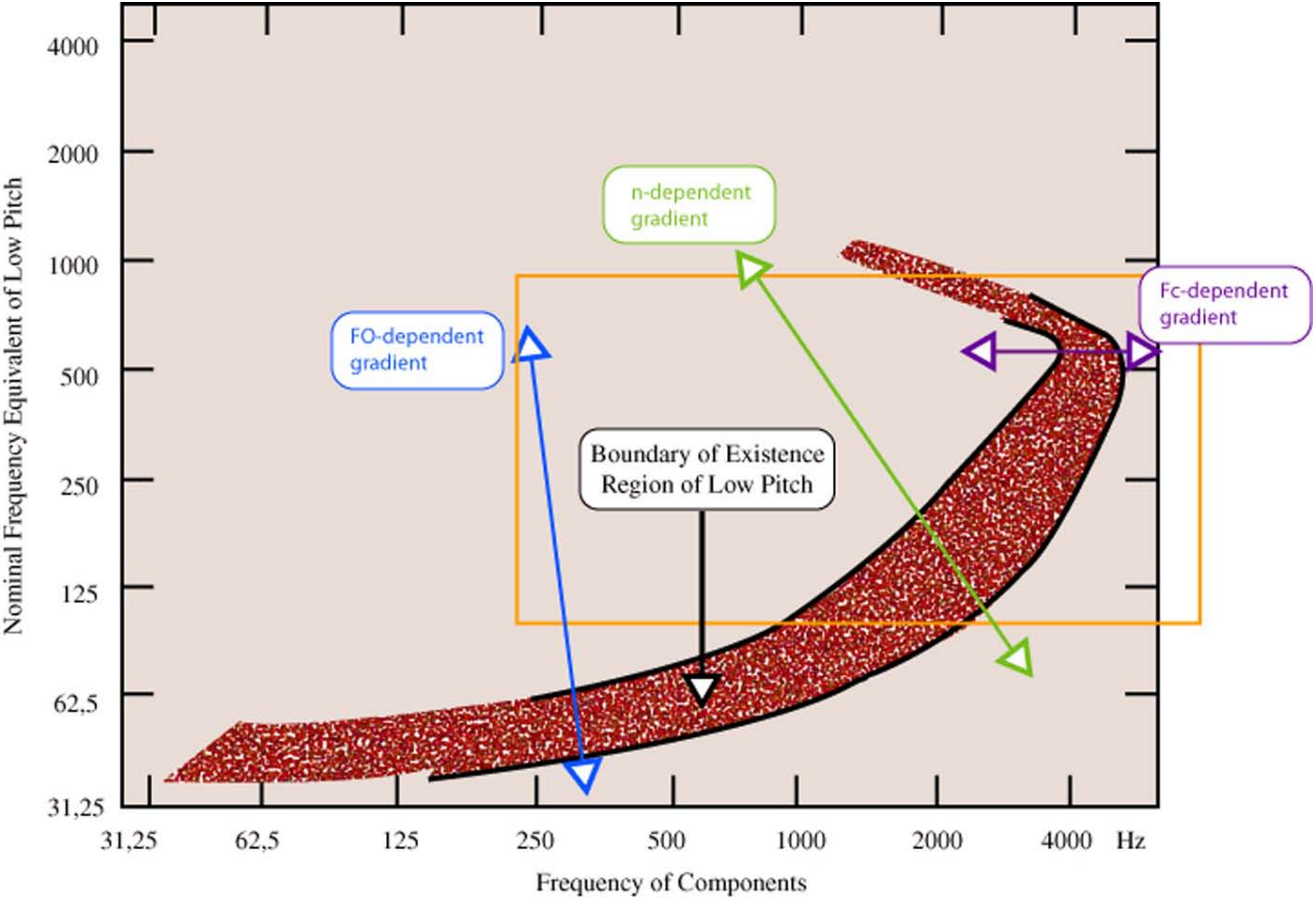


Figure by MIT OpenCourseWare.

van Norden, 1981; after Ritsma (1962)

Level-invariance   Pitch equivalence

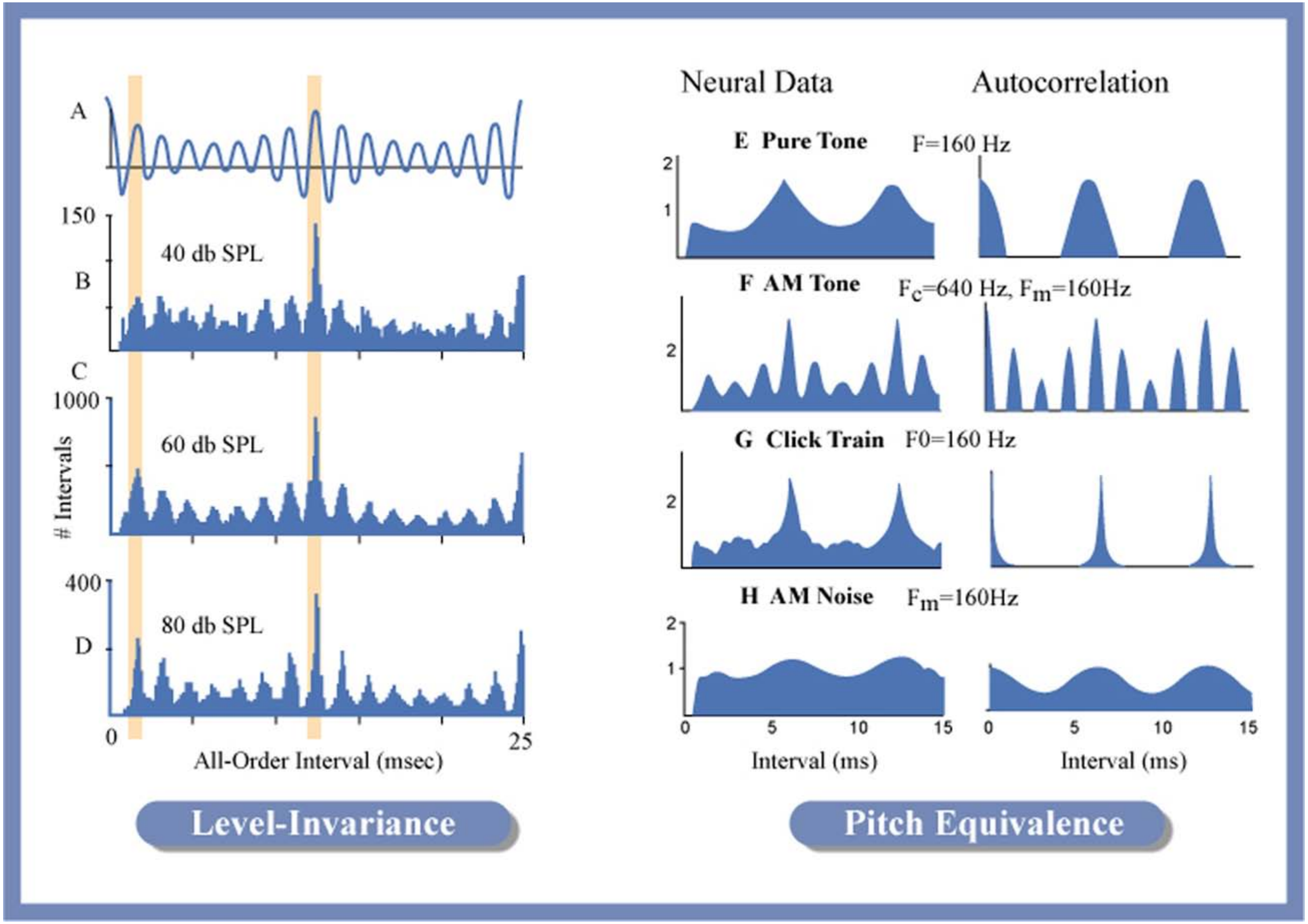
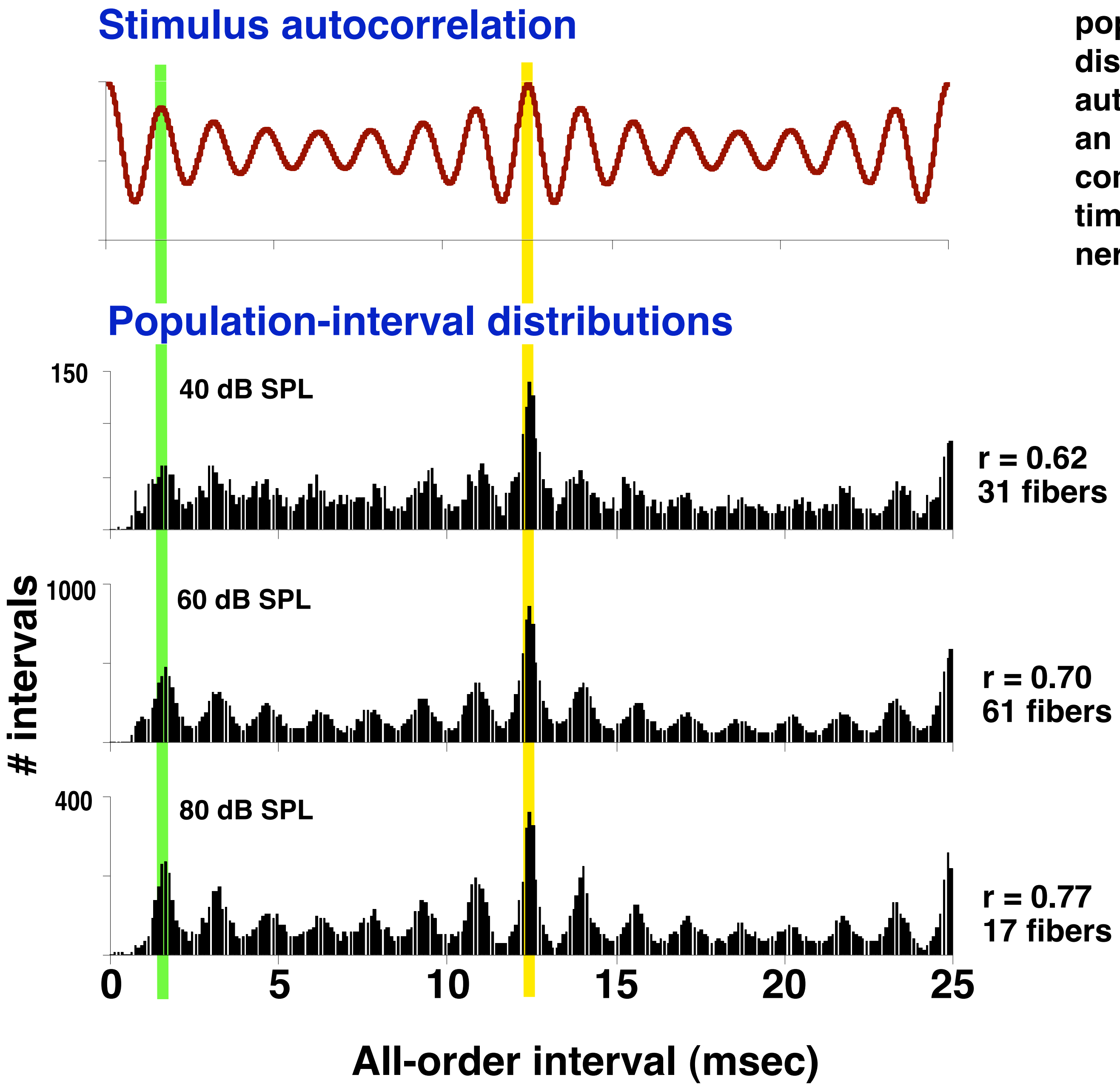


Figure by MIT OpenCourseWare.

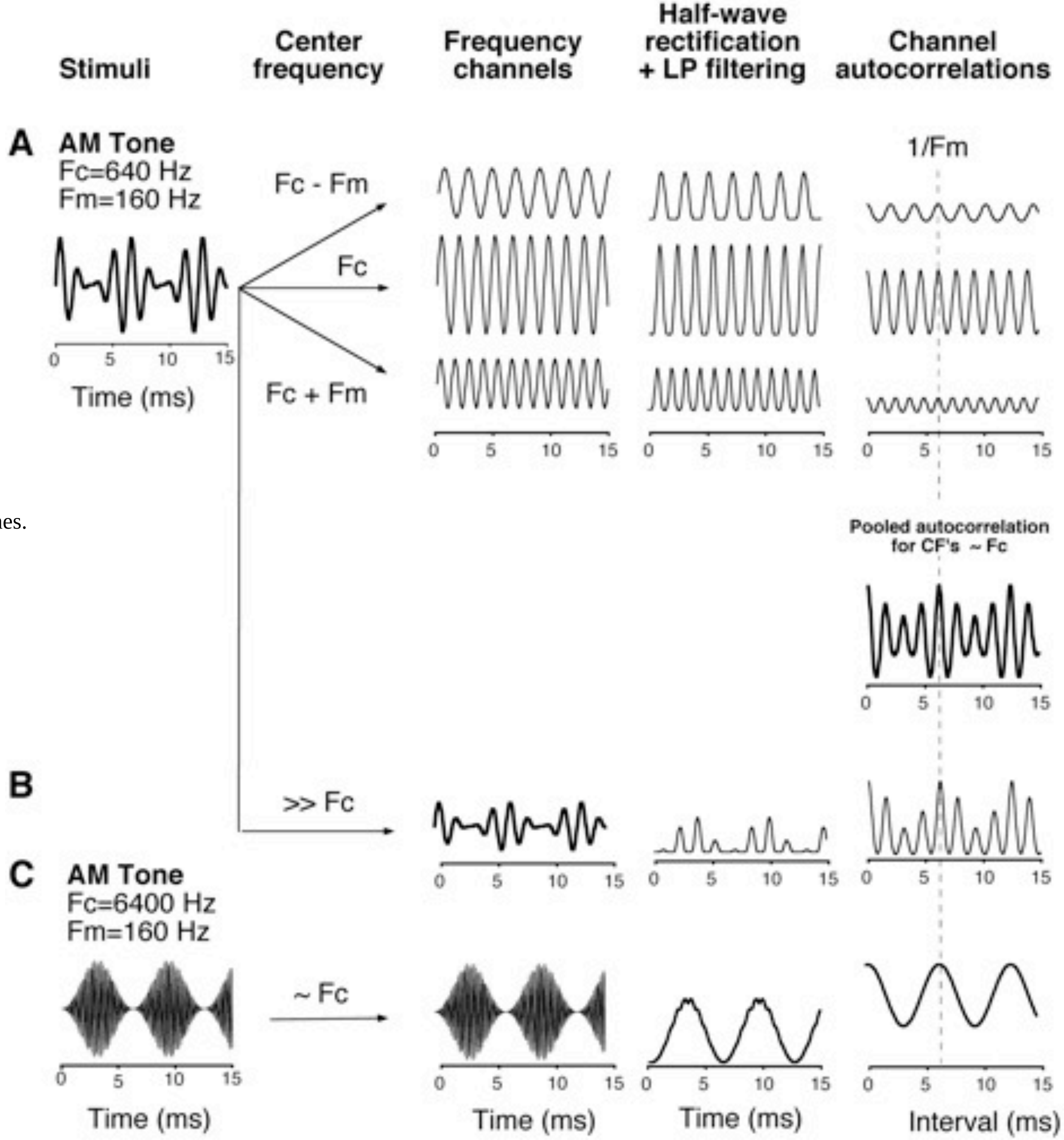
A correlational representation for loudness?



Correlation between population-interval distributions and stimulus autocorrelation serves as an index of the amount of common, stimulus-driven time structure in the auditory nerve array.

As the sound pressure level increases, a progressively greater fraction of the activity in the array is stimulus locked and mutually correlated.

This representation makes use of the dynamic range of the entire auditory nerve array, and leads to a theory of partial loudnesses of multiple auditory objects in a very straightforward way.



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]. Courtesy of the American Physiological Association. Used with permission.

Population-interval distributions and autocorrelation functions

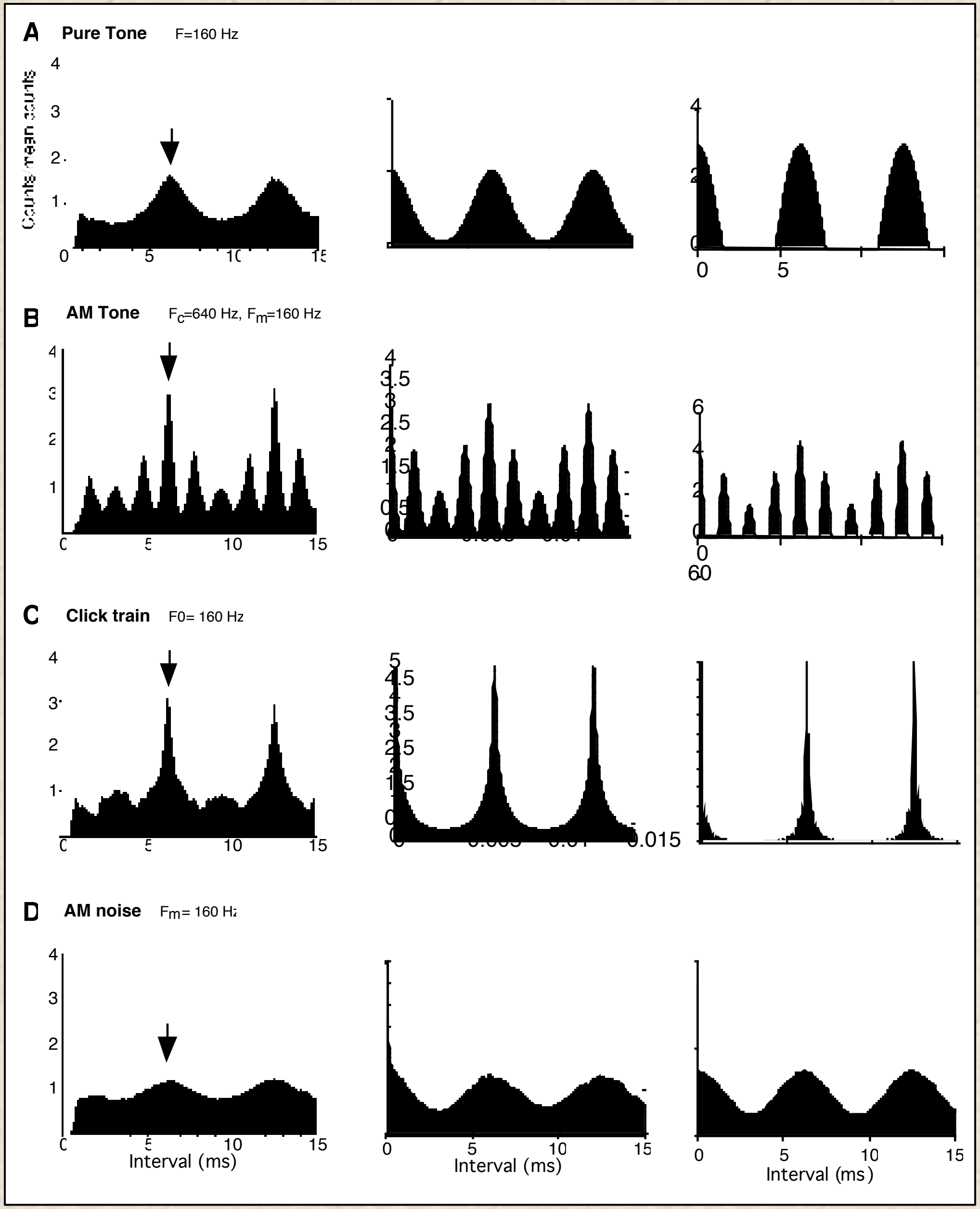
Pure Tone
160 Hz

AM Tone
Fm =160 Hz
Fc = 640 Hz

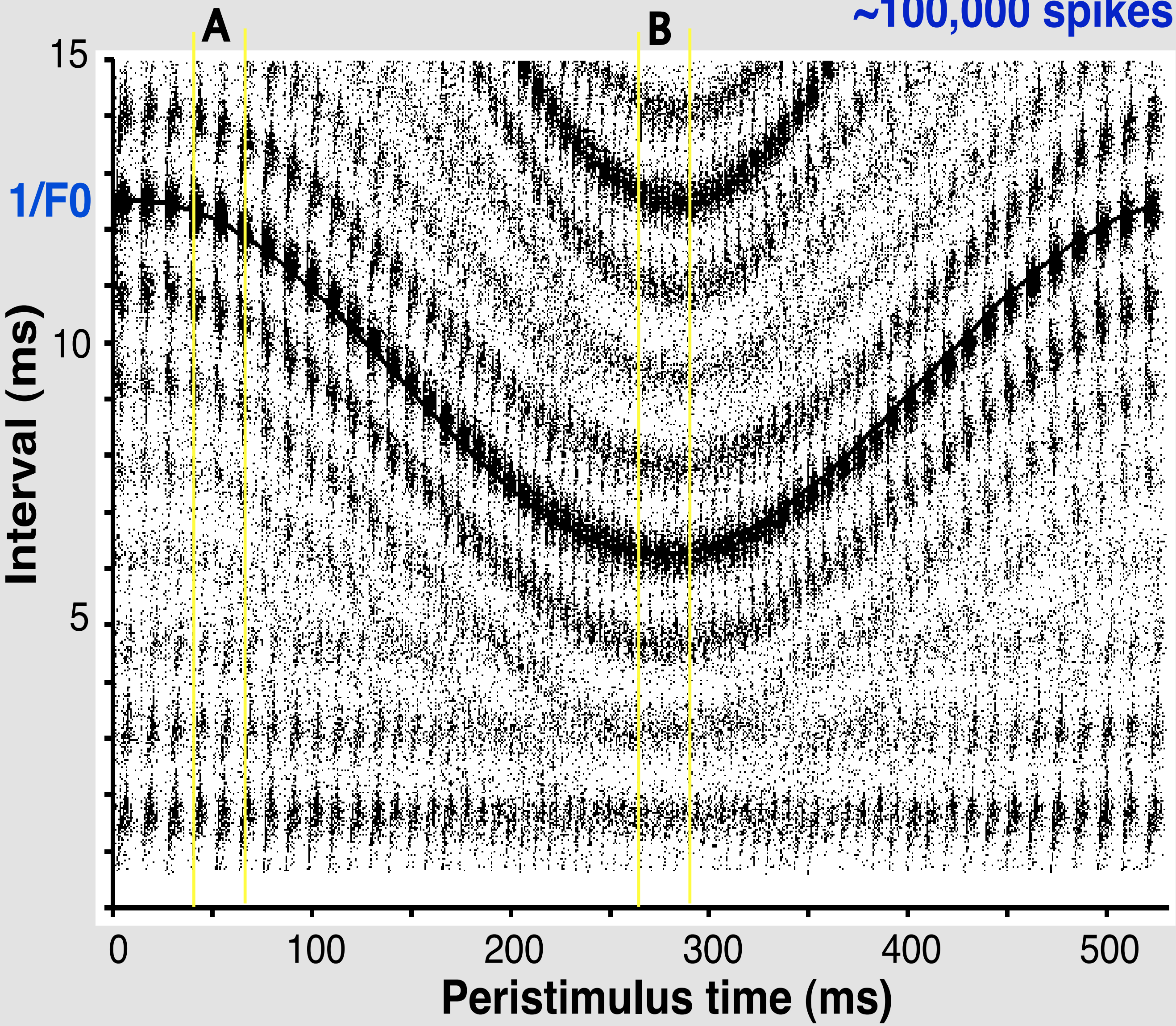
Click Train
F0 = 160 Hz

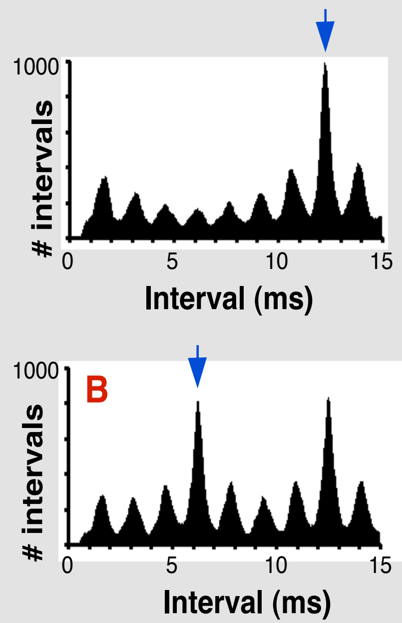
AM Broadband Noise
Fm = 160 Hz

ANF DATA | **ANF SIMULATION** | **AUTO CORRELATION**



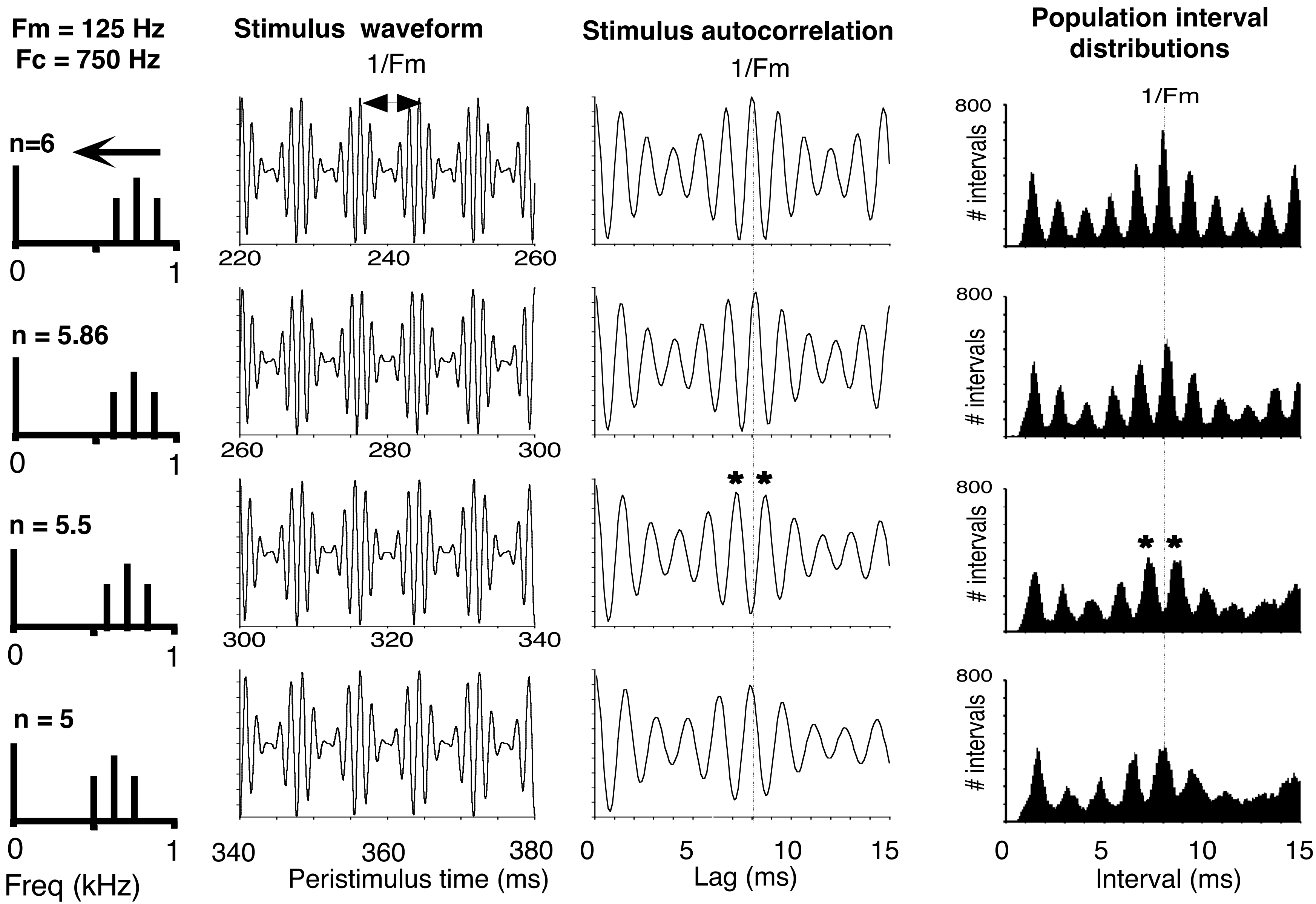
The running population-interval distribution



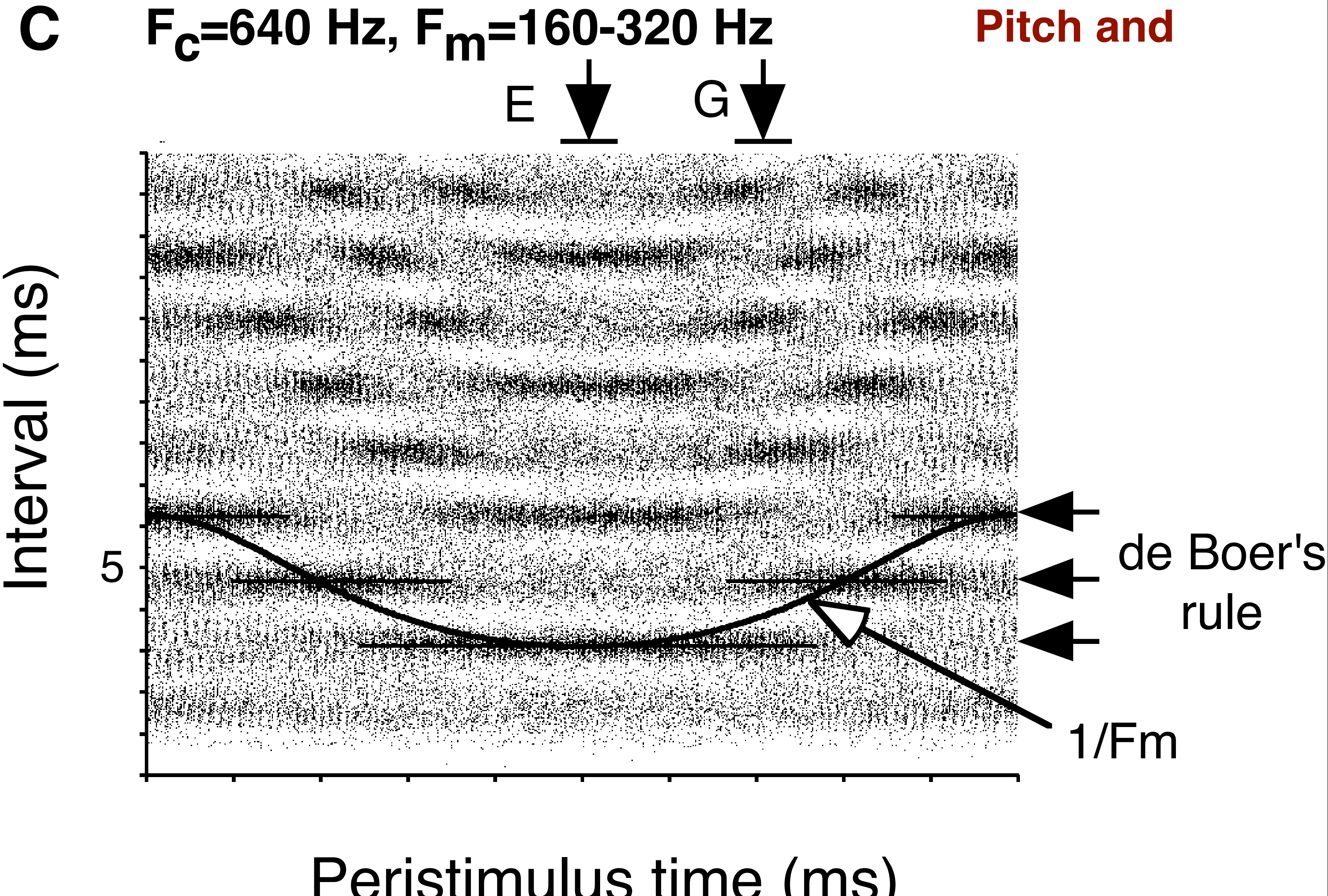


Wednesday, February 11, 2009

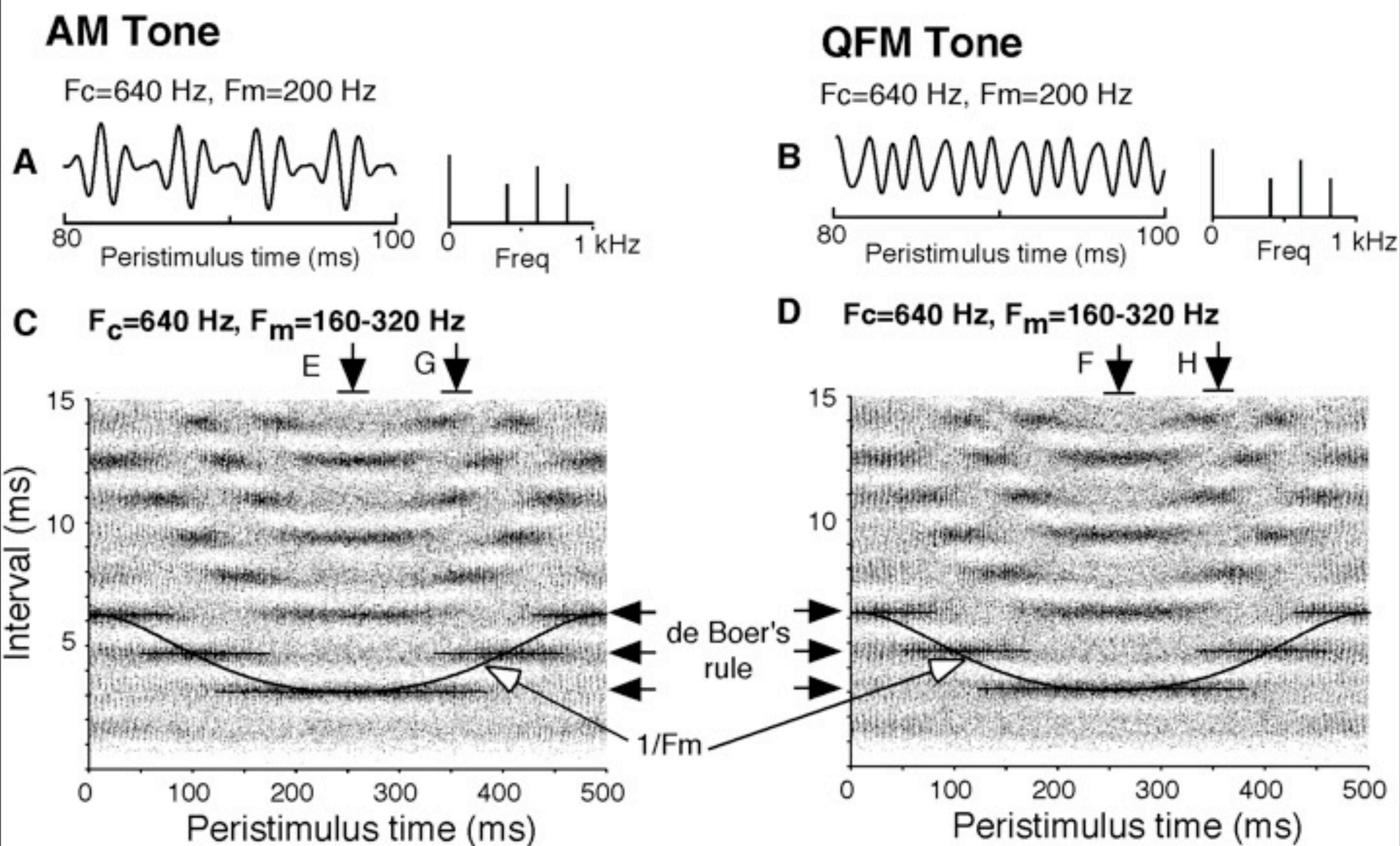
Pitch shift of inharmonic complex tones

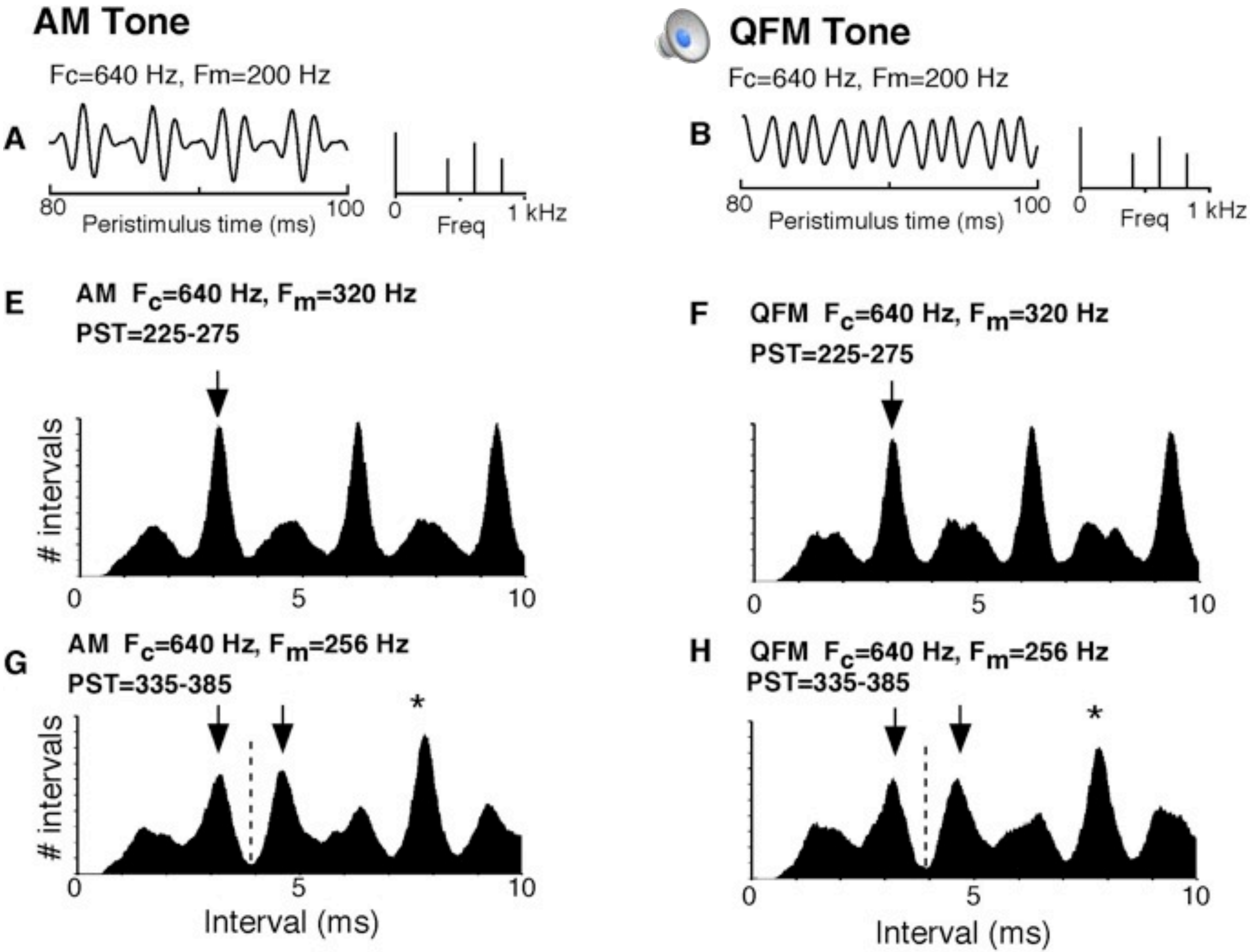


Pitch and


C
F_c=640 Hz, F_m=160-320 Hz
E
G
Interval (ms)
5
de Boer's rule
1/Fm
Peristimulus time (ms)

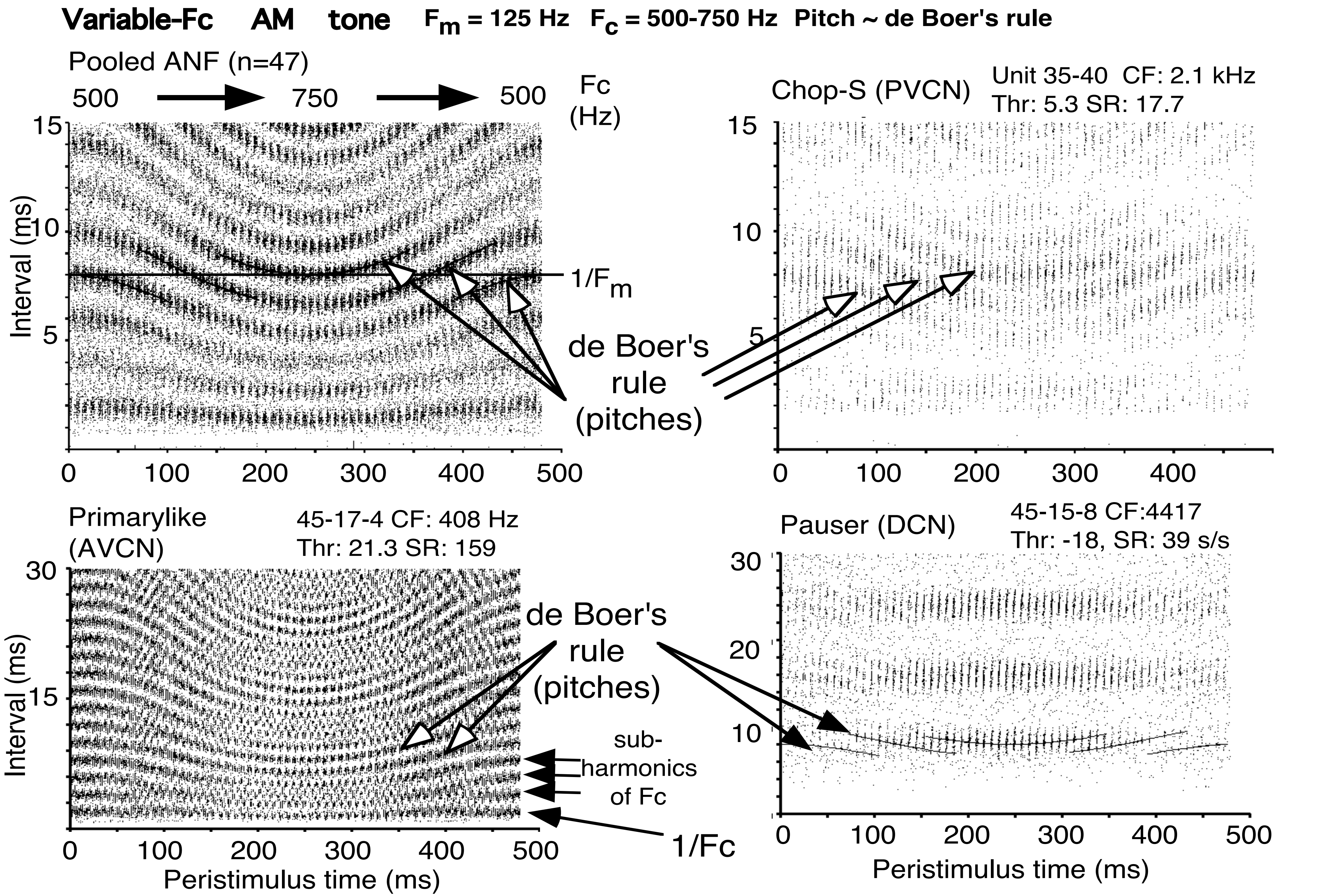
Pitch shift of inharmonic complex tones
Phase-invariant nature of all-order interval code



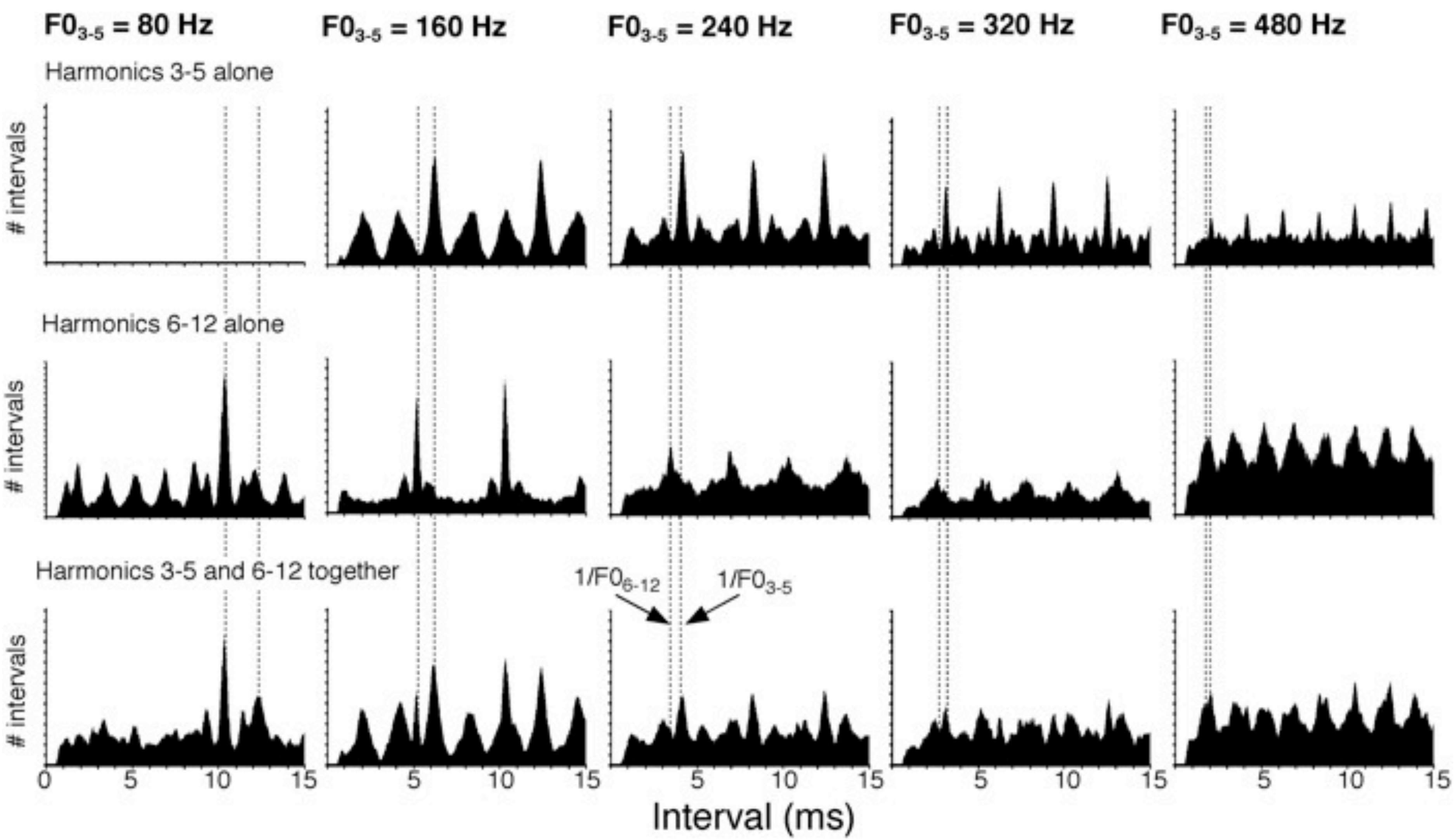


Wednesday, February 11, 2009

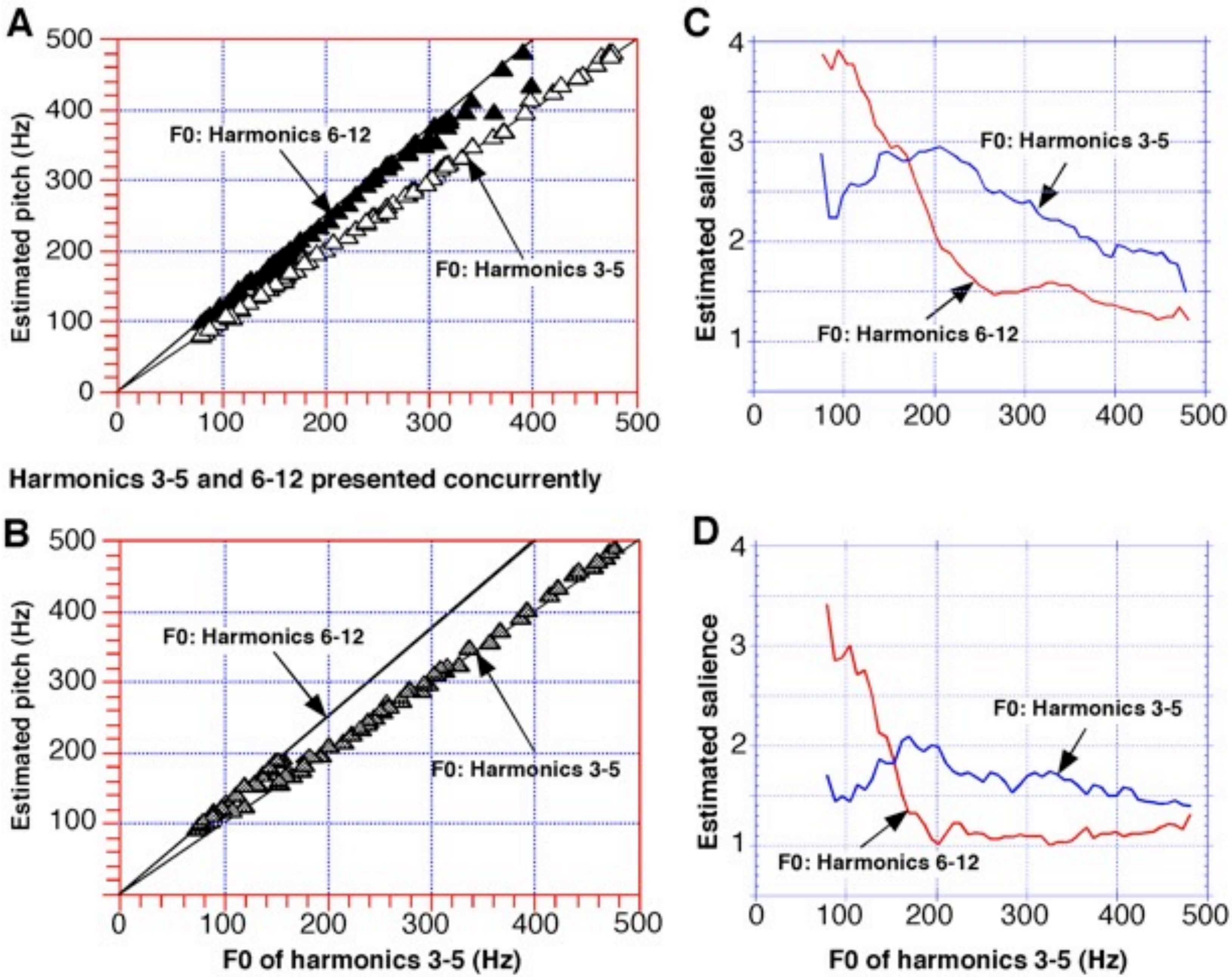
Cochlear nucleus IV: Pitch shift



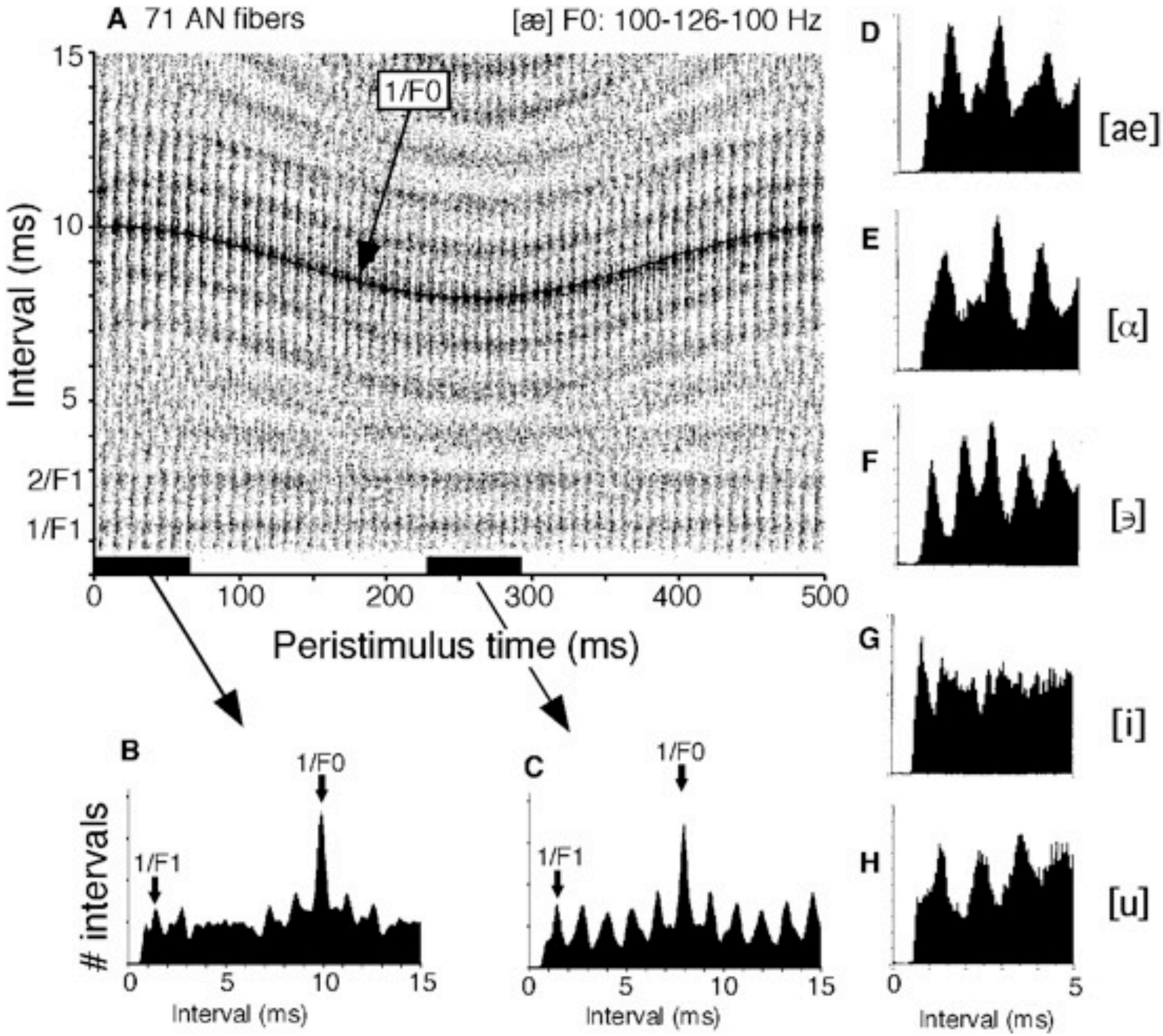
Dominance region for pitch (harmonics 3-5 or partials 500-1500 Hz)



Harmonics 3-5 and 6-12 presented separately



Coding of vowel quality (timbre)



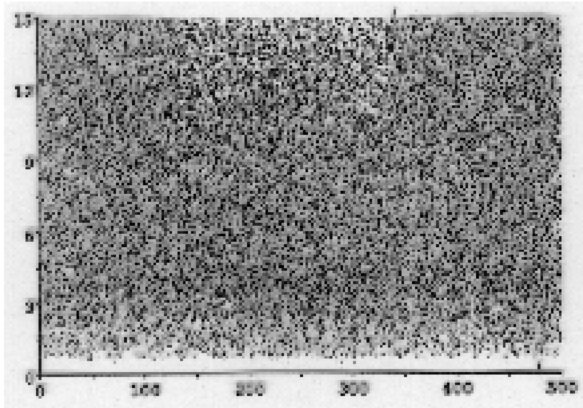
Pitch masking

Variable-F0 click train in broad-band noise
Click train: F0= 160-320 Hz, positive polarity, 80 dB SPL
Peak/background ratio: intervals @ 1/F0 ± 150 usec/mean over all intervals
Informal pitch thresholds, s/n dB: 12 (MT), 13 (PC), 16 (BD)

What degree of temporal correlation is necessary for pitch to become audible?

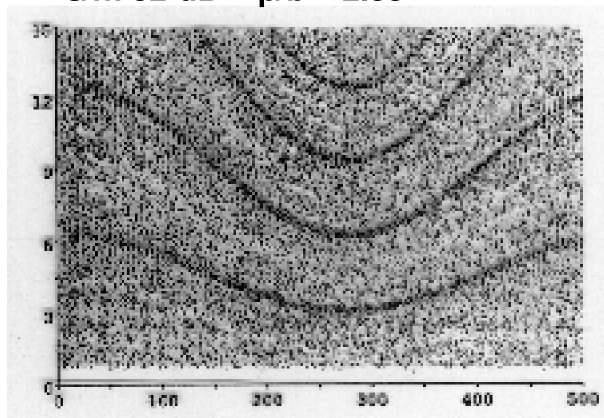
Below threshold (no pitch)

s/n: 8 dB p/b = 1.11



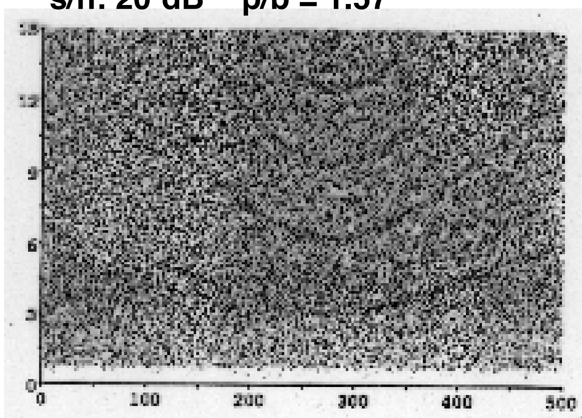
Above threshold (strong pitch)

s/n: 32 dB p/b = 2.38



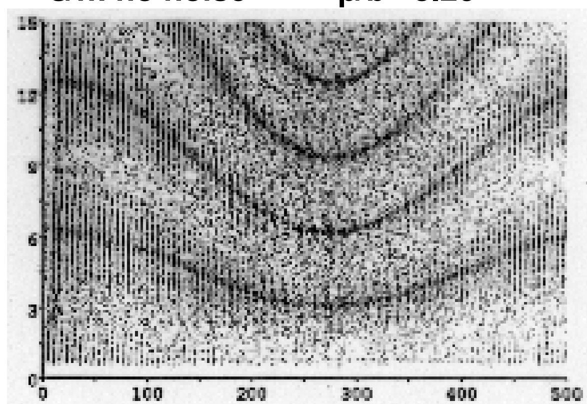
Above threshold (weak pitch)

s/n: 20 dB p/b = 1.57



Above threshold (strong pitch)

s/n: no noise p/b =3.26

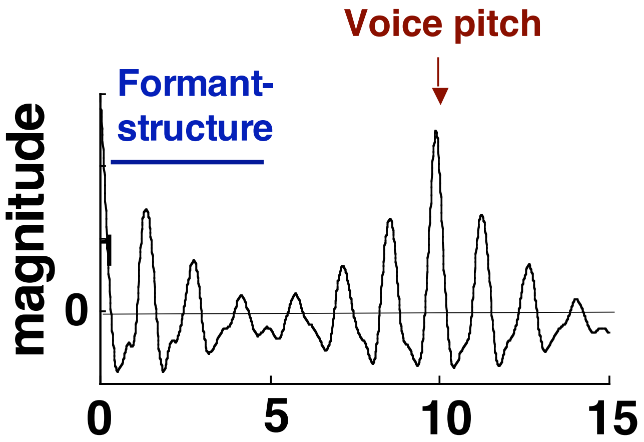


Peristimulus time (ms)

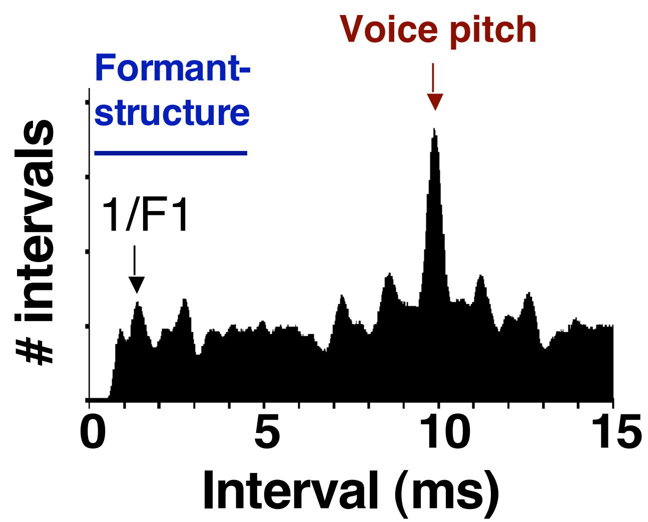
Vowels

Population-interval coding of timbre (vowel formant structure)

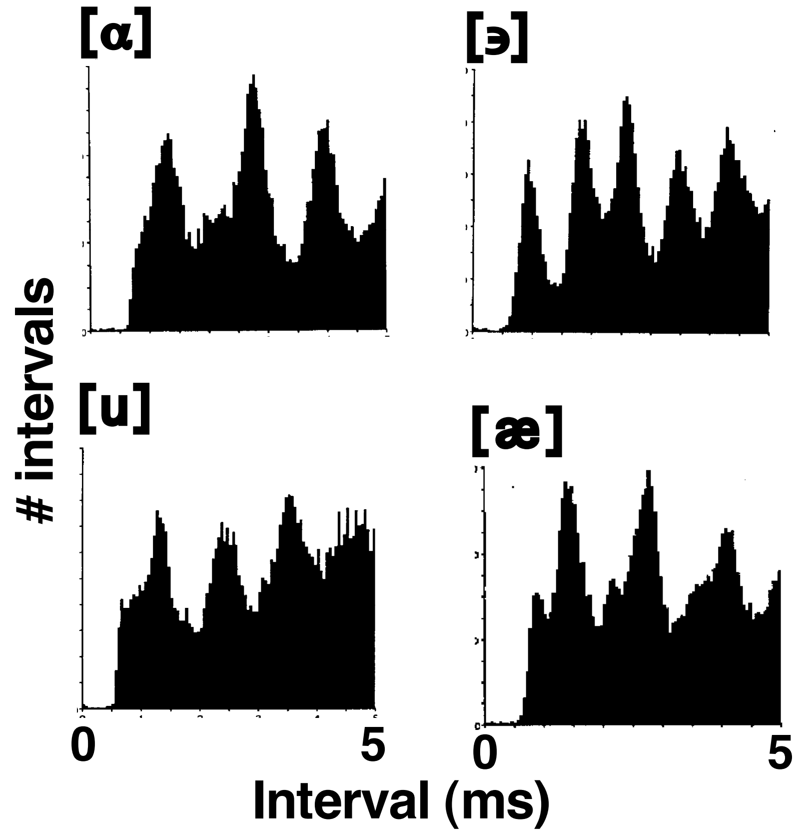
Signal autocorrelation [ae]



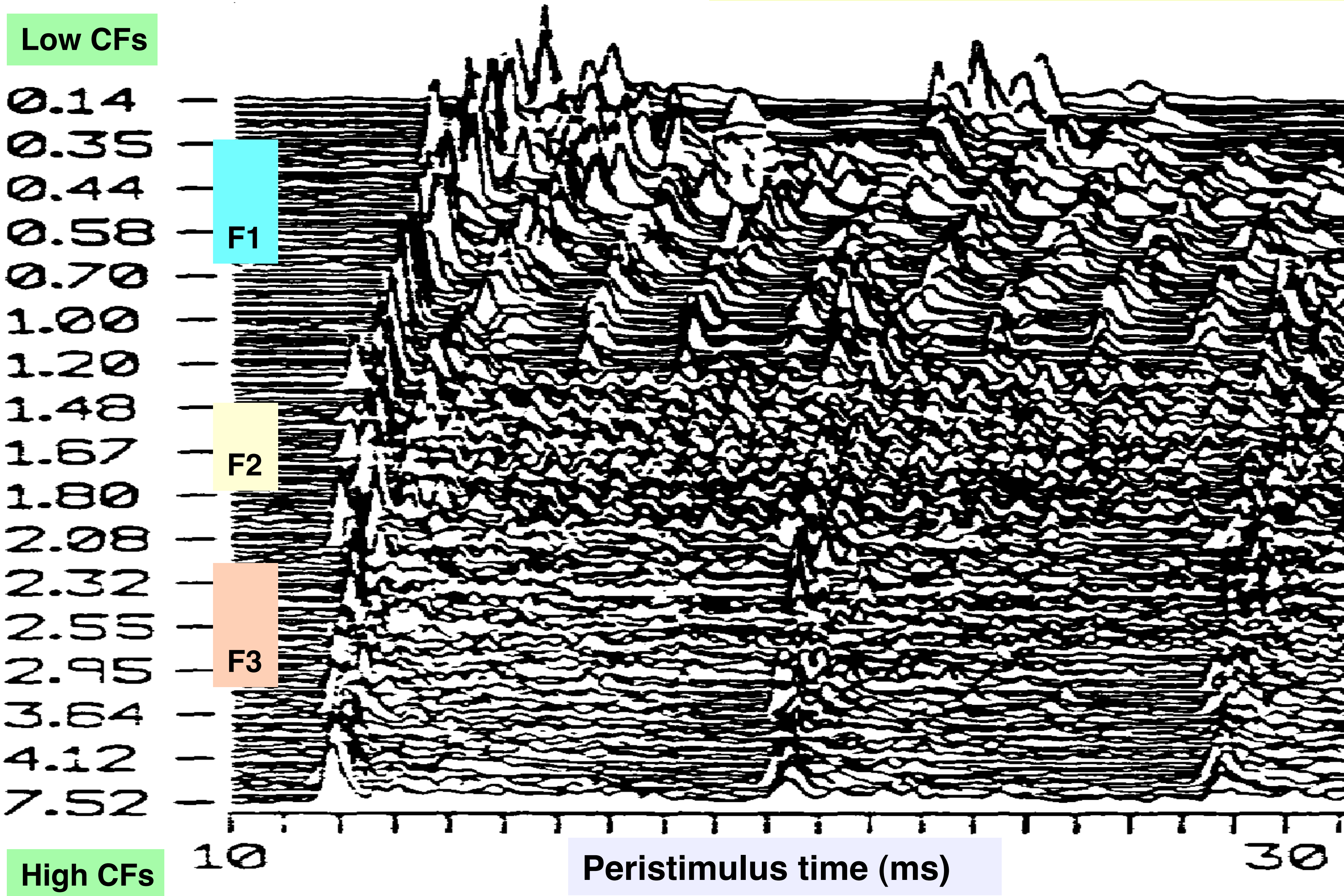
Population interval histogram



Population-wide distributions of short intervals for 4 vowels



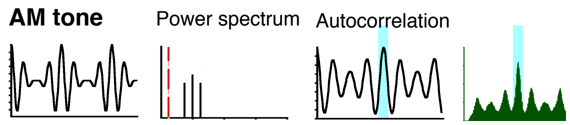
Time domain analysis of auditory-nerve fiber firing rates.
Hugh Secker-Walker & Campbell Searle, J. Acoust. Soc. 88(3), 1990
Neural responses to /da/ @ 69 dB SPL from Miller and Sachs (1983)



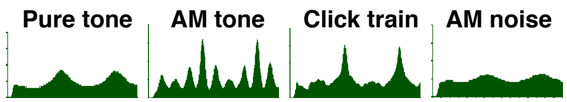
Summary

Population-interval representation of pitch at the level of the auditory nere

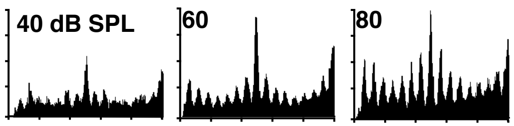
Pitch of the "missing fundamental"



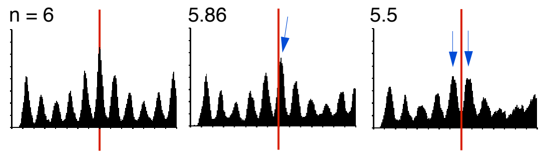
Pitch Equivalence



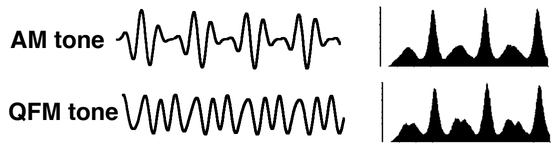
Level invariance



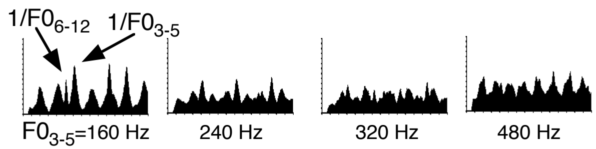
Pitch shift of inharmonic AM tones



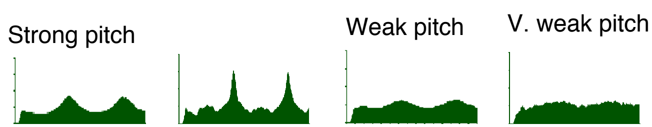
Phase invariance



Dominance region



Pitch salience



Temporal coding of pitch in the auditory nerve

Pitch = predominant all-order interspike interval

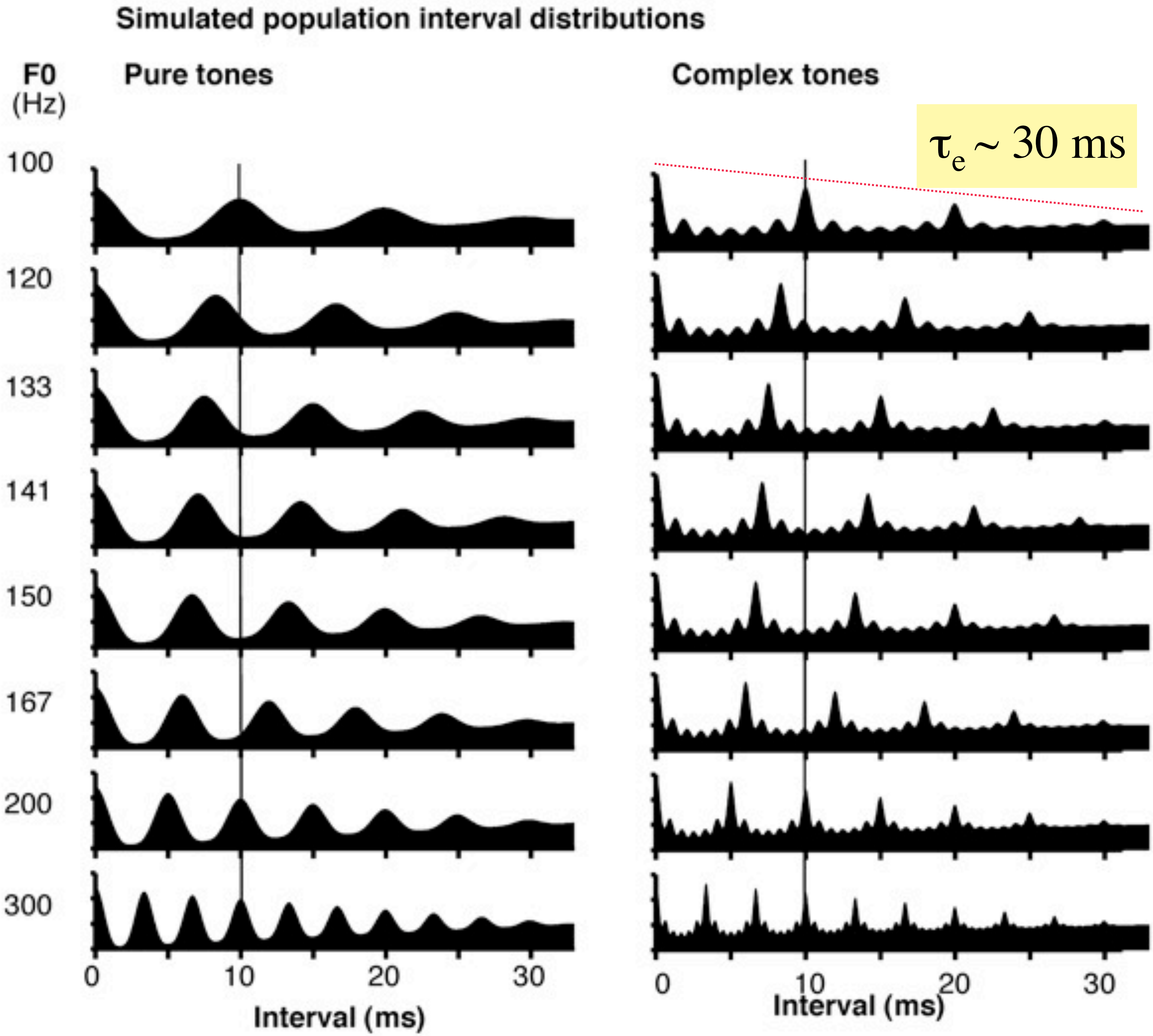
Pitch strength = relative proportion of pitch-related intervals

Timbre (tone quality) = pattern of other intervals

Stimulus autocorrelation ~ population-interval distribution

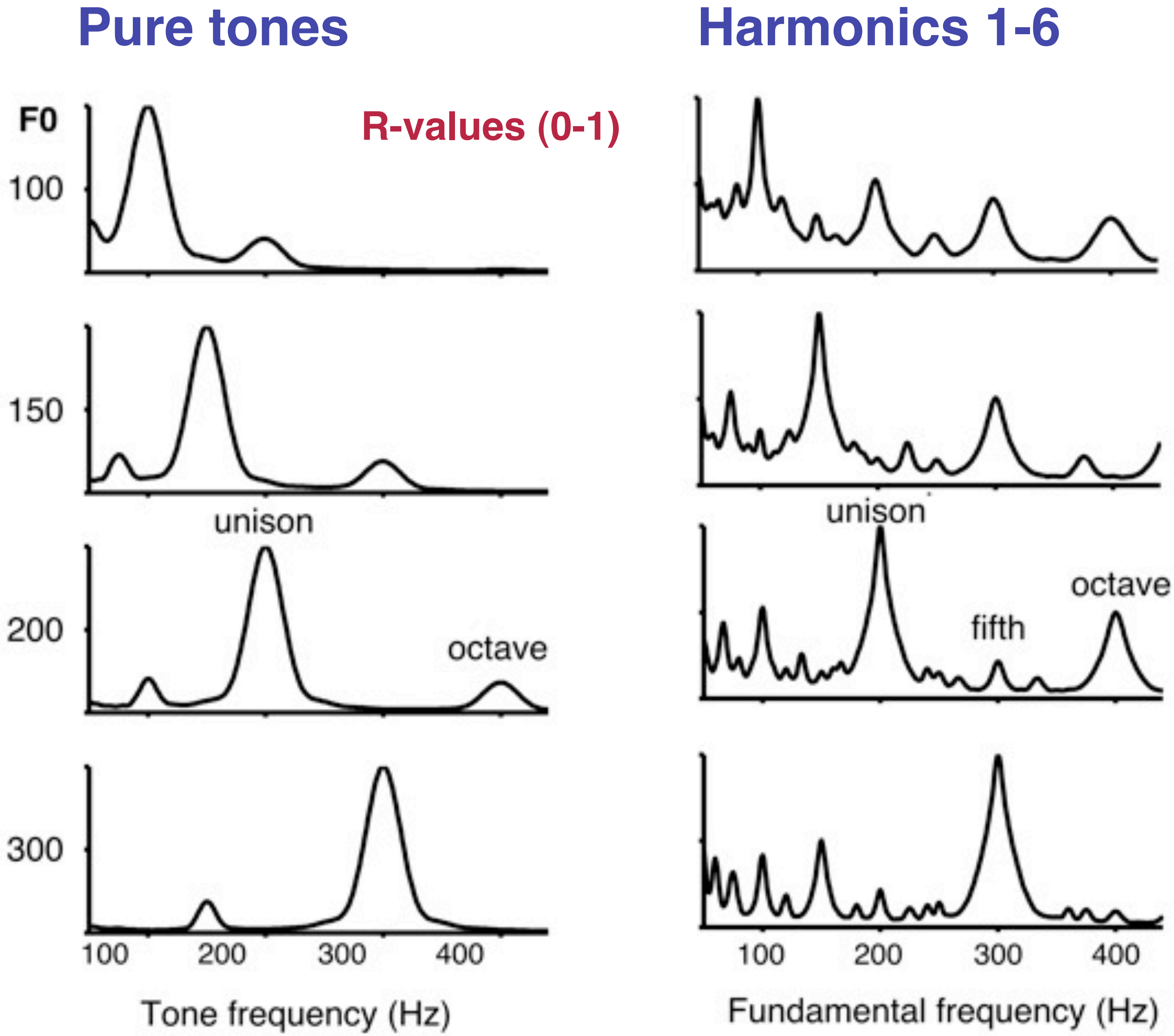
Readily explains:

- Pitch equivalence classes
- Invariance w. Δ sound pressure level
- Invariance w. Δ waveform envelope, phase spectrum
- Existence region of musical tonality (octaves, melody)
- Pitches of resolved & unresolved harmonics
- Pitches of harmonic & inharmonic tone complexes



OCTAVE SIMILARITY FROM SIMILARITY OF INTERSPIKE INTERVAL REPRESENTATIONS

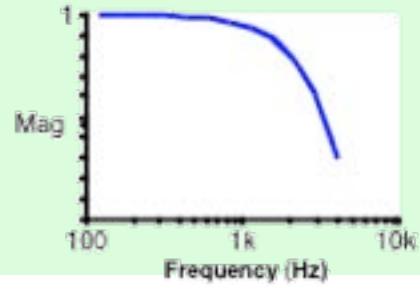
Correlations between population-interval patterns

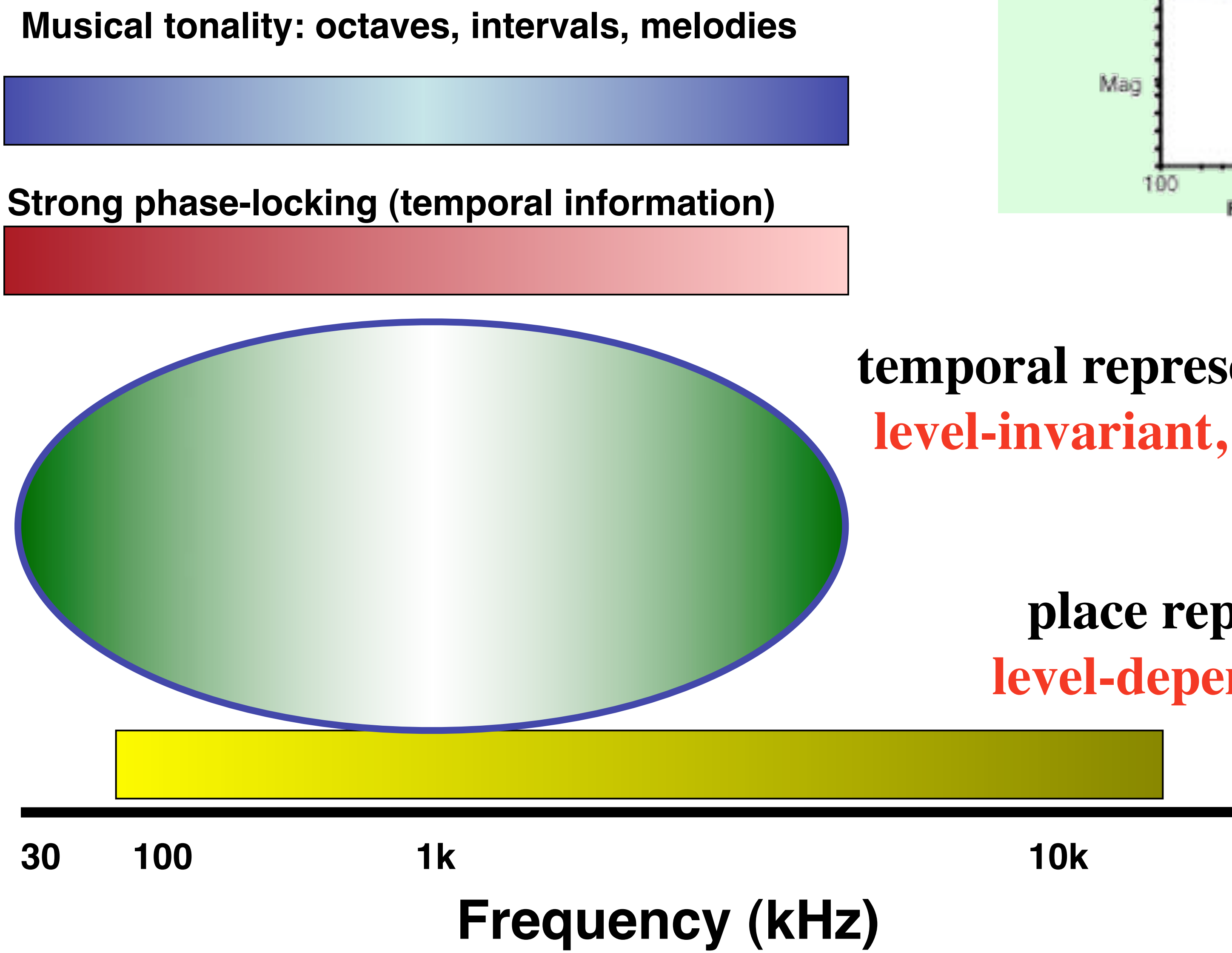


This image is from the article Cariani, P. "Temporal Codes, Timing Nets, and Music Perception." *Journal of New Music Research* 30, no. 2 (2001): 107-135. DOI: 10.1076/jnmr.30.2.107.7115. This journal is available online at http://www.ingentaconnect.com/content/routledg/jnmr/

Existence region of musical tonality is coextensive with spike timing information

Phase-locking as measured by synchronization index declines dramatically above ~4 kHz





Duplex time-place representations

temporal representation

level-invariant

- strong (low fc, low n)
- weak (high fc, high n; F0 < 100 Hz)

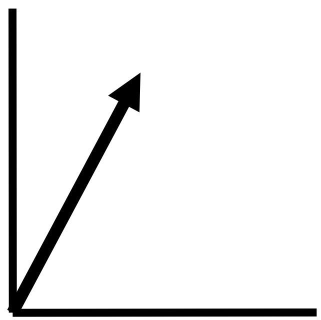
place-based representation

level-dependent

coarse

30 **100** **1k** **10k**

Similarity to interval pattern



Similarity to place pattern

cf. Terhardt's spectral and virtual pitch

Two codes for “pitch

Place code	Time code
Pitch = place of excitation	Pitch = dominant periodicity
Pitch height	Musical pitch "Chroma", Tonality
Absolute	Relational
Low vs. high	Musical intervals, tonality Melodic recognition & transposition
Existence region f: .100-20,000 Hz	Existence region F0: 30-4000 Hz
Frequency analysis **Fourier spectrum**	**Periodicity analysis** **Autocorrelation**

Pitch dimensions: height & chroma

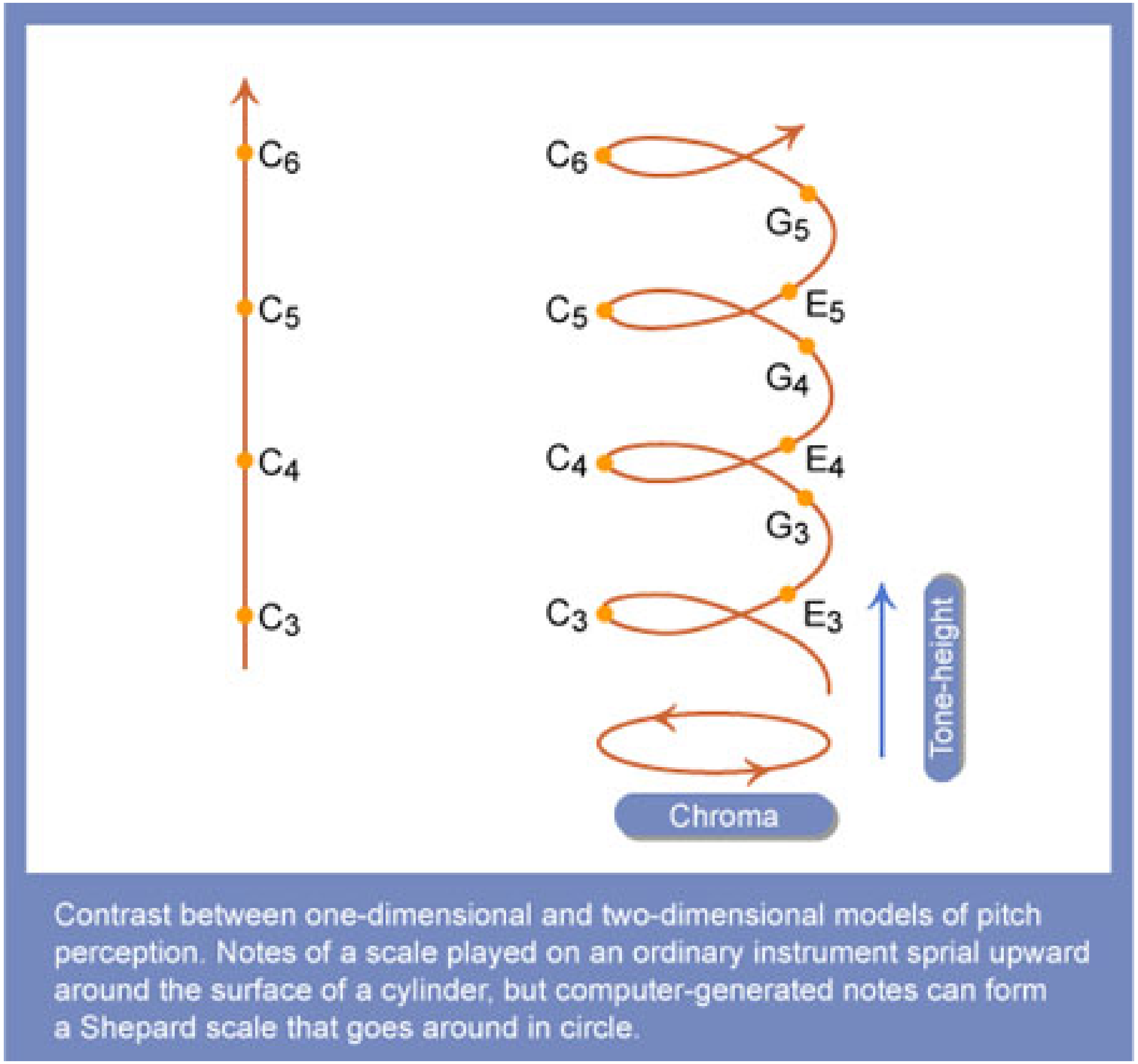


Figure by MIT OpenCourseWare.

Pitch height and pitch chroma

Images removed due to copyright restrictions.
Figures 1, 2, and 7 in Shepard, R. N. "Geometrical Approximations to the Structure of Musical Pitch."
Psychological Review 89, no. 4 (1982): 305-322.

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

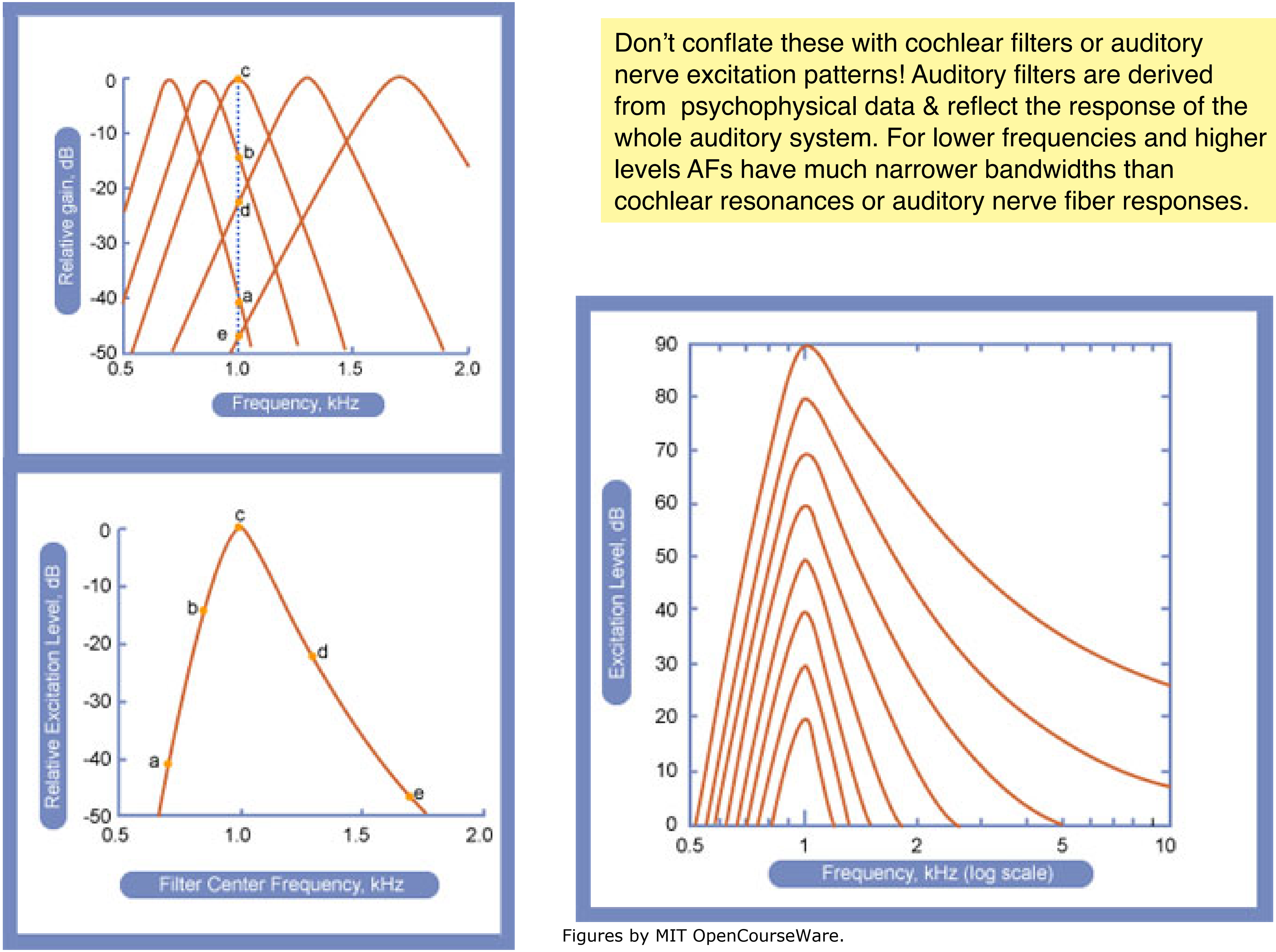
Spectral pattern models

Mostly conceived within a frequency analysis framework

“Auditory filters” derived from psychophysics, not physiology

Shapes of perceptually-derived "auditory filters" (Moore)

Don't conflate these with cochlear filters or auditory nerve excitation patterns! Auditory filters are derived from psychophysical data & reflect the response of the whole auditory system. For lower frequencies and higher levels AFs have much narrower bandwidths than cochlear resonances or auditory nerve fiber responses.

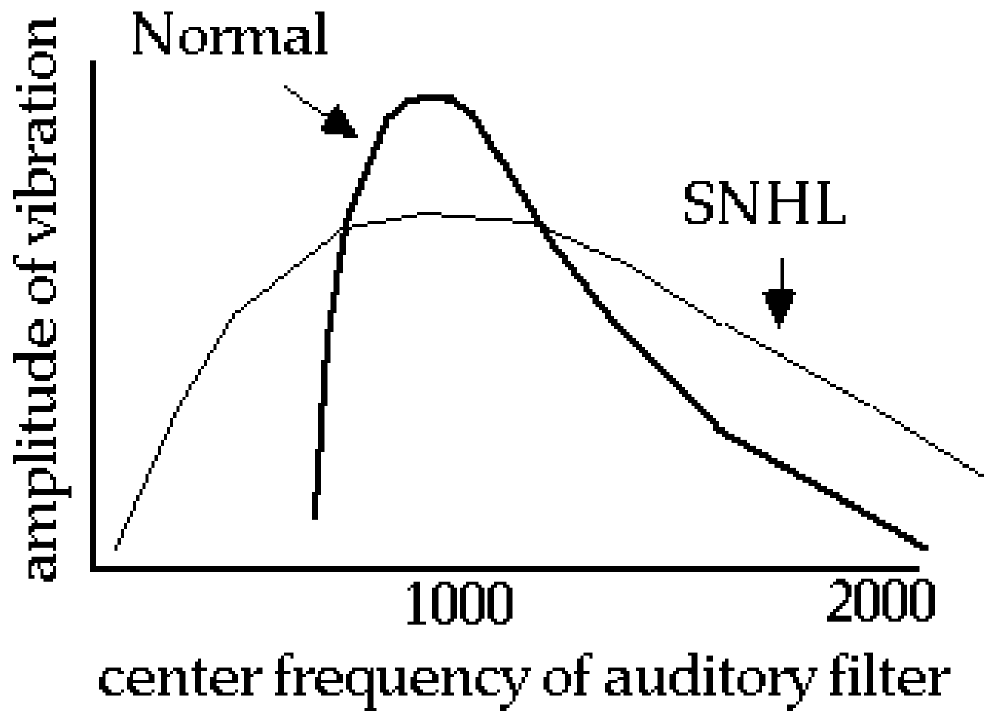


Figures by MIT OpenCourseWare.

From masking patterns to "auditory filters" as a model of hearing

Power spectrum
Filter metaphor

Notion of one central spectrum that subserves



2.2. Excitation pattern Using the filter shapes and bandwidths derived from masking experiments we can produce the excitation pattern produced by a sound. The excitation pattern shows how much energy comes through each filter in a bank of auditory filters. It is analogous to the pattern of vibration on the basilar membrane. For a 1000 Hz pure tone the excitation pattern for a normal and for a SNHL (sensori-neural hearing loss) listener look like this: The excitation pattern to a complex tone is simply the sum of the patterns to the sine waves that make up the complex tone (since the model is a linear one). We can hear out a tone at a particular frequency in a mixture if there is a clear peak in the excitation pattern at that frequency. Since people suffering from SNHL have broader auditory filters their excitation patterns do not have such clear peaks. Sounds mask each other more, and so they have difficulty hearing sounds (such as speech) in noise. --Chris Darwin, U. Sussex, http://www.biols.susx.ac.uk/home/Chris_Darwin/Perception/Lecture_Notes/Hearing3/hearing3.html

Resolvability of harmonics (Plomp, 1976)

Fundamental
frequency
(F0)

Image removed due to copyright restrictions.
Graph of frequency separation between partials vs. frequency of the partial.
From Plomp, R. *Aspects of Tone Sensation*. New York, NY: Academic Press, 1976.

Resolution of harmonics (based on psychophysics)

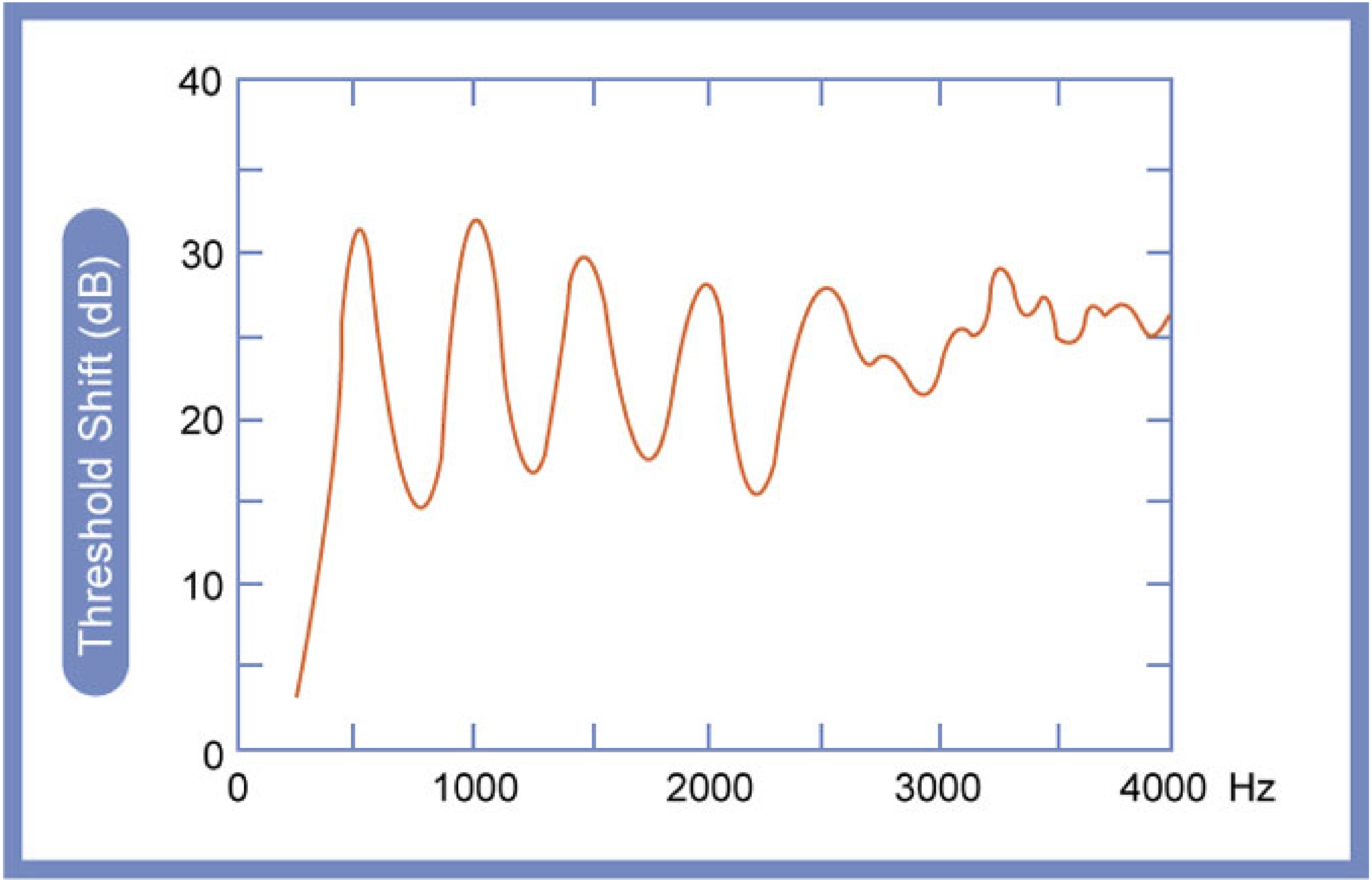
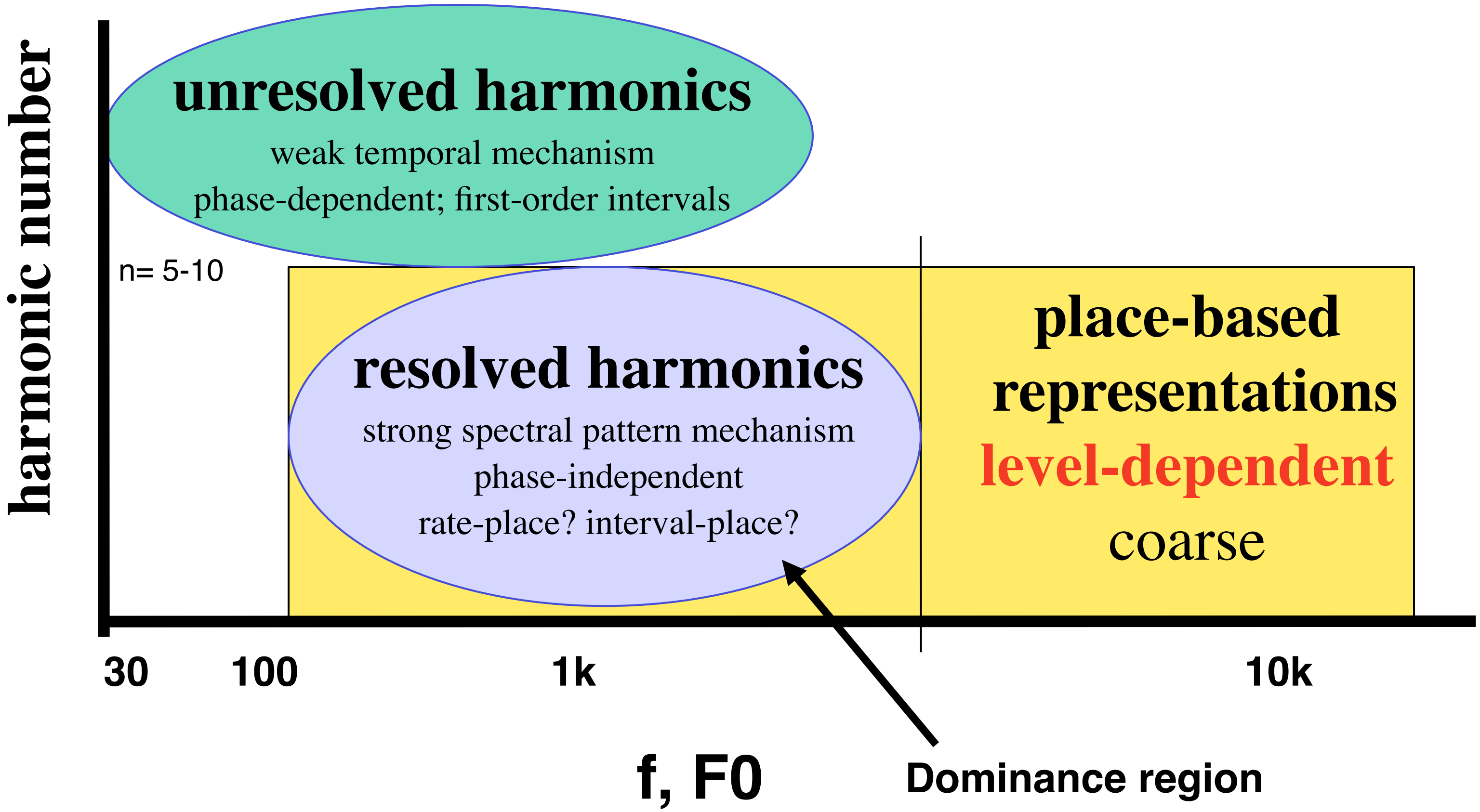


Figure by MIT OpenCourseWare.

Goldstein's harmonic templates

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

A "two-mechanism" perspective (popular with some psychophysicists)



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 component
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

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.

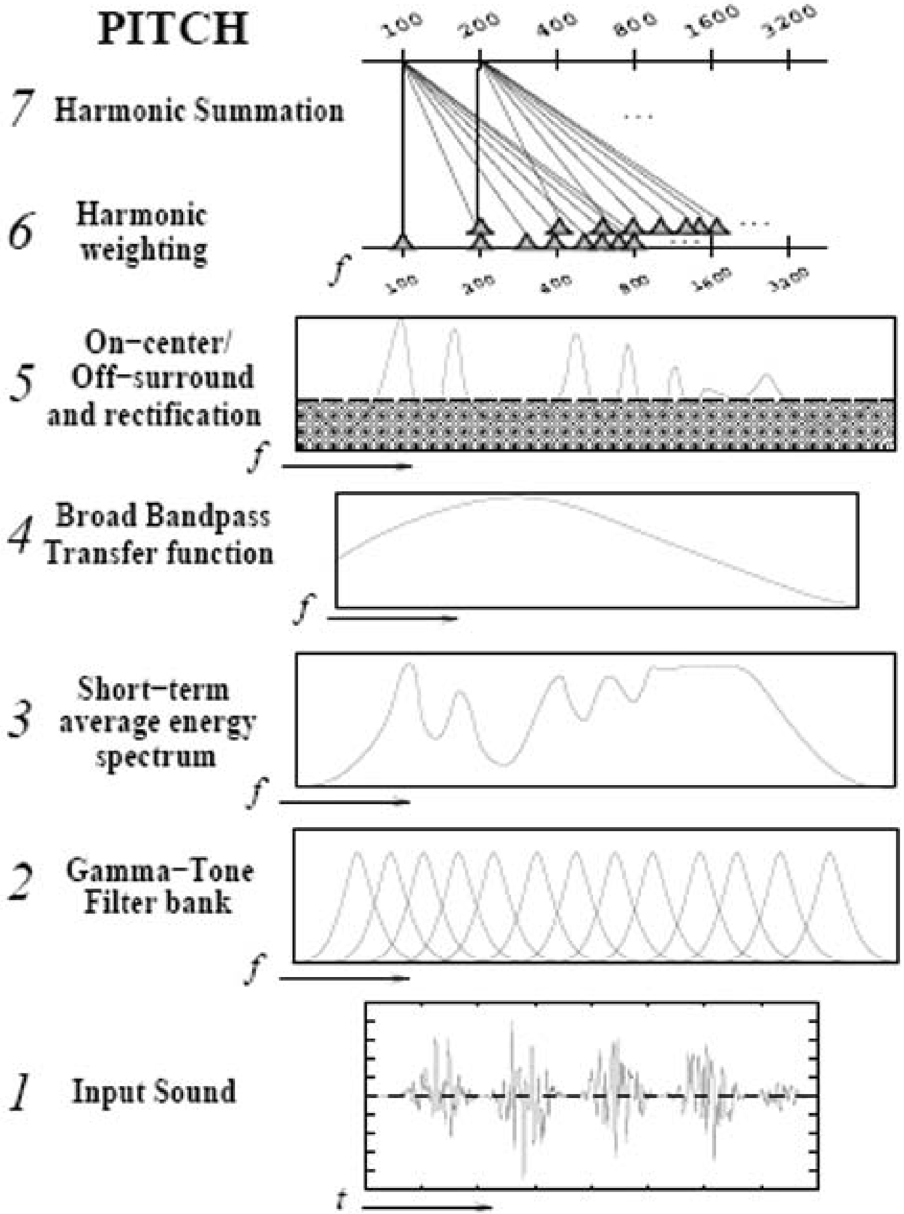
Terhardt E (1984) The concept of musical consonance: a link between music and psychoacoustics. Music Perception 1:276-295.

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.

SPINET:
Cohen Grossberg, Wyse JASA

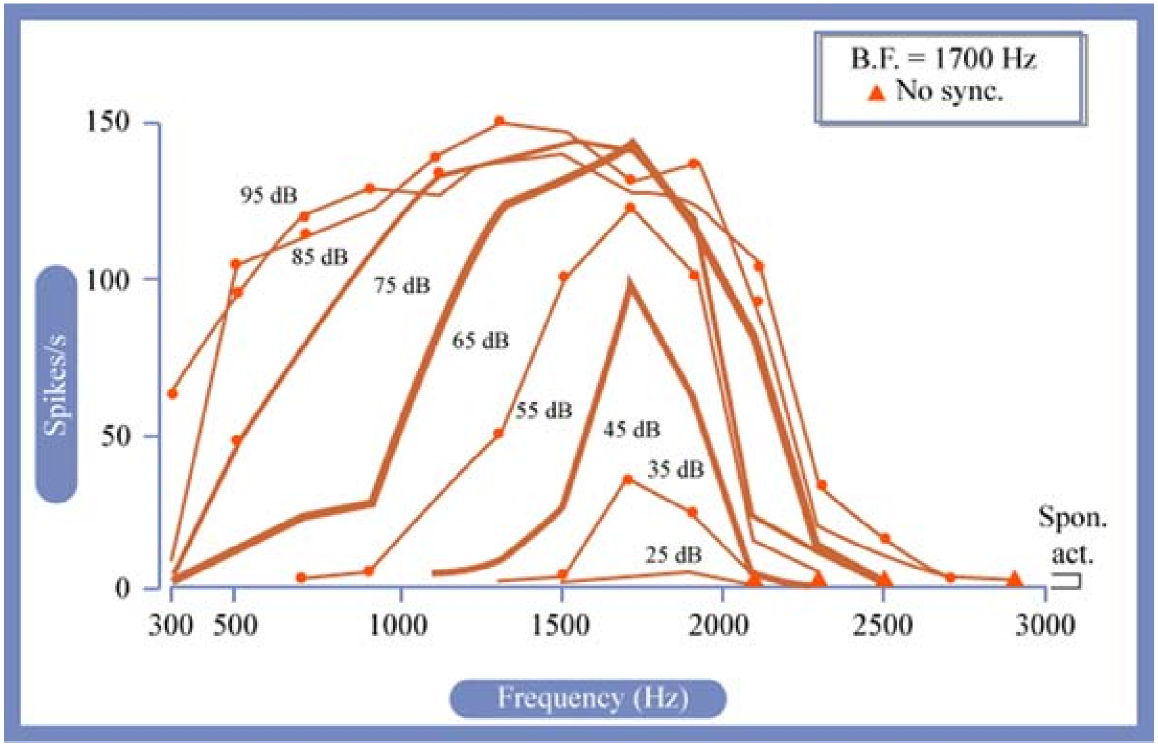
Fixed neural network: connection weights arranged so as to form pitch-equivalence classes



Courtesy of Prof. Stephen Grossberg. Used with permission.
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.

Broad tuning and rate saturation at moderate levels in low-CF auditory nerve fibers confounds rate-based resolution of harmonics.

Low SR auditory nerve fiber



Rose, 1971

Figure by MIT OpenCourseWare.

Spectral pattern theories - pros & cons

Do make use of frequency tuning properties of
elements in the auditory system
No clear neural evidence of narrow (< 1/3 octave)
frequency channels in low-BF regions (< 2 kHz)

Operate on perceptually-resolved harmonics
Do not explain low pitches of unresolved harmonics

Require templates or harmonic pattern analyzers
Little or no neural evidence for required analyzers
Problems w. templates: relative nature of pitch

Do not explain well existence region for F0
Learning theories don't account for F0 ranges
or for phylogenetic ubiquity of periodicity pitch

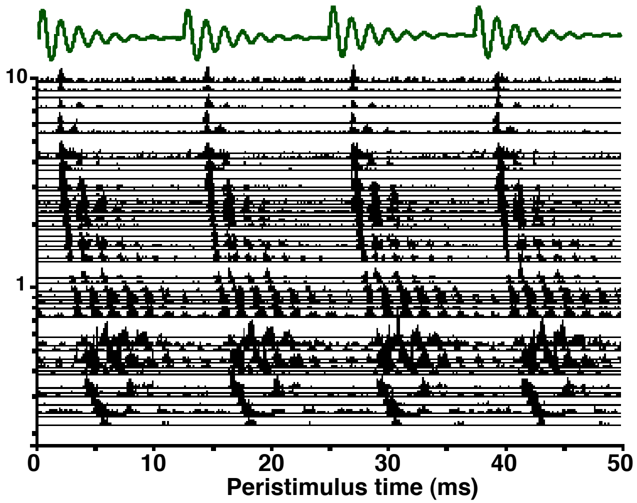
Problems with rate-place models

In contrast to musical (F0) pitch percepts....

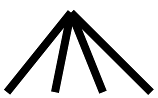
Rate-place spectral profiles

- have coarse resolution ($\geq$ 1 octave)
- change with sound level
- worsen dramatically at higher sound levels
- should work better for high frequencies
- cannot account for F0 pitches of unresolved harmonics

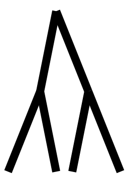
Some possible auditory representations



All-at-once



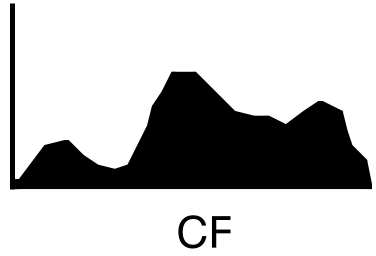
Stages of integration



Local

Rate-place

Central spectrum

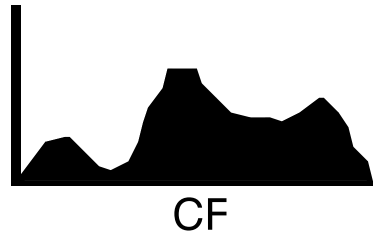


Masking phenomena
Loudness

Synchrony-place

Central spectrum

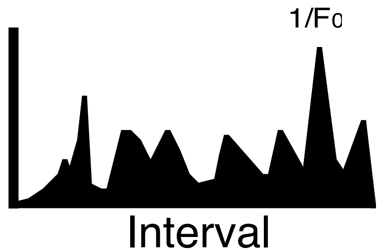
Interval-place



Pure tone pitch JNDs

Population interval

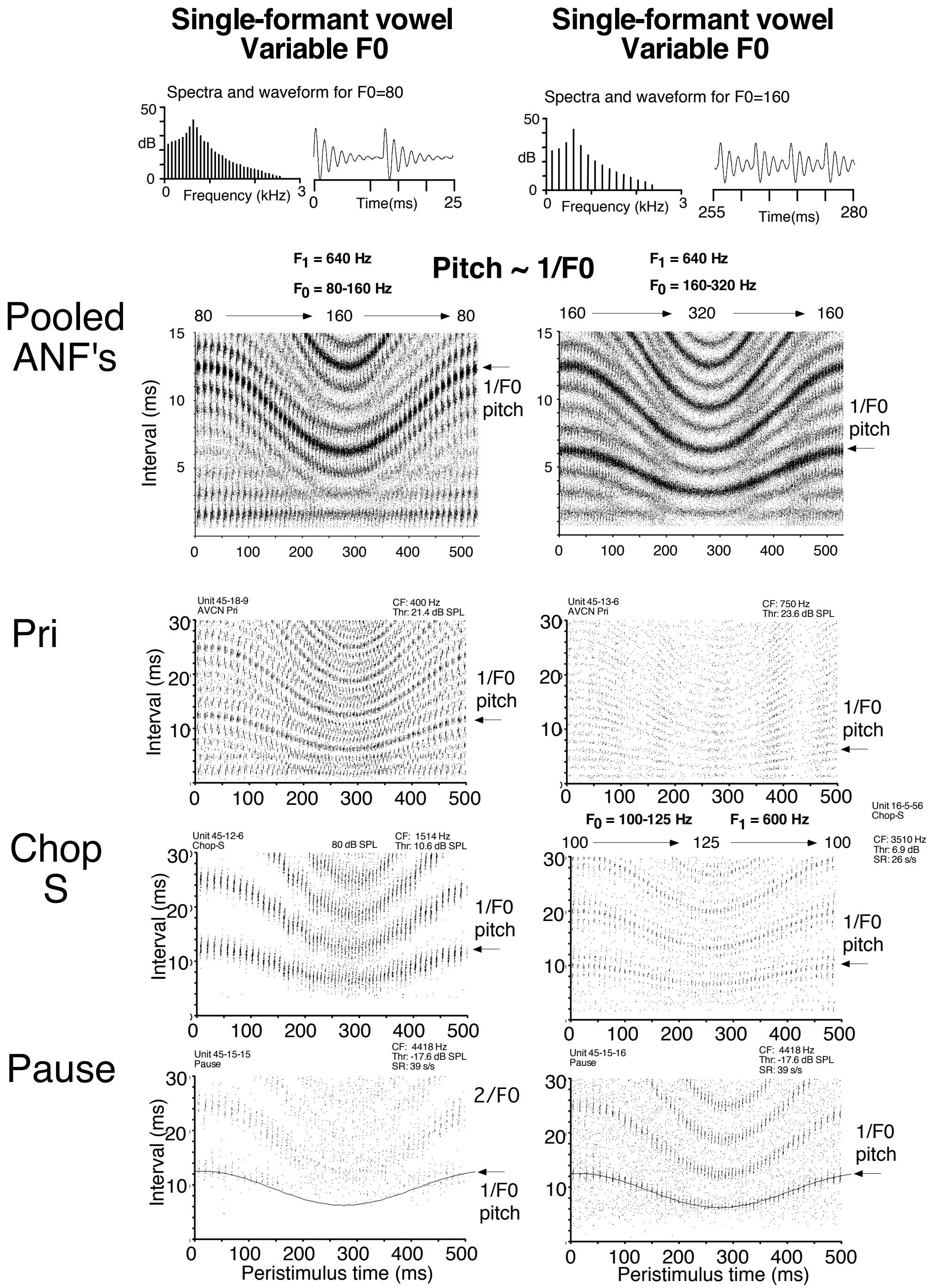
Population-interval



Complex tone pitch

Global

Cochlear nucleus II



Modulation detectors in the midbrain

Stimulus-related temporal discharge patterns

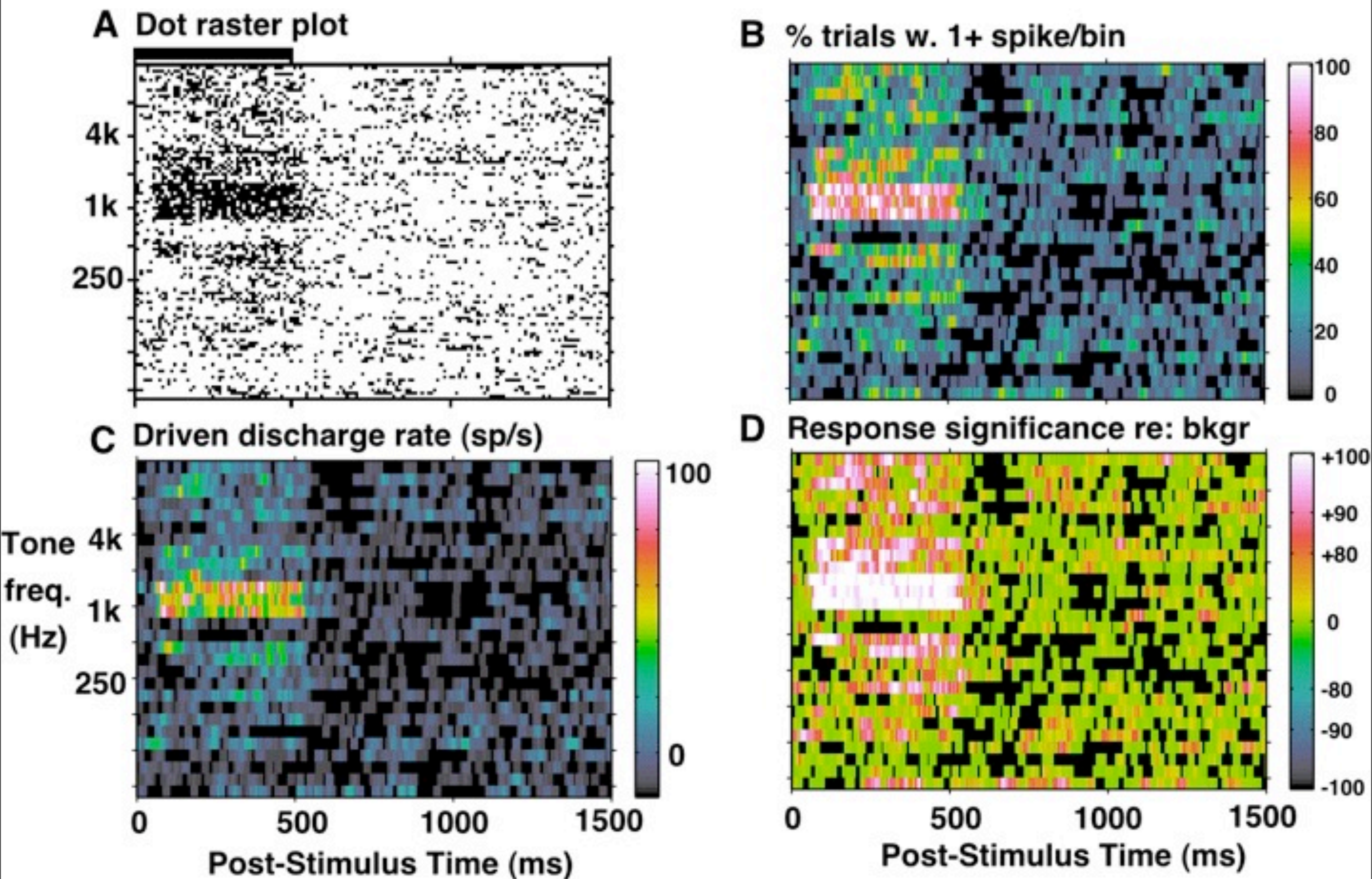
Modulation-tuning
of discharge rate

Images removed due to copyright restrictions.
See Fig. 2 and 3 in Langner, G. and C. E. Schreiner.
"Periodicity Coding in the Inferior Colliculus of the Cat. I. Neuronal Mechanisms."
J Neurophysiol 60 (1988): 1799-1822.

See: 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 Press.
Langner, G. and Schreiner, C.E. Periodicity coding in the inferior colliculus of the cat. I. Neuronal mechanisms.
J. Neurophysiol. 60:1799-1822.
Langner (1992) review, Periodicity coding in the auditory system. *Hearing Research*, 60:115-142.

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

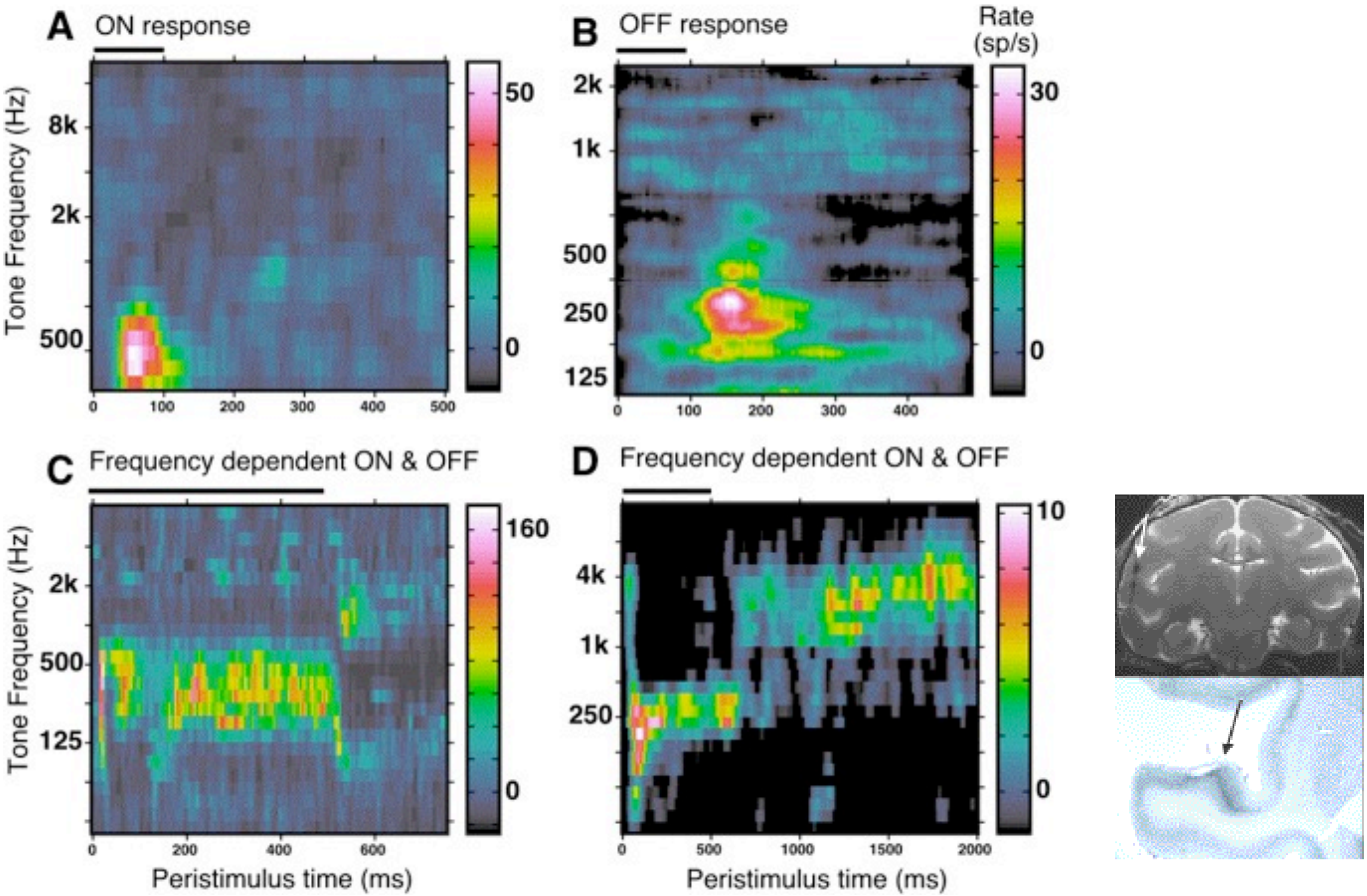
Figure. Averaged cortical field potentials (current source density analysis, lower lamina 3, site BF=5 kHz) in response to 50 ms click trains F0=100-500 Hz. Ripples up to 300-400 Hz show synchronized component of the ensemble-response, From Steinschneider (1999).



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.

Pure tone temporal response profiles in auditory cortex (A1)



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.

Phase-locking in thalamus and cortex

De Ribaupierre: 10% of thalamic units in awake cats with synchronization indices of 0.3 or better to 1-2 kHz tones

Reports of isolated cortical units with phase-locking to Fm of AM tones up to 1 kHz (Semple)

Pitch detectors in thalamus and cortex

Schwarz & Tomlinson failed to find true F0 detectors in their study of > 200 cortical units in awake macaque

Riquimaroux found 16 units that responded both to pure tones and harmonic complexes that would evoke the same pitch

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.

The enigma of the central representation of pitch

The tight correspondences between psychophysics and the population interspike interval code **strongly** suggest that our perception of pitch depends on this information.

Yet, despite some recent advances, we still do not understand how the central auditory system uses this information.

What happens to neural timing information as one ascends the auditory pathway from auditory nerve to cortex?

Is time converted to some sort of place code (e.g. pitch detectors) or is some other kind of code involved?

Discharge rate as a function of frequency and intensity

Auditory nerve fiber

Rose (1971)

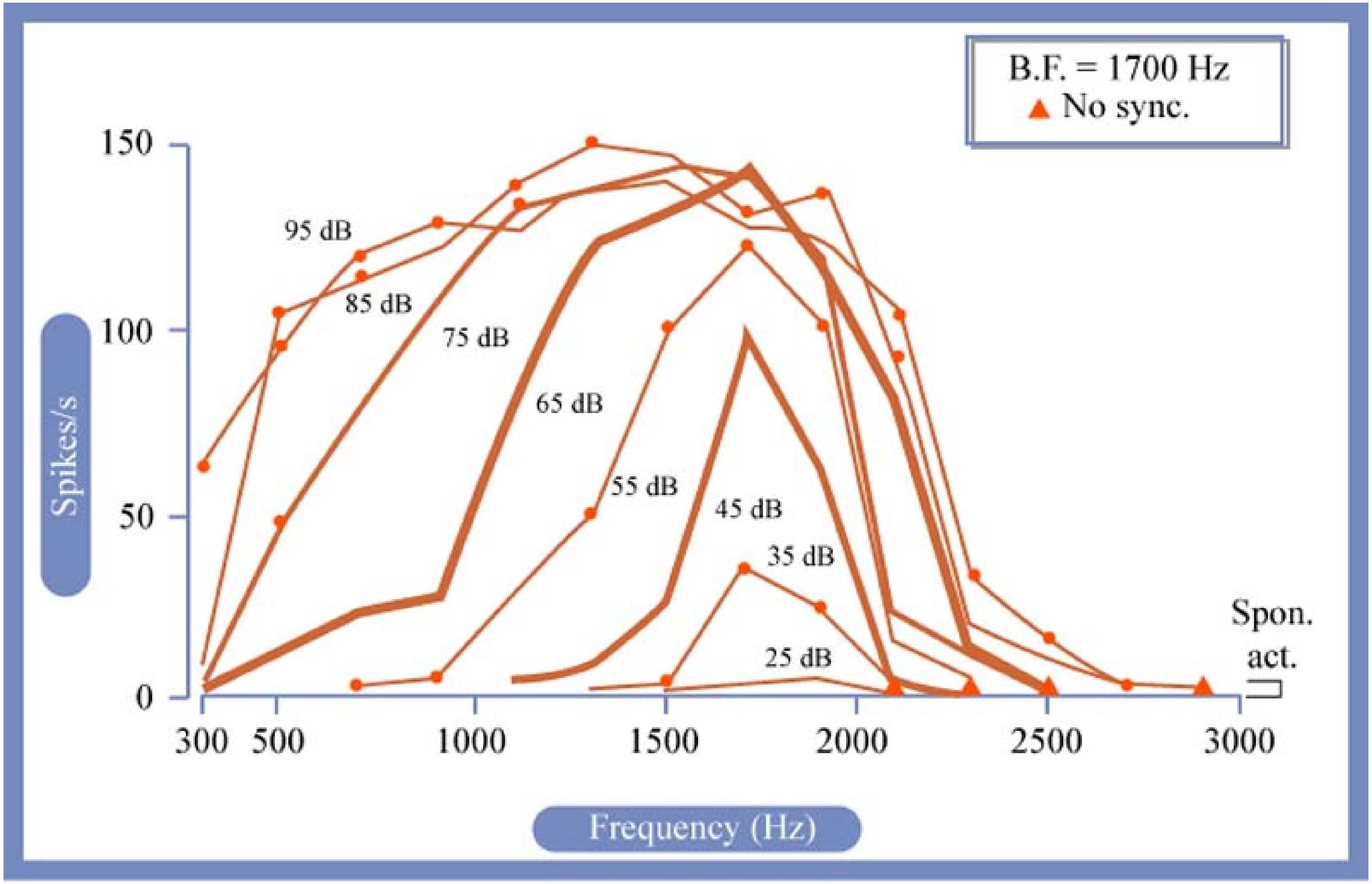


Figure by MIT OpenCourseWare.

Four graphs removed due to copyright restrictions.
See Fig. 2 and Fig. 4 in Rose, J. E., et al. “Some Effects of Stimulus Intensity on Response of Auditory Nerve Fibers in the Squirrel Monkey.” *J Neurophysiol* 34 (1971): 685–699.

Frequency response curves for fibers in the cochlear nerve of the squirrel monkey.
Left: low spontenous activity; right: high spontaneous activity.

From Rose, J. E., et al. “Some Effects of Stimulus Intensity on Response of Auditory Nerve Fibers in the Squirrel Monkey.” *J Neurophysiol* 34 (1971): 685–699.

Alan Palmer
In Hearing, Moore ed.

Figure 4 in Palmer, Alan. "Neural Signal Processing." Chapter 3 in *Hearing*. 2nd ed. Edited by B. C. J. Moore. Academic Press, 1995. [Preview this image in Google Books]

HST.725 Music Perception and Cognition
Spring 2009