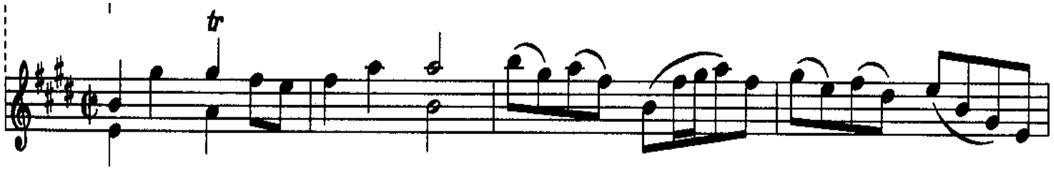
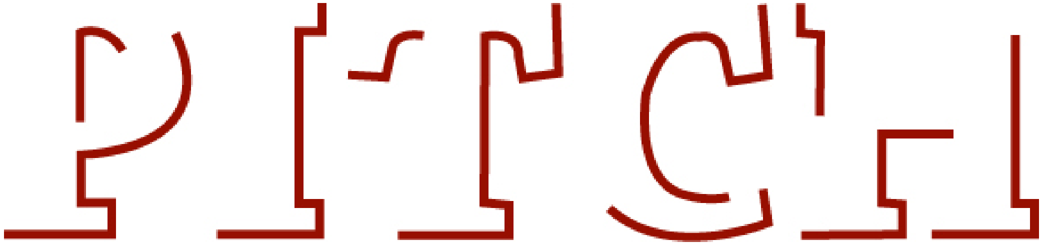
Harvard-MIT Division of Health Sciences and Technology
HST.725: Music Perception and Cognition
Prof. Peter Cariani

Theories & models of musical pitch

(Image removed due to copyright considerations.)



www.cariani.com

The search for the missing fundamental: theories & models of musical pitch

- Brief review of basics of sound & vibration
- Brief review of pitch phenomena
- 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)
  - Population-interval models (Meddis & Hewitt, Cariani & Delgutte)
- Problems & prospects

Vibrations create compressions and expansions of air

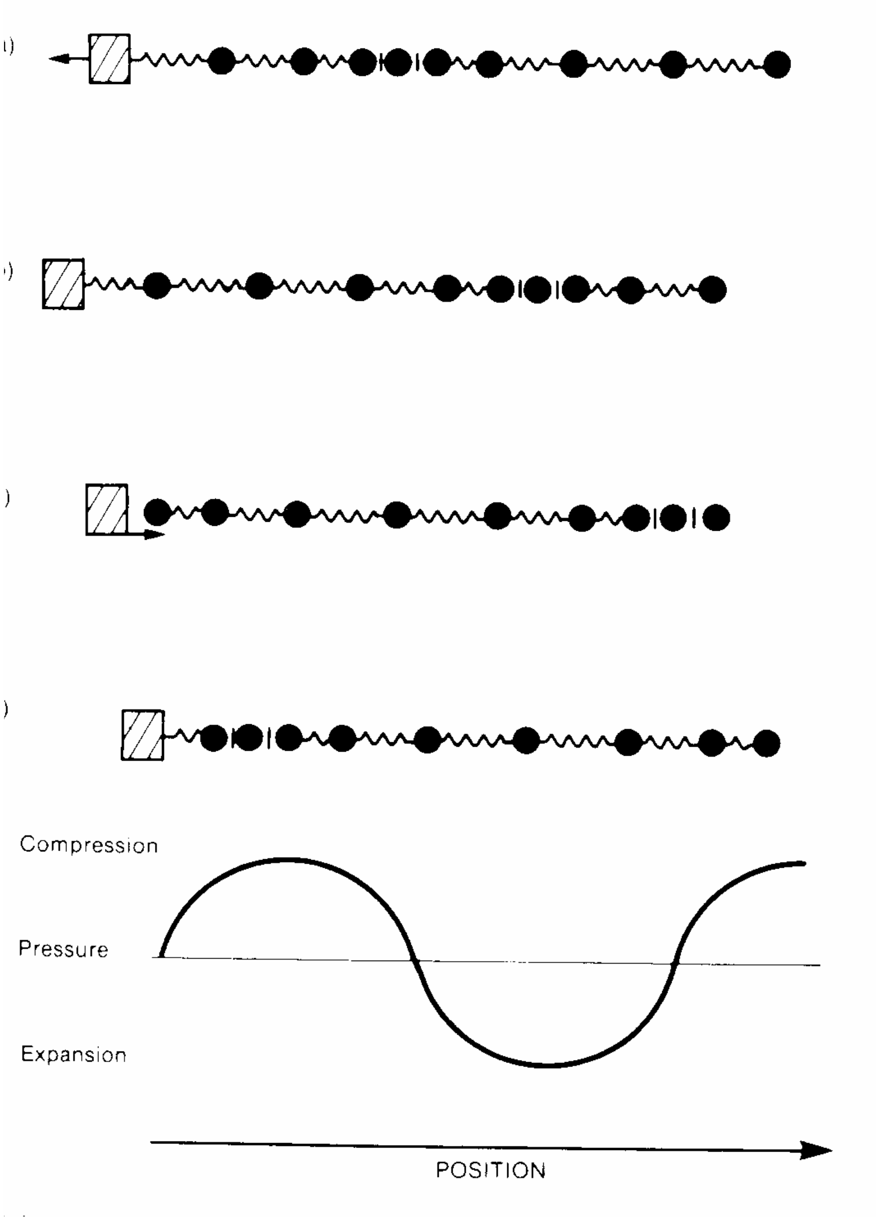
(Series of figures from Handel, S. 1989. Listening: an Introduction to the Perception of Auditory Events. MIT Press. Used with permission.)

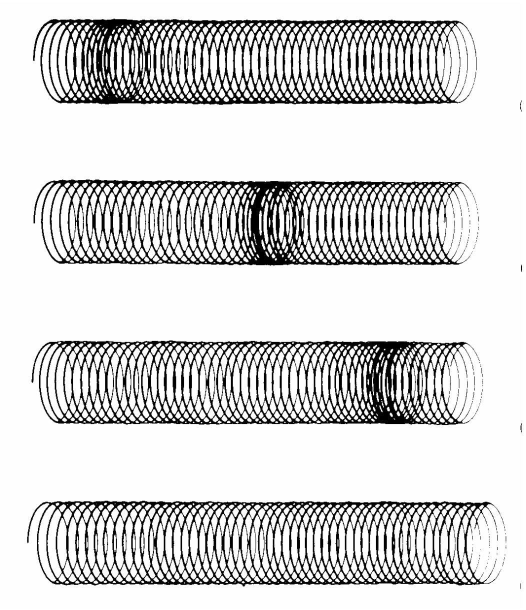
Sound waves are alternating local changes in pressure

These changes propagate through space as “longitudinal” waves

Condensation phase **(compression)**: pressure increases

Rarefaction **(expansion)** **phase**: pressure decreases



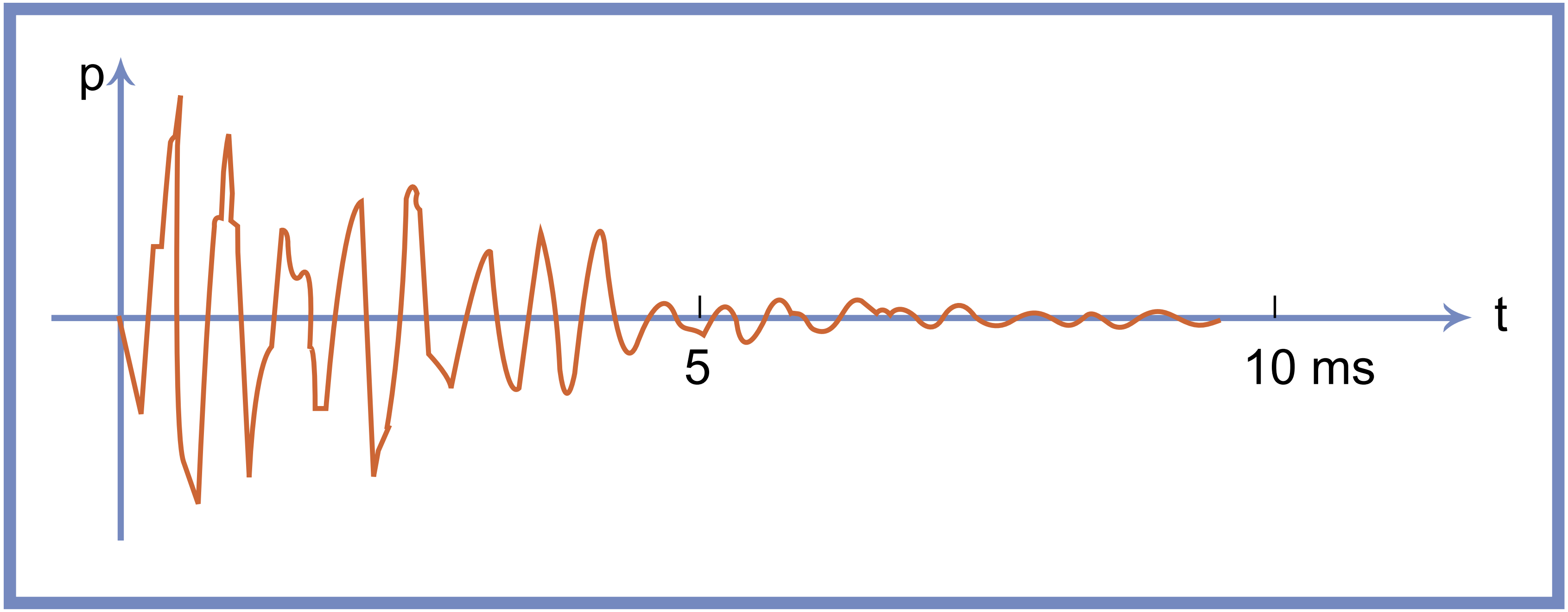


Waveforms

Microphones convert sound pressures to electrical voltages.
Waveforms plot pressure as a function of time, i.e. a "time-series" of amplitudes.
Waveforms are complete descriptions of sounds.

Audio CD's sample sounds at 44,100 samples/sec.

Oscilloscope demonstration.



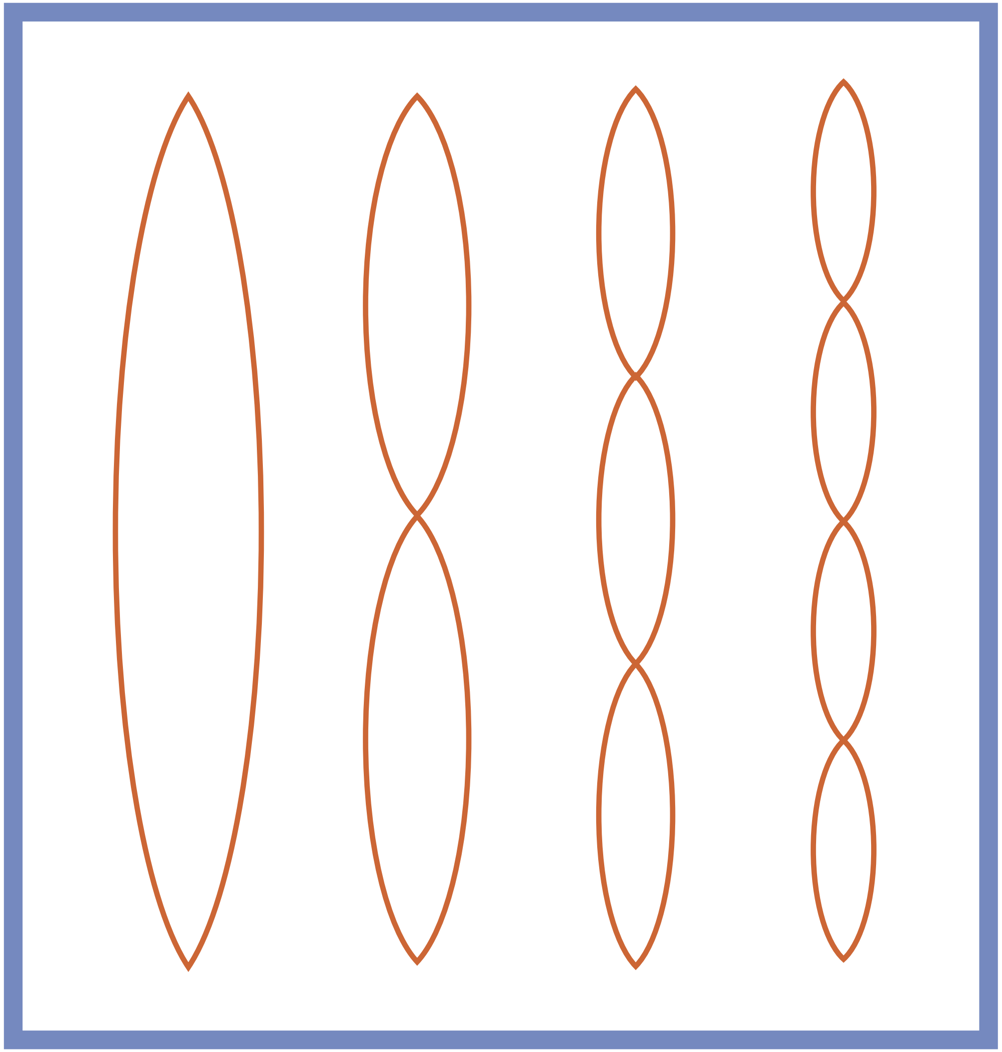
Complex modes of vibration

Most physical systems have multiple modes of vibrations that create resonances that favor particular sets of frequencies.

Vibrating strings or vibrating columns of air in enclosures exhibit harmonic resonance patterns.

Material structures that are struck (bells, xylophones, percussive instruments) have resonances that depend partly on their shape and therefore canproduce frequencies that are not harmonically related.

More later on what this means for pitch and sound quality.



Frequency spectra

Joseph Fourier (1768-1830) showed that any waveform can be represented as the sum of many sinusoids (*Fourier spectrum*).

George Ohm (1789-1854) and Hermann von Helmholtz (1821-1894) postulated that the ear analyzes sound analogously, first breaking sounds into their partials.

Each sinusoid of a particular frequency (*frequency component, partial*) has 2 parameters:

- 1) its magnitude (amplitude of the sinusoid)
- 2) its phase (relative starting time)

A sound with only one frequency component is called a *pure tone*.

A sound with more than one is called a *complex tone*.

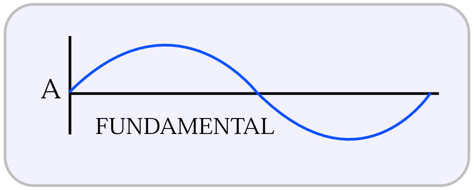
A complex tone whose partials are all part of the same harmonic series is called a *harmonic complex*.

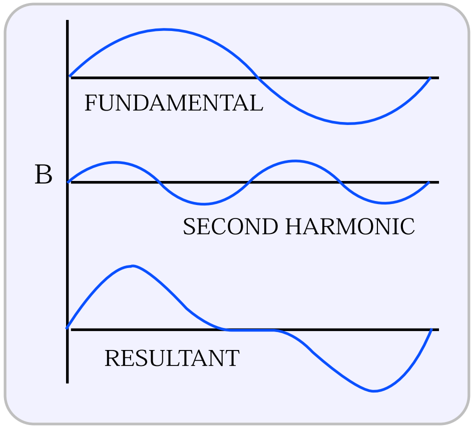
Such a tone is periodic -- its waveform repeats with a period equal to 1/fundamental frequency (i.e. the fundamental period).

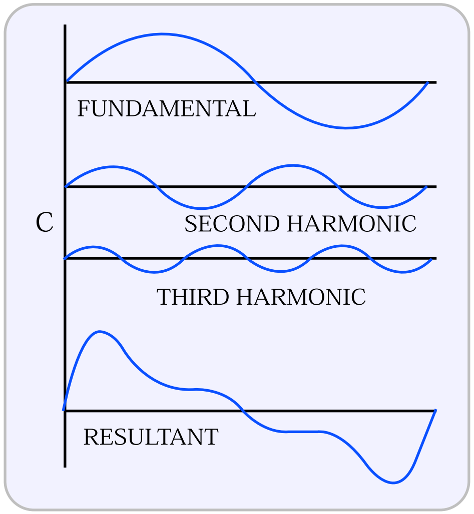
If any of the partials are not part of a harmonic series, then the sound is *inharmonic*.

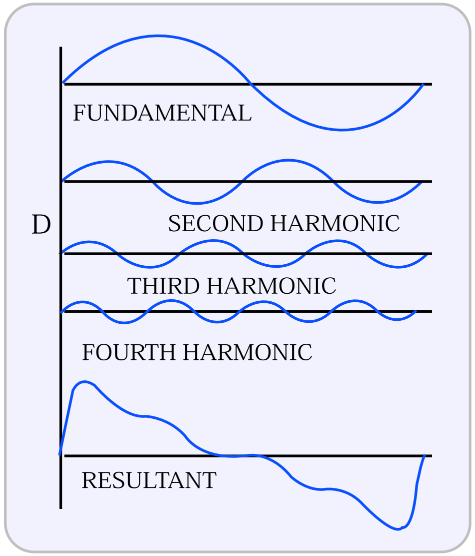
Fundamentals and harmonics

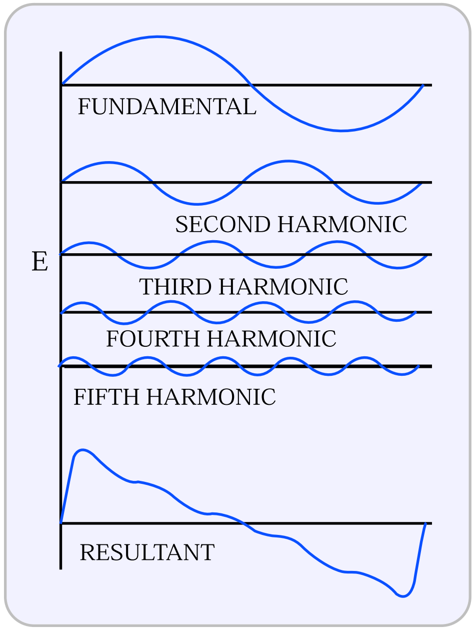
- Periodic sounds (30-20kHz) produce pitch sensations.
- Periodic sounds consist of repeating time patterns.
- The fundamental period (F0) is the duration of the repeated pattern.
- The fundamental frequency is the repetition frequency of the pattern.
- In the Fourier domain, the frequency components 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.











Demonstrations

- **Oscilloscope demonstrations**
- Armenian flute
- Human voice

- **Spectrum analyzer demonstrations**
- Flute
- Violins
- Human voice

- **Real time spectrogram demonstrations (iTunes/SpectroGraph plugin)**
- Armenian flute
- Violins
- Vocal music

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*.

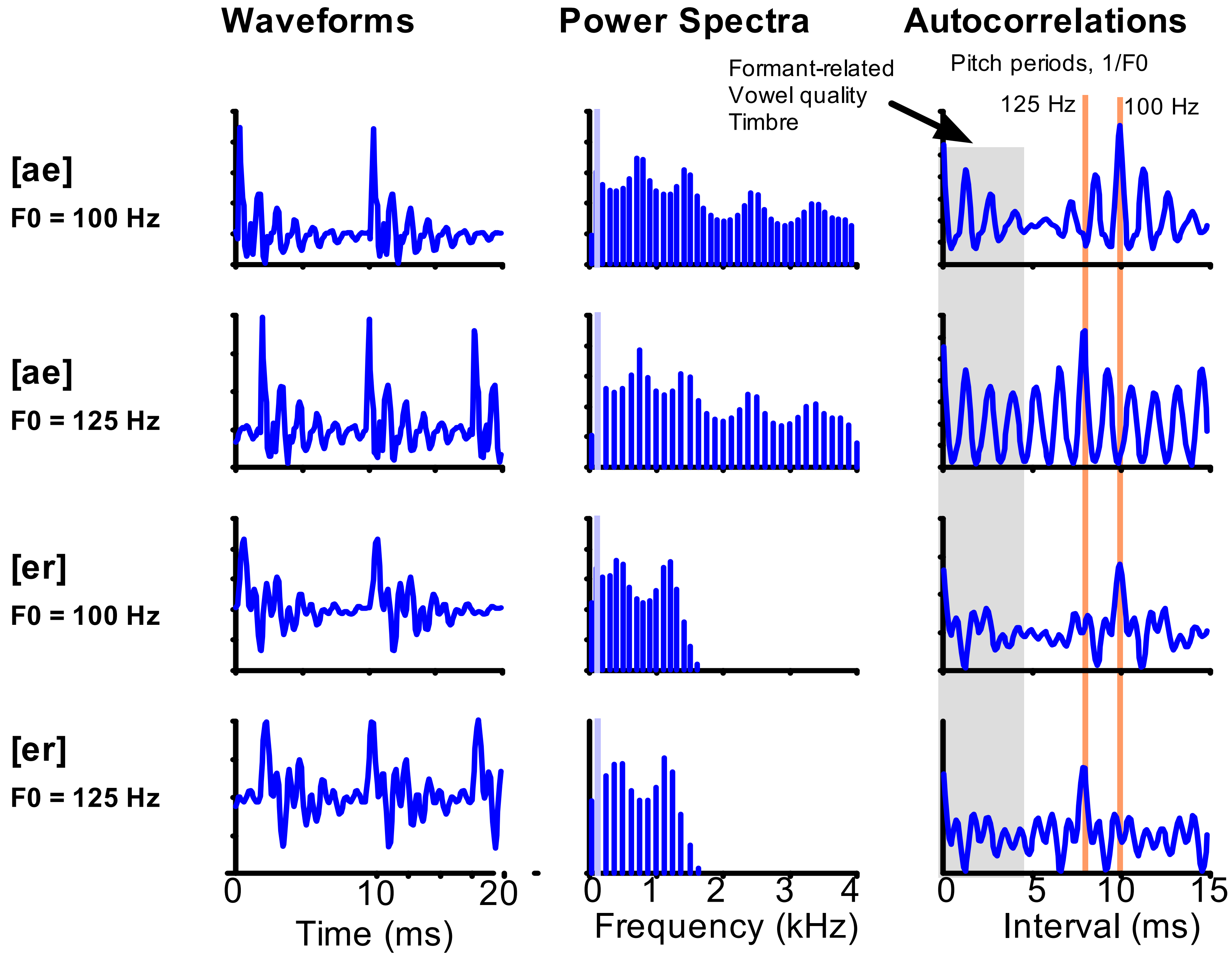
Subharmonics are integer divisions of the fundamental:

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 msec. Periods of the subharmonics are integer multiples of the fundamental period.

Synthetic vowels



Pitch : basic properties

- **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

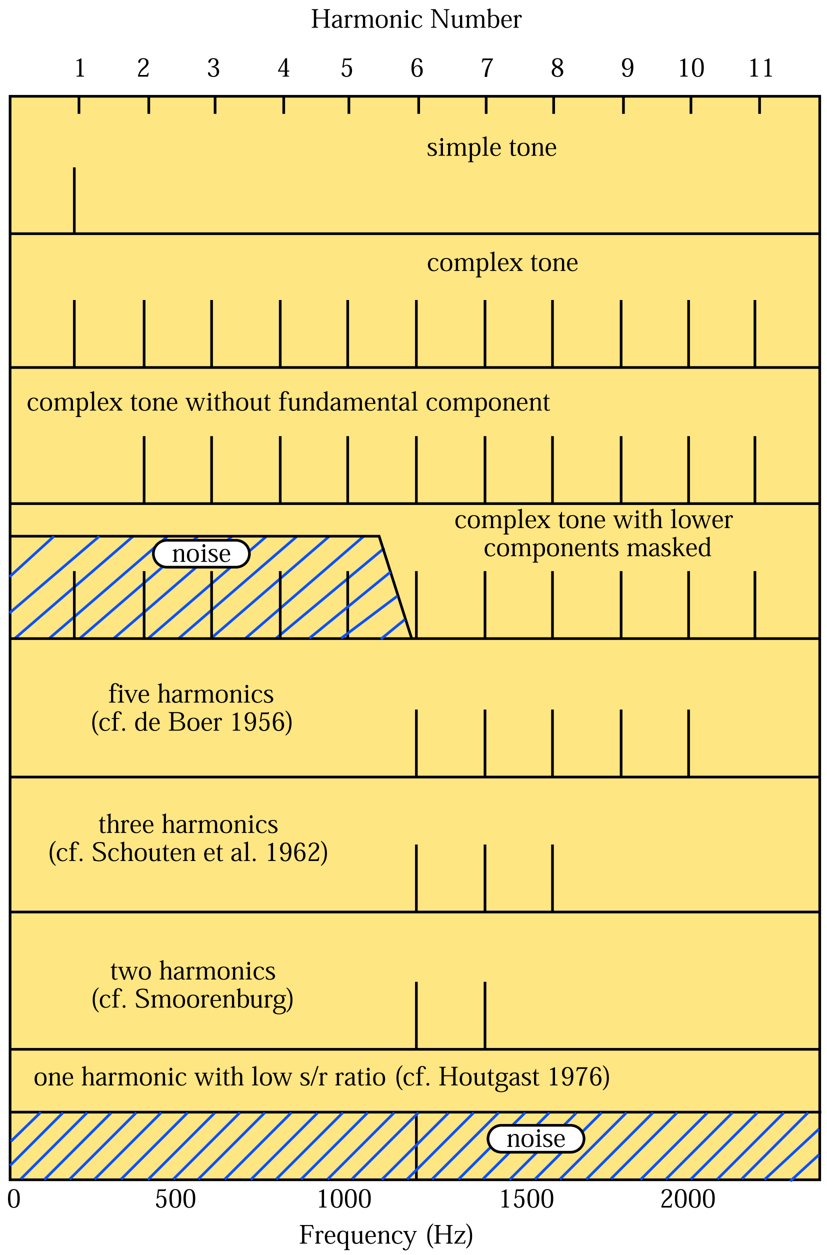
Strong
- Pure tones
- Harmonic complexes
- Iterated noise

Weaker
- High harmonics
- Narrowband noise

Very weak
- AM noise
- Repeated noise

Strong pitches

Weaker low pitches



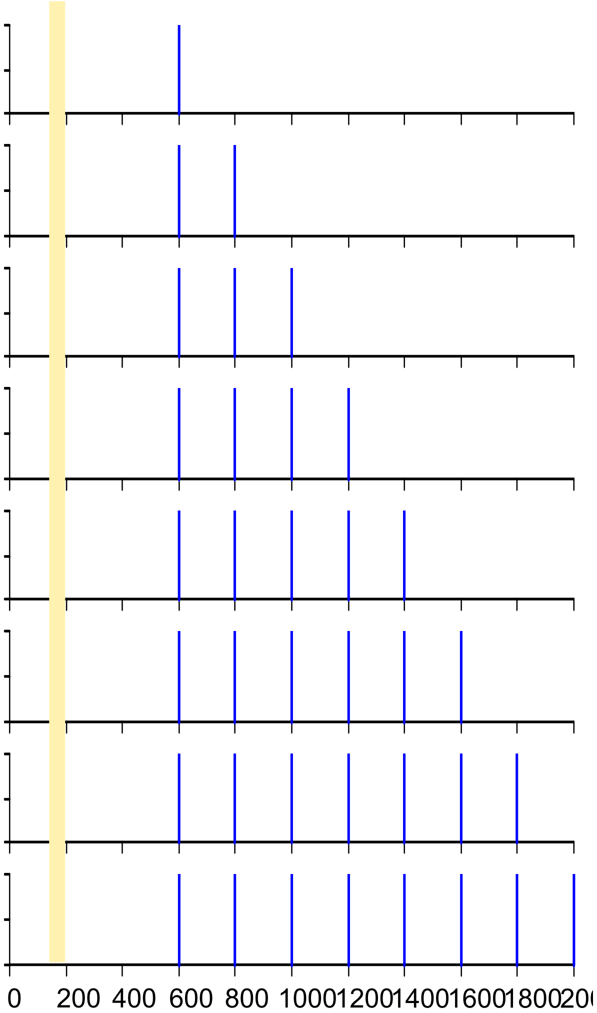
Schematic diagram representing eight signals with the same low pitch.

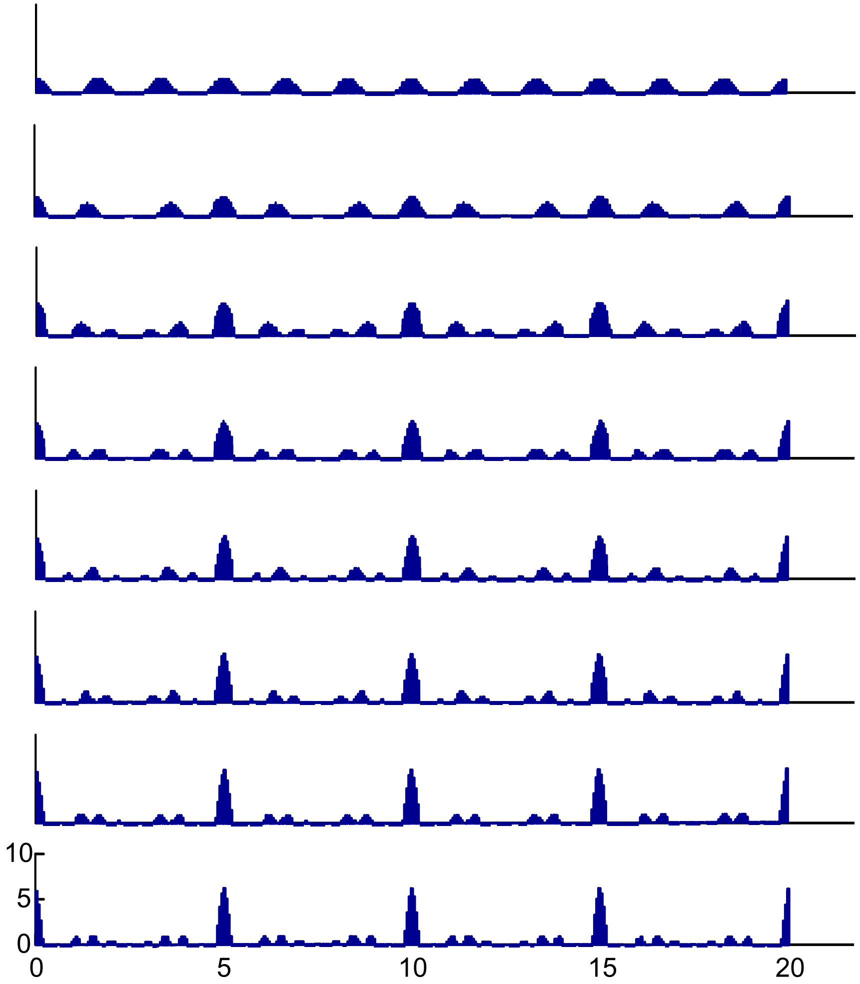
Pitch as a perceptual emergent

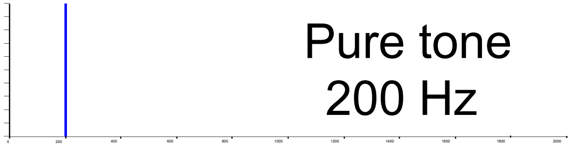
Missing F0

Line spectra

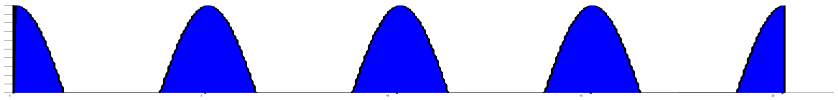
Autocorrelation (positive part)



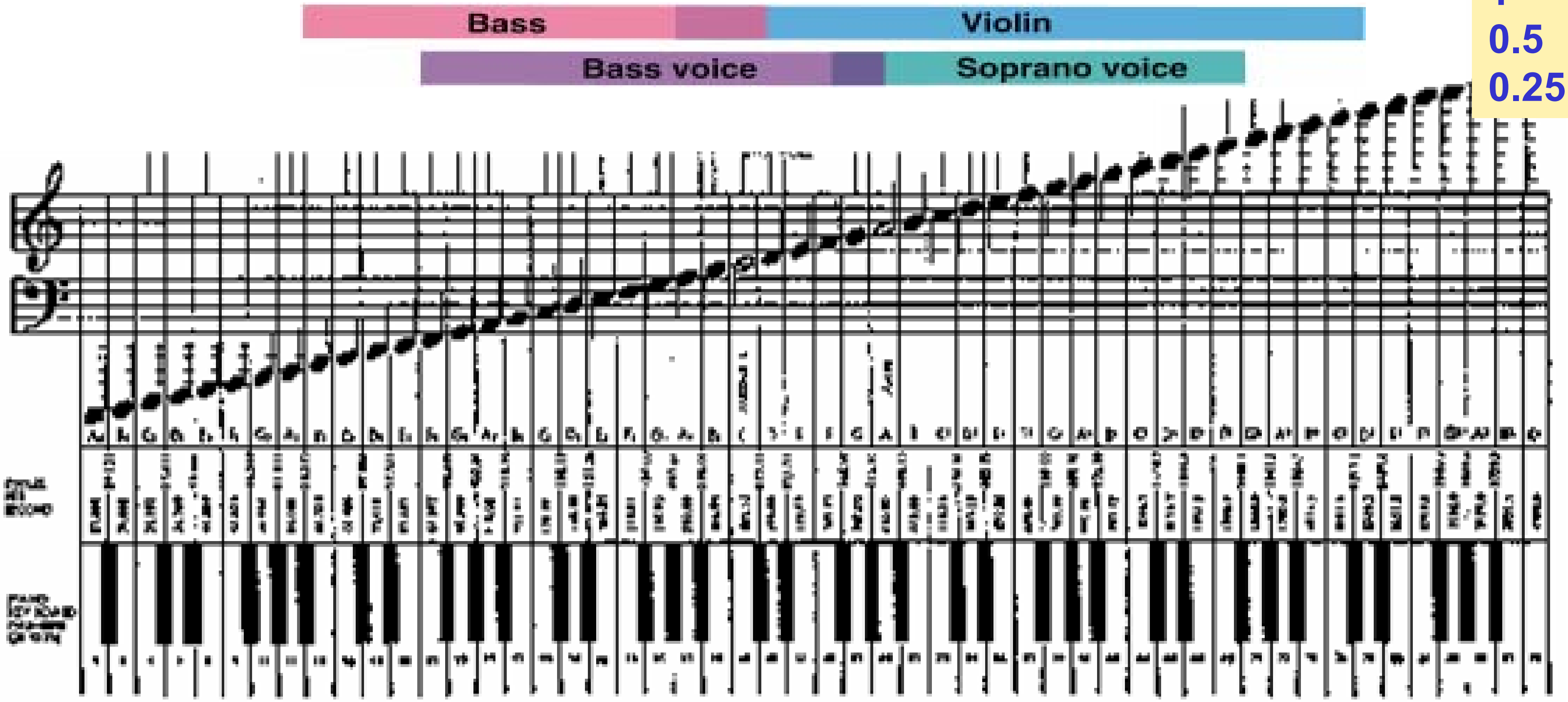




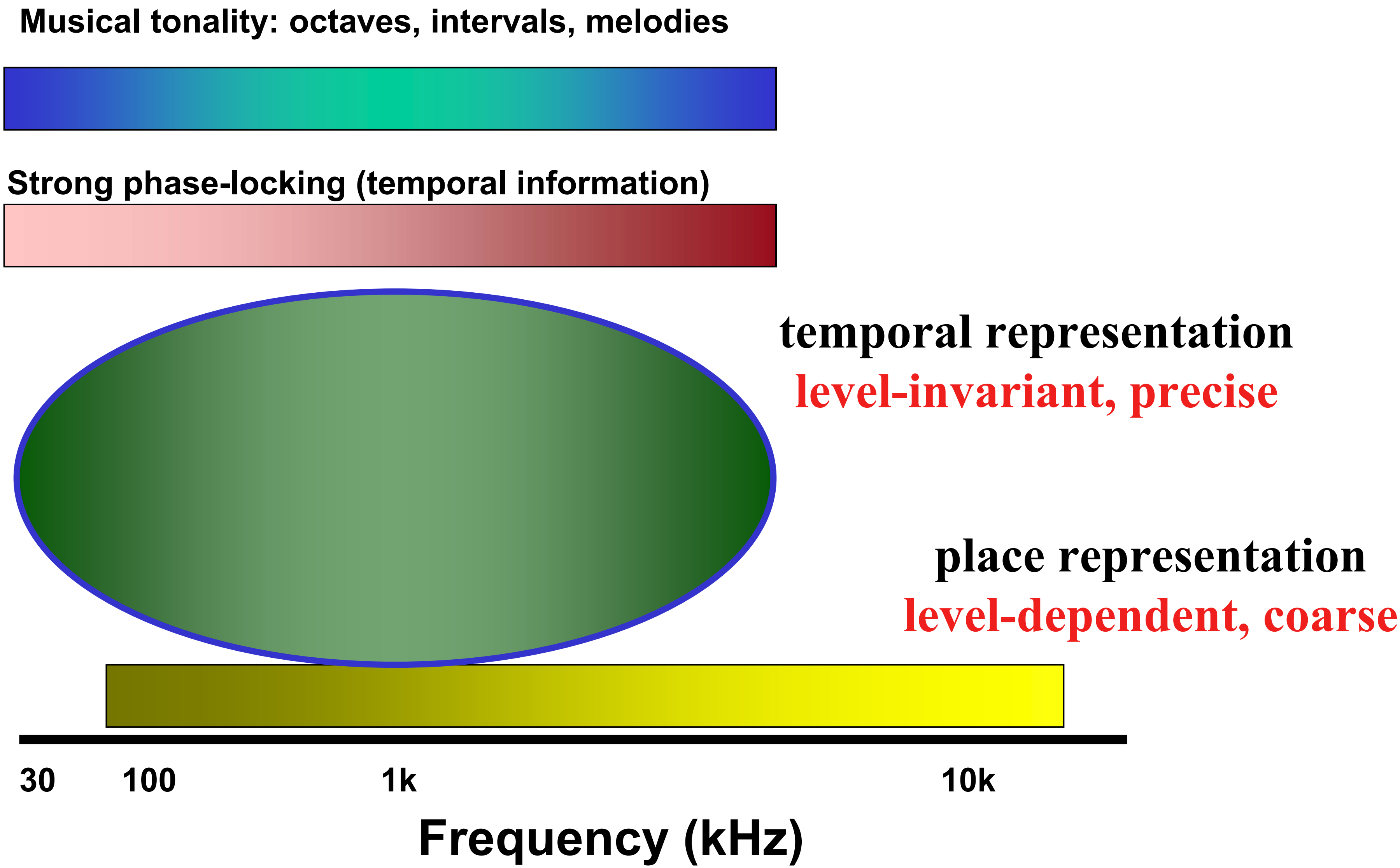
Pure tone
200 Hz



Frequency ranges of (tonal) musical instruments



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



Musical tonality: octaves, intervals, melodies

Strong phase-locking (temporal information)

temporal representation

level-invariant, precise

place representation

level-dependent, coarse

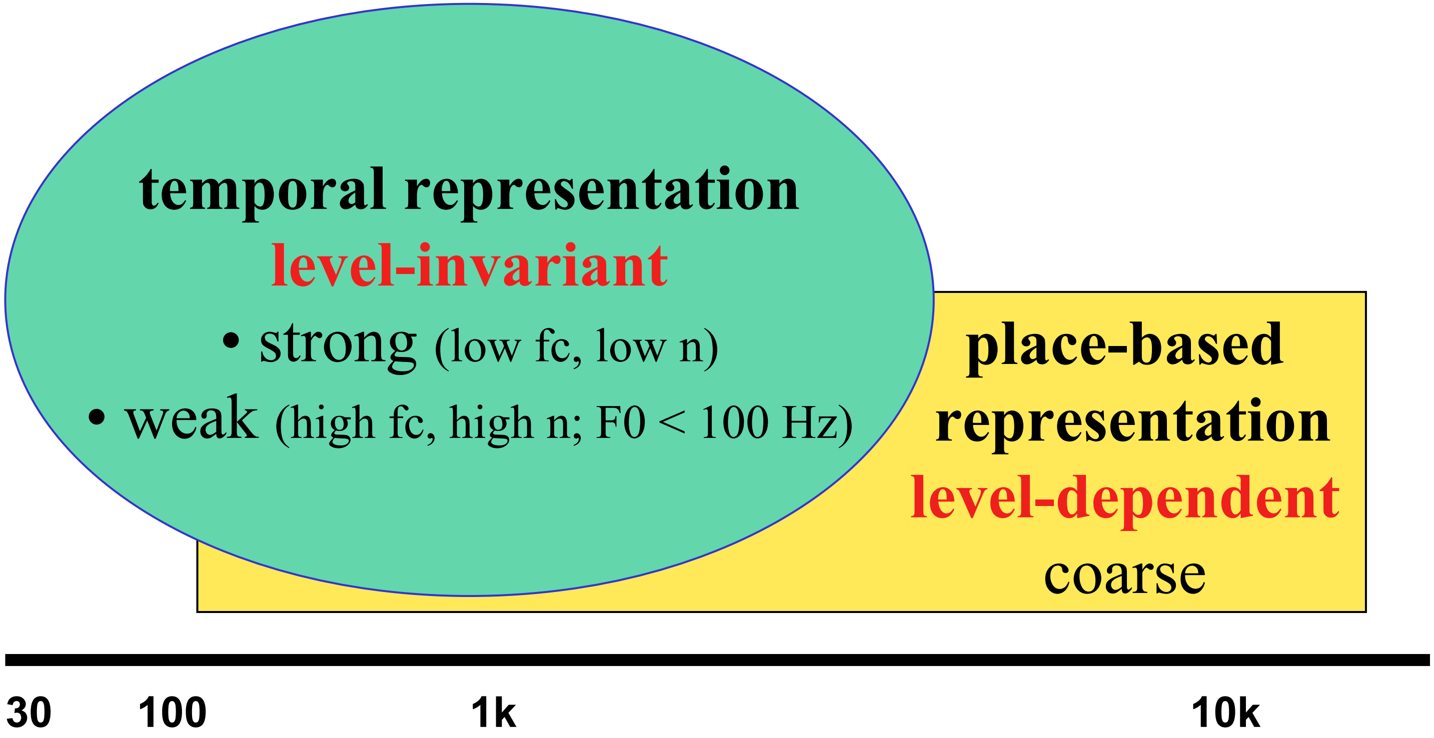
30 100 1k 10k

Frequency (kHz)

Pitch height and pitch chroma

Please see Figure 1, 2, and 7 in Roger N. Shepard. "Geometrical Approximations to the Structure of Musical Pitch." *Psychological Review* 89 no. 4 (1982): 305-322.

Duplex time-place representations



temporal representation

level-invariant

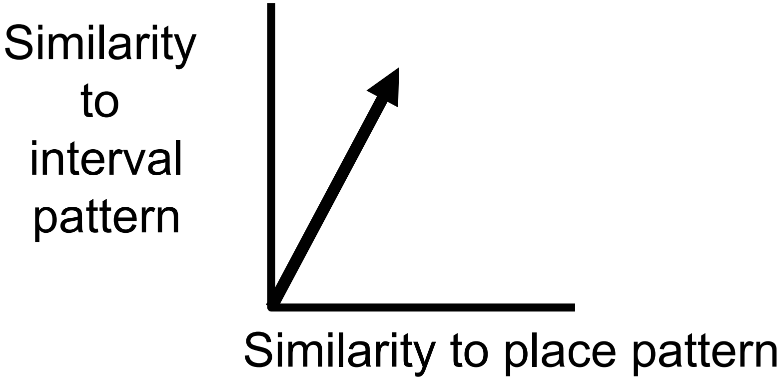
- strong (low fc, low n)
- weak (high fc, high n; F0 < 100 Hz)

place-based representation

level-dependent

coarse

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

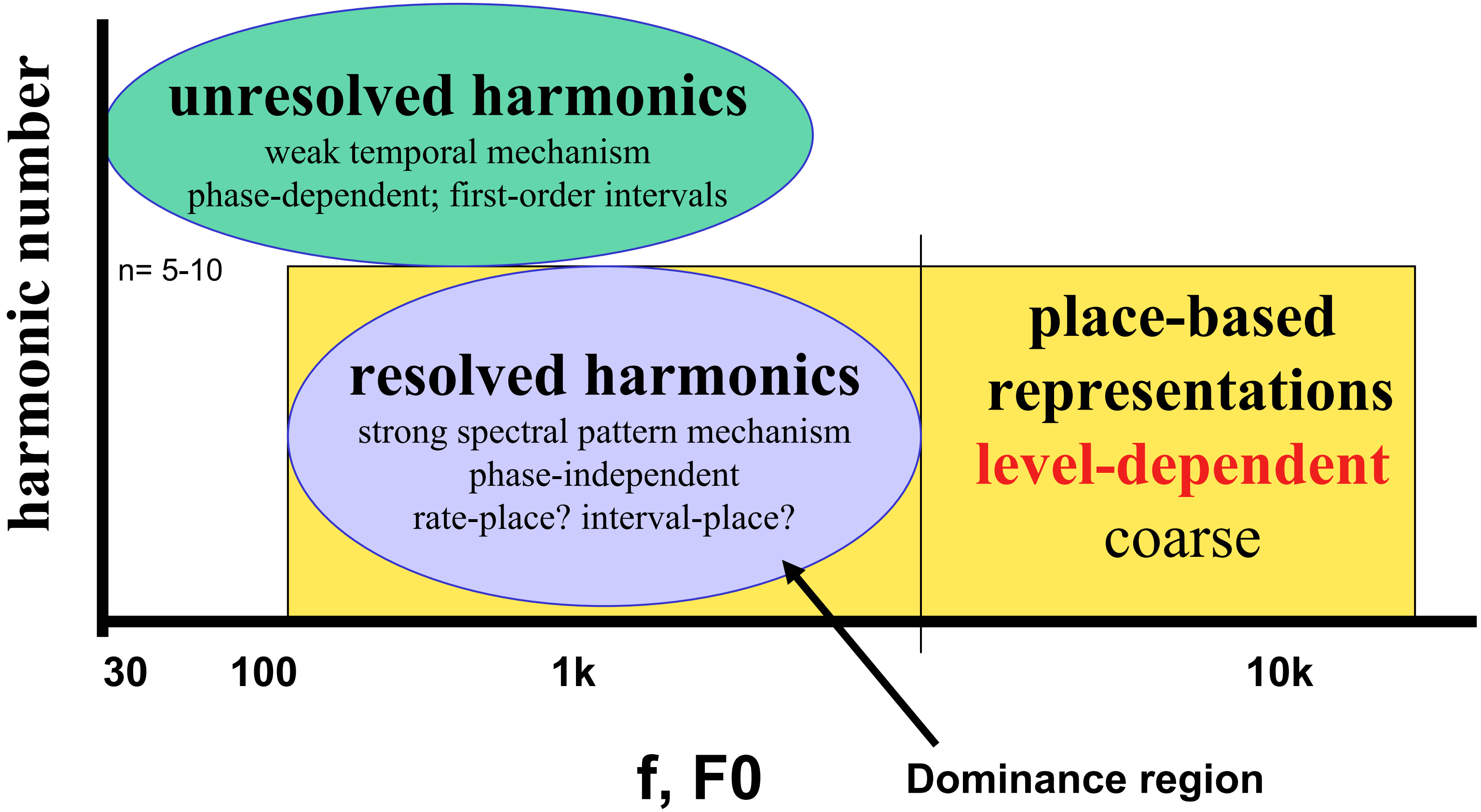


Similarity to interval pattern

Similarity to place pattern

<u>cf. Terhardt's spectral and virtual pitch</u>

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



unresolved harmonics

weak temporal mechanism
phase-dependent; first-order intervals

resolved harmonics

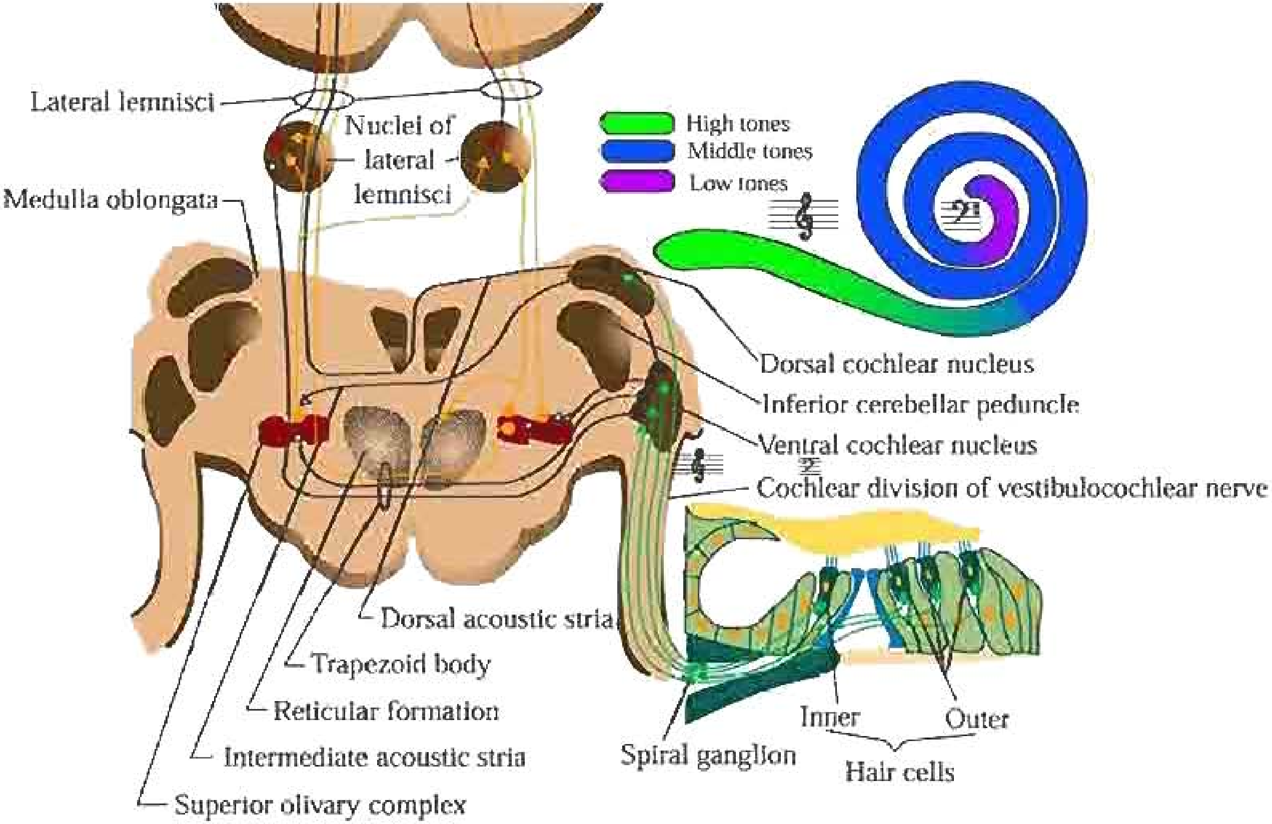
strong spectral pattern mechanism
phase-independent
rate-place? interval-place?

place-based representations

level-dependent

coarse

Dominance region

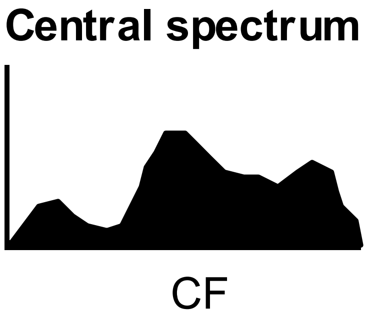

Lateral lemnisci
Nuclei of lateral lemnisci
Medulla oblongata
High tones
Middle tones
Low tones
Dorsal cochlear nucleus
Inferior cerebellar peduncle
Ventral cochlear nucleus
Cochlear division of vestibulocochlear nerve
Dorsal acoustic stria
Trapezoid body
Reticular formation
Intermediate acoustic stria
Superior olivary complex
Spiral ganglion
Inner
Outer
Hair cells

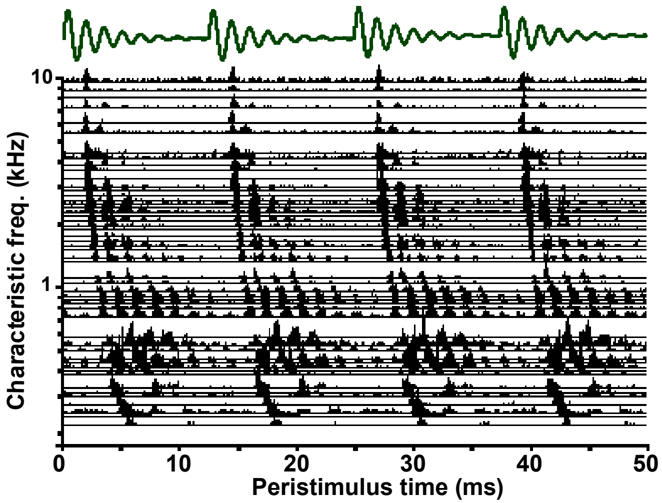
Some possible auditory representations

Local

Rate-place → Central spectrum

Masking phenomena
Loudness

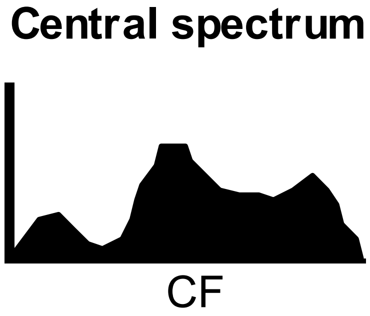




Synchrony-place

Phase-place

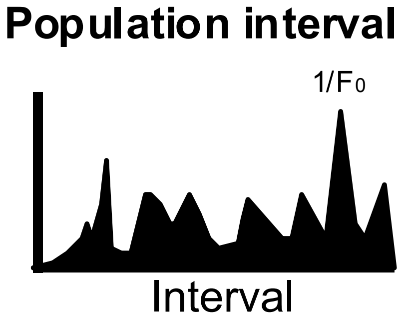
Interval-place

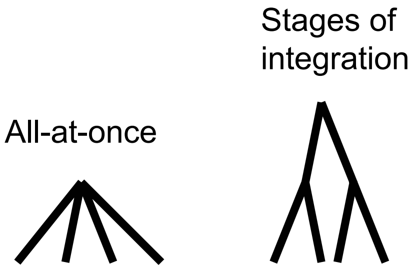


Pure tone pitch JNDs: Goldstein

Population-interval → Population interval

Complex tone pitch





Global

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

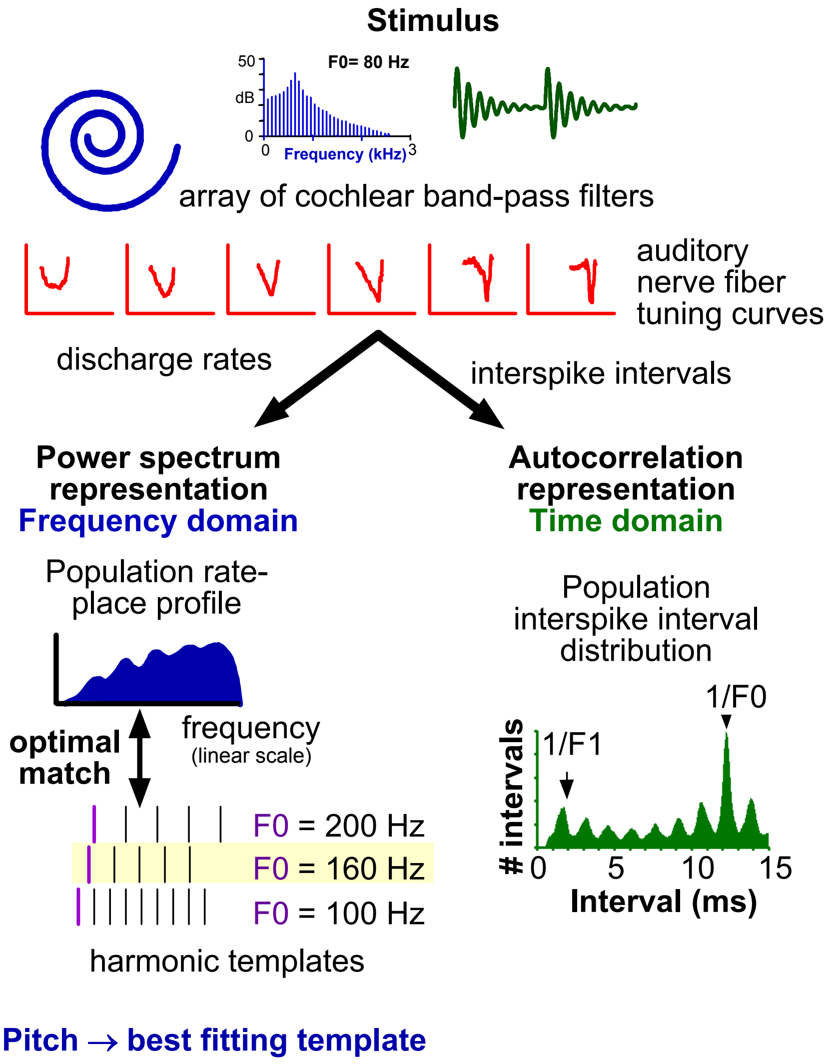
–Operate in frequency domain

–Recognize harmonic relations on resolved components

3. Temporal pattern theories

–Operate in time domain

–Analyze interspike interval dists.



Stimulus

array of cochlear band-pass filters

auditory nerve fiber tuning curves

discharge rates

interspike intervals

Power spectrum representation
Frequency domain

Population rate-place profile

harmonic templates

Pitch → best fitting template

Autocorrelation representation
Time domain

Population interspike interval distribution

Psychological perspectives on pitch

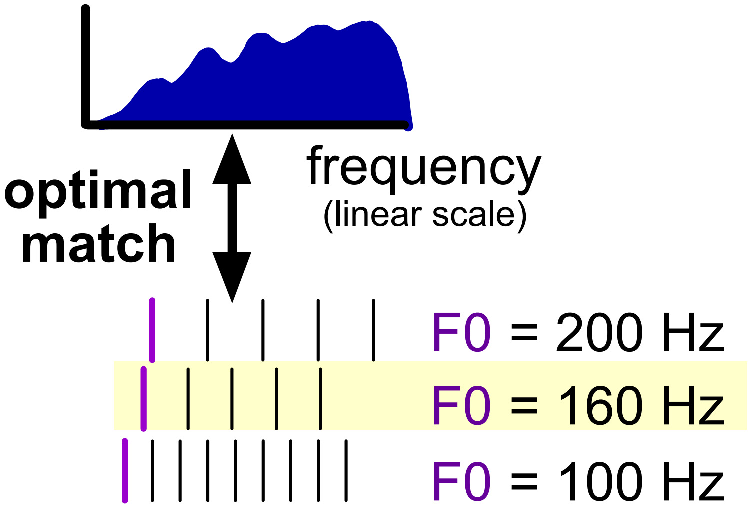
Analytical: break sounds into frequencies
    (peceptual atoms), then analyze patterns (templates)
    (British empiricism; machine perception)
Relational: extract invariant relations from patterns
    (Gestaltists, Gibsonians, temporal models)

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
    structure of experience (self-organizing systems)

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
- 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

Central spectrum



optimal match

frequency
(linear scale)

F0 = 200 Hz

F0 = 160 Hz

F0 = 100 Hz

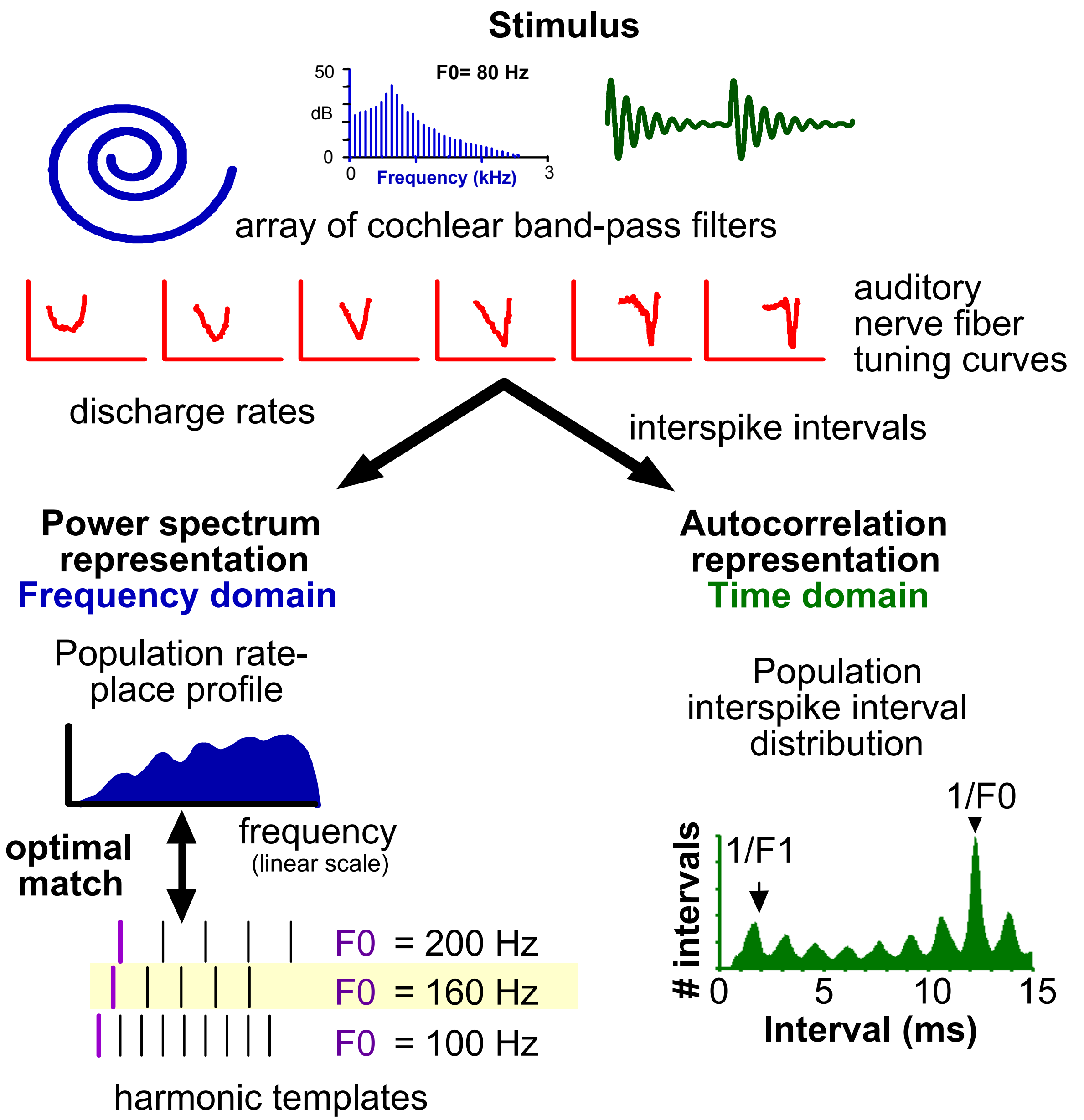
harmonic templates

Pitch → best fitting template

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 placeinformation as the basis of the central representation, and
2) use of spectral vs. autocorrelation-like central representations



Stimulus

array of cochlear band-pass filters

auditory nerve fiber tuning curves

discharge rates

interspike intervals

Power spectrum representation
Frequency domain

Population rate-place profile

F0 = 200 Hz
F0 = 160 Hz
F0 = 100 Hz

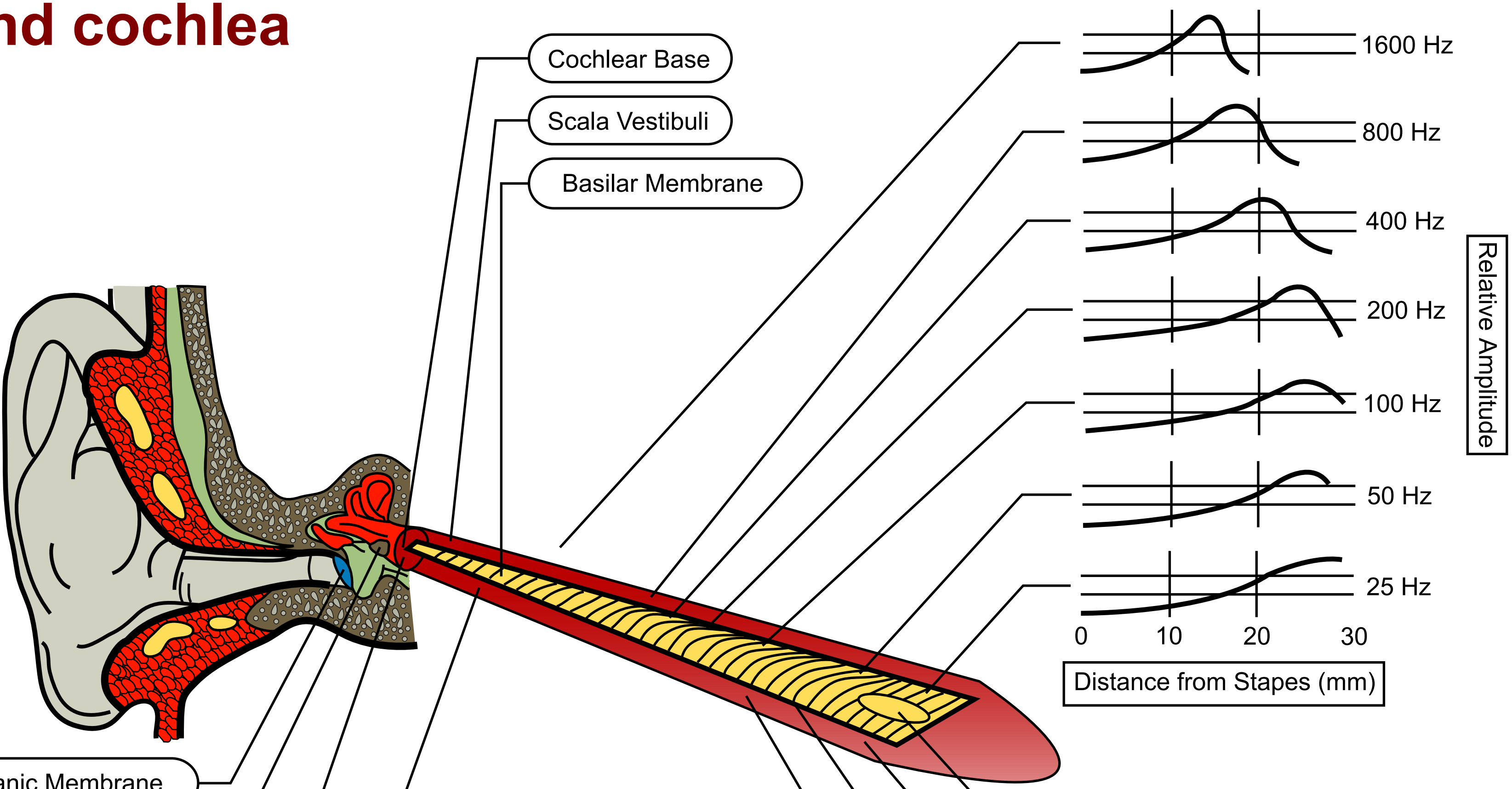
harmonic templates

Autocorrelation representation
Time domain

Population interspike interval distribution

Pitch → best fitting template

Ear and cochlea



Cochlear Base

Scala Vestibuli

Basilar Membrane

Tympanic Membrane

Stapes on Oval Window

Narrow Base of Basilar Membrane is “tuned” for high frequencies

Scala Tympani

Helicotrema

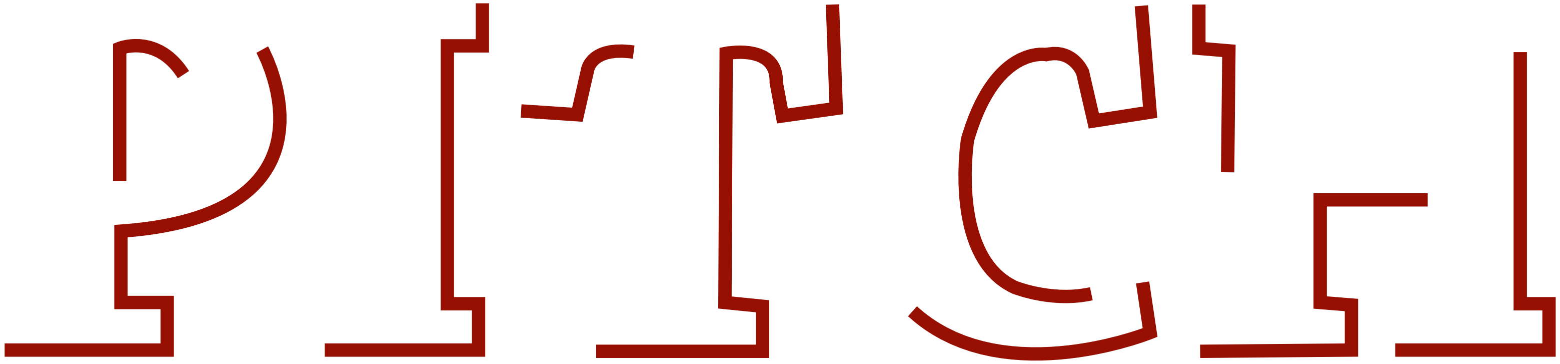
Cochlear Apex

Wider apex is “tuned” for low frequencies

“Unrolled” 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)

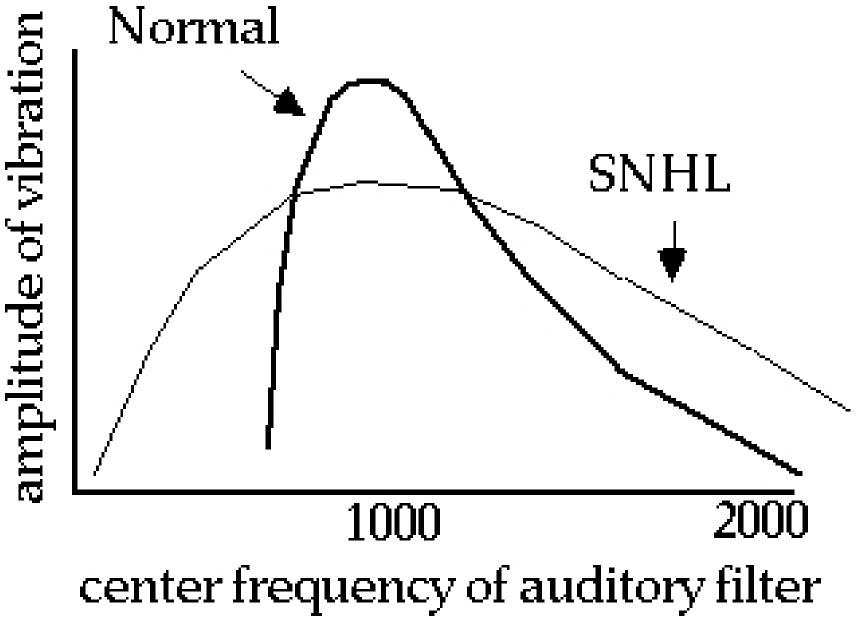
“Virtual” pitch: F0-pitch as pattern completion



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

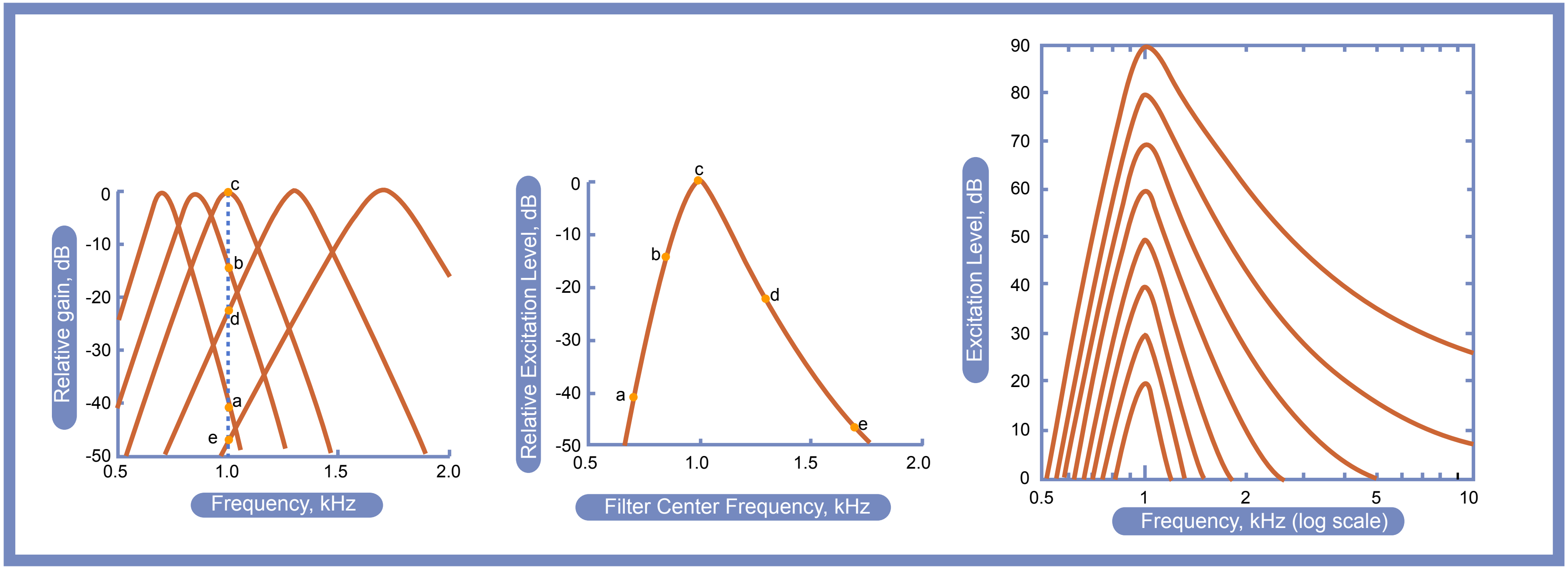
Power spectrum
Filter metaphor

Notion of one central spectrum that subserves perception of pitch, timbre, and loudness

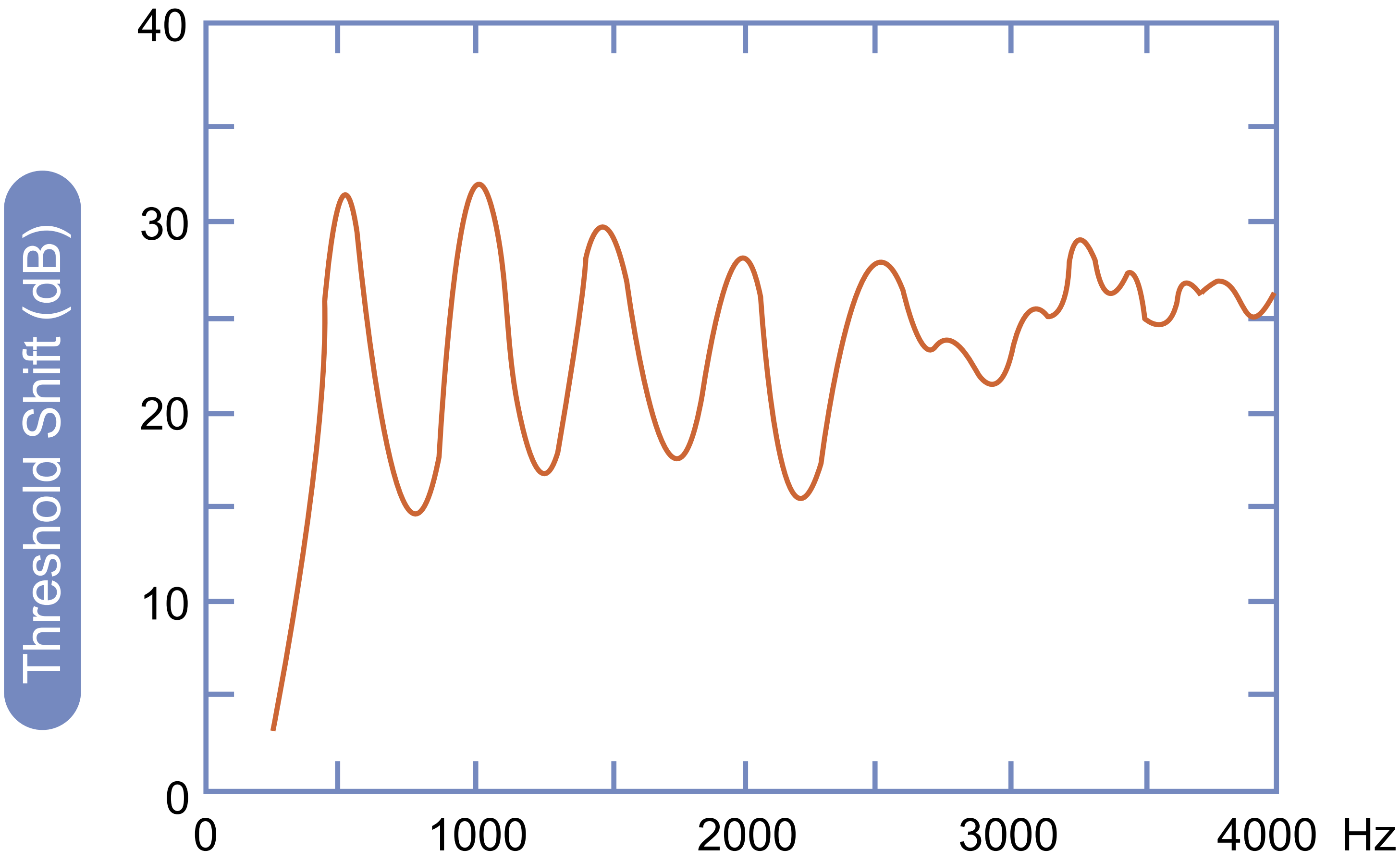


2.2. Excitation pattern Using the filter shapes and bandwidths derived from masking experiments we can produce the excitation pattern produced by a sound. The excitation pattern shows how much energy comes through each filter in a bank of auditory filters. It is analogous to the pattern of vibration on the basilar membrane. For a 1000 Hz pure tone the excitation pattern for a normal and for a SNHL 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

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



Resolution of harmonics



Goldstein's harmonic templates

Please see Figure 3 in Goldstein J. L., A. Gerson, P. Srulovicz, M. Furst. "Verification of the Optimal Probabilistic Basis of Aural Processing in Pitch of Complex Tones." *J AcoustSoc Am*. 63, no. 2 (Feb, 1978): 486-97.

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
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.

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

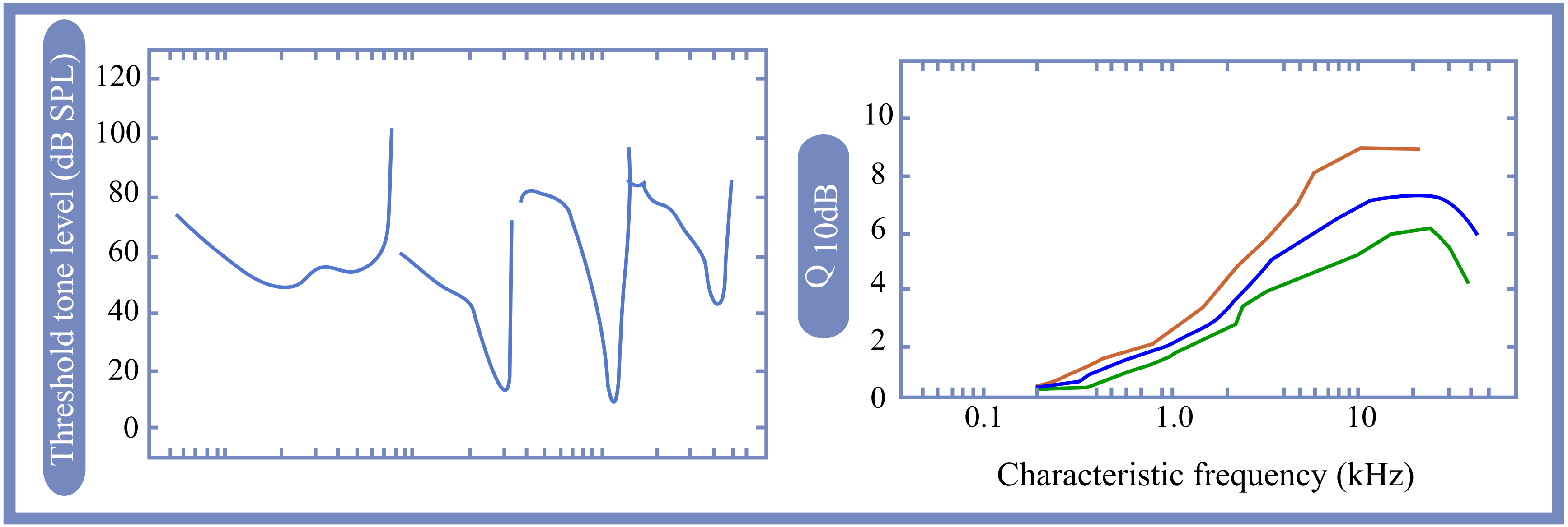
SPINET:

Cohen Grossberg, Wyse JASA

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

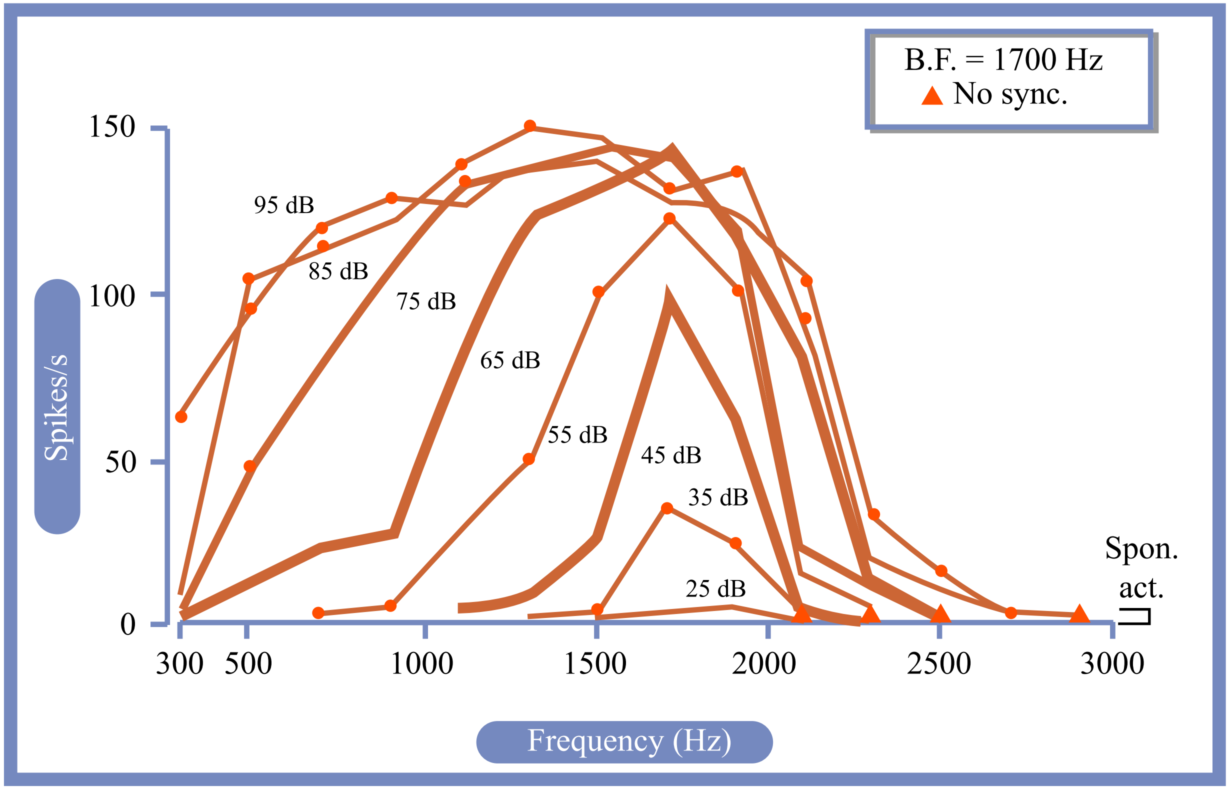
Please see Cohen M. A., S. Grossberg, L. L. Wyse. "A Spectral Network Model of Pitch Perception." *J AcoustSoc Am*. 98 no. 2 Pt 1 (Aug, 1995): 862-79.

Neural tuning as a function of CF



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



Spectral pattern theories - pros & cons

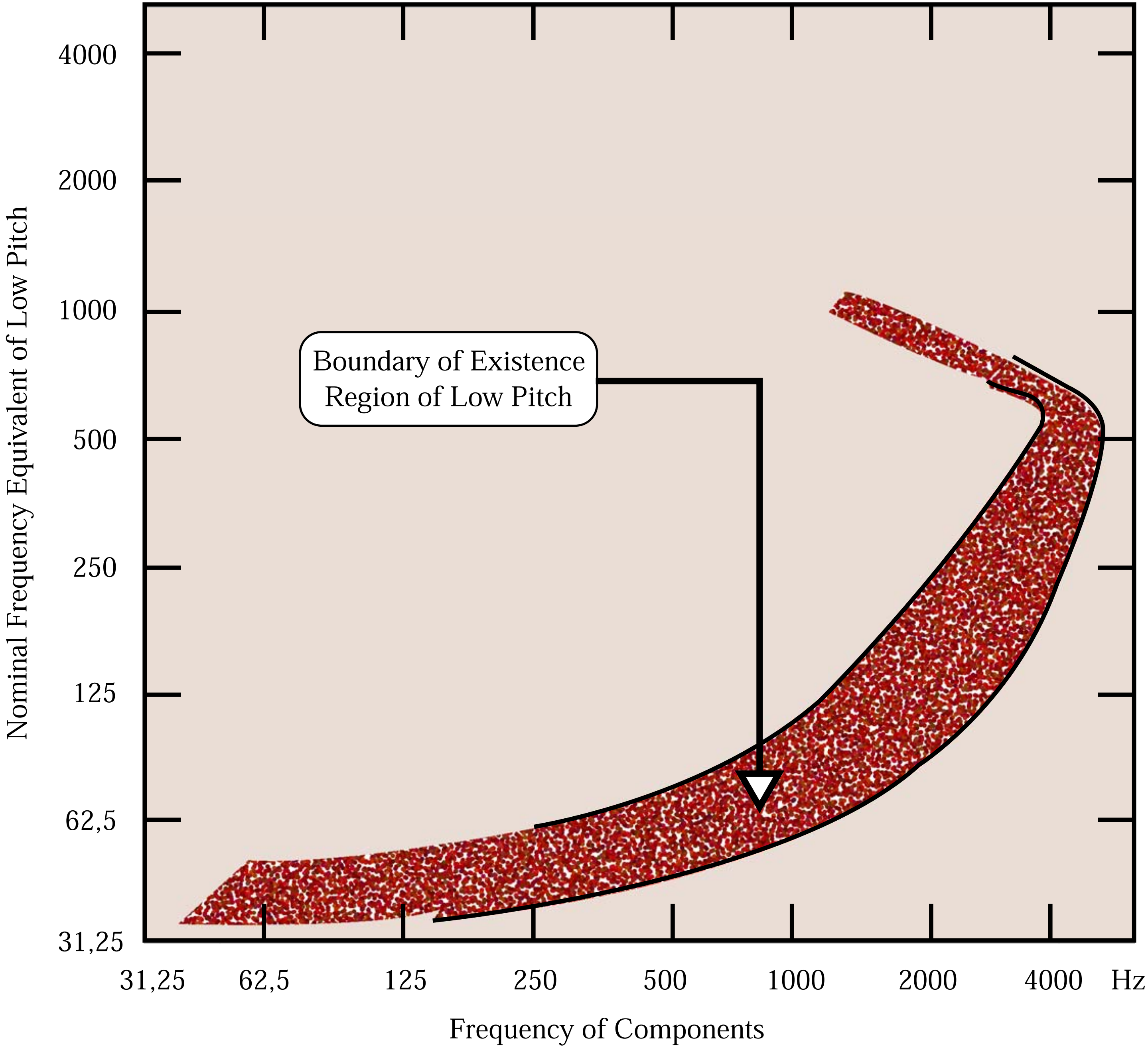
Do make use of frequency tuning properties of elements in the auditory system
No clear neural evidence of narrow (< 1/3 octave) frequency channels in low-BF regions (< 2 kHz)

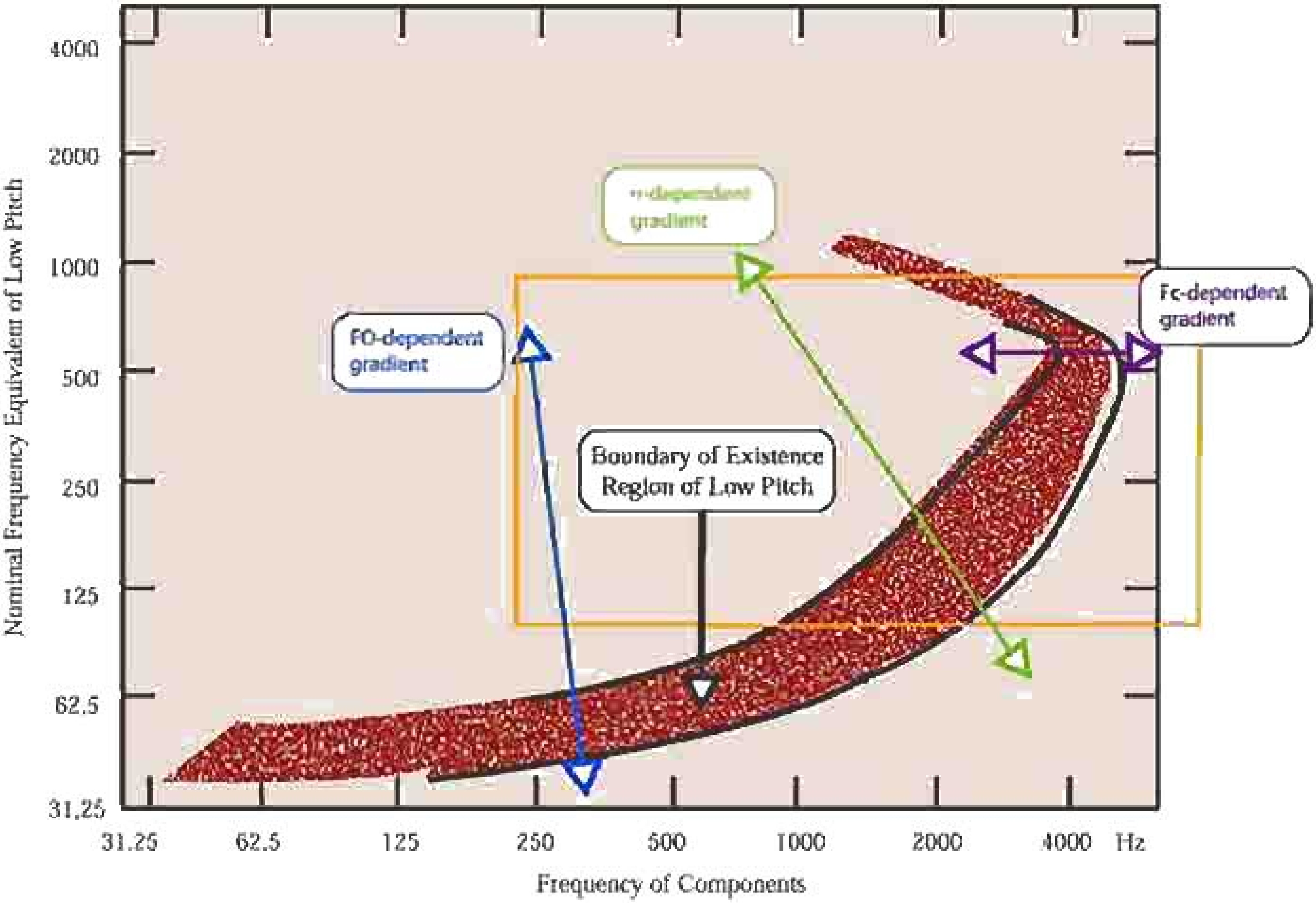
Operate on perceptually-resolved harmonics
Do not explain low pitches of unresolved harmonics

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

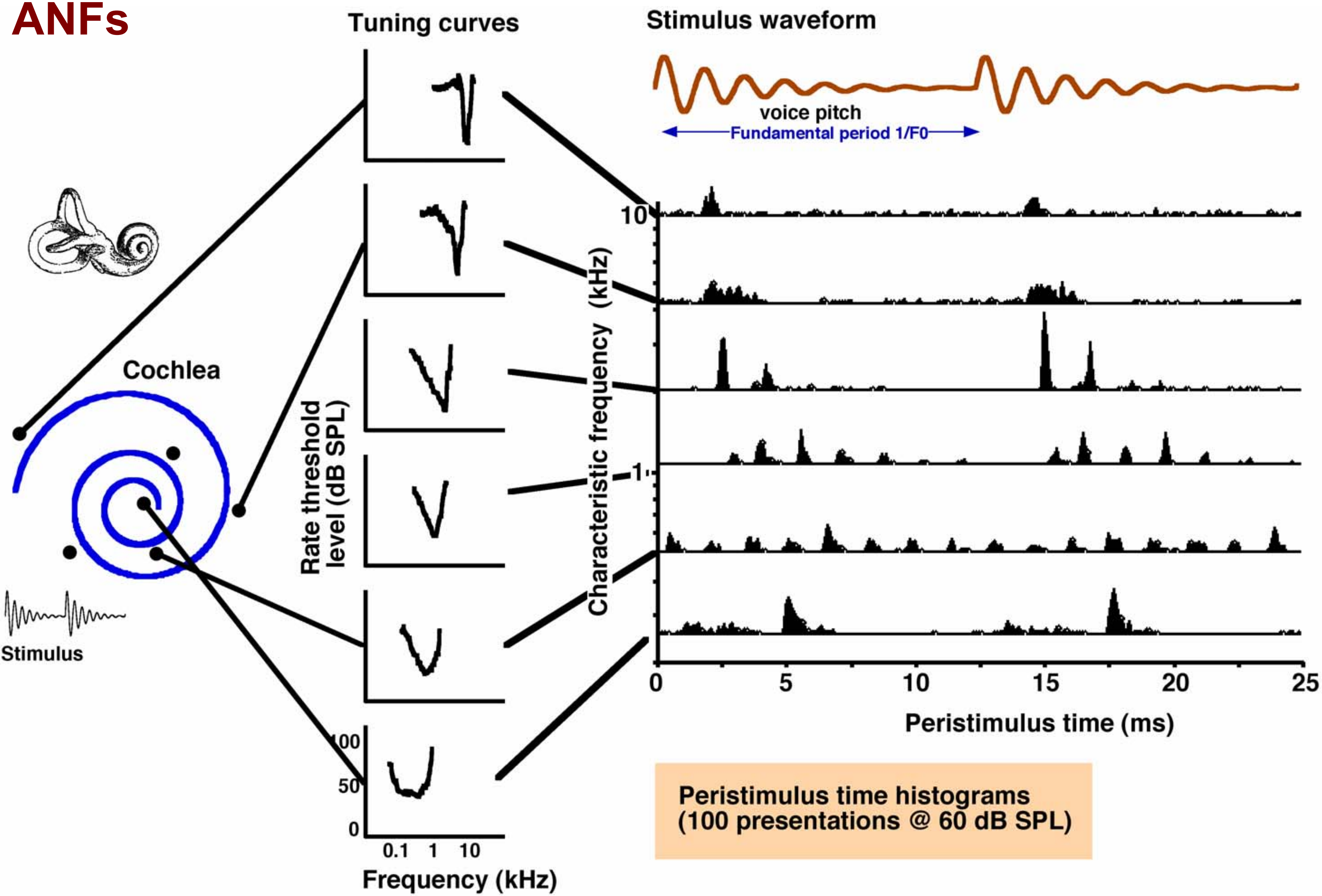
Do not explain well existence region for F0
Learning theories don't account for F0 ranges
or for phylogenetic ubiquity of periodicity pitch

Existence region for missing fundamentals of AM tones




4000
2000
1000
500
250
125
62.5
31.25
Nominal Frequency Equivalent of Low Pitch
n-dependent gradient
F0-dependent gradient
Fc-dependent gradient
Boundary of Existence Region of Low Pitch
31.25
62.5
125
250
500
1000
2000
4000
Hz
Frequency of Components

ANFs



Peristimulus time histograms (100 presentations @ 60 dB SPL)

Temporal pattern theories

Σ First-order intervals (renewal density)

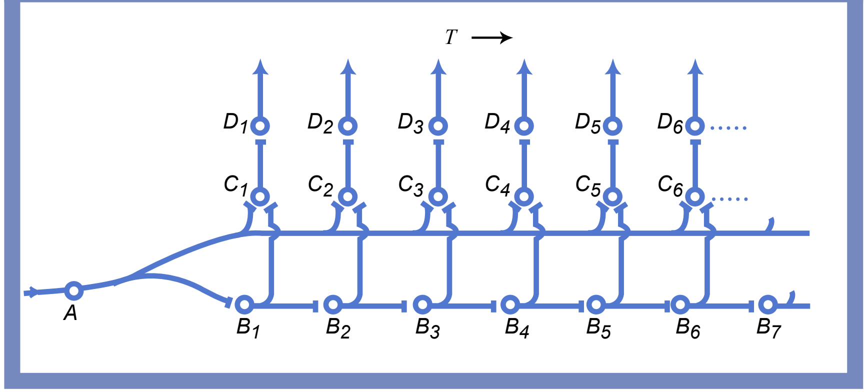
- **Schouten's temporal theory (1940's) depended on interactions between unresolved (high) harmonics. It was displaced by discovery of dominance region and binaural combination pitches in the 1960's. The idea persists, however in the form of spectral mechanisms for resolved harmonics and temporal ones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

Please see van Noorden, L. "Two channel pitch perception." In *Music, Mind and Brain*. Edited by M. Clynes. New York: Plenum.

Σ All-order intervals (temporal autocorrelation)

Licklider (1951)

Basic schema of neuronal autocorrelator. A is the input neuron, *B*$_1$, *B*$_2$, *B*$_3$, ... is a delay chain.



Please see Moore, BCJ. *An Introduction to the Psychology of Hearing*. 5th ed. San Diego: Academic Press. 2003.

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.

Interval-based theories of pitch

First-order intervals

(renewal density)

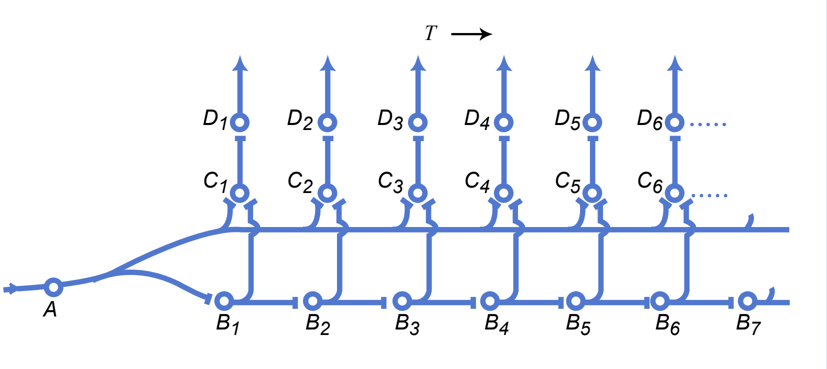
Please see van Noorden, L. Two channel pitch perception. In Music, Mind and Brain. Edited by M. Clynes. New York: Plenum.

Please see Moore, BCJ. An introduction to the psychology of hearing. 5th ed. San Diego: Academic Press. 2003.

All-order intervals

(temporal autocorrelation)

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

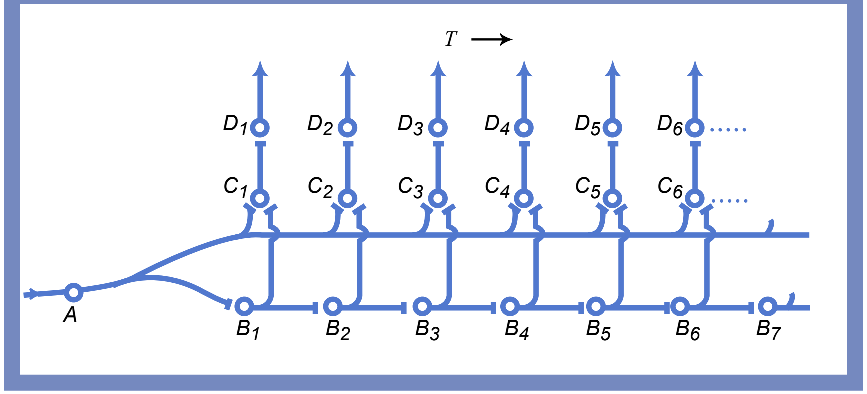


Licklider (1951)

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.

Licklider’s (1951) duplex model of pitch perception

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



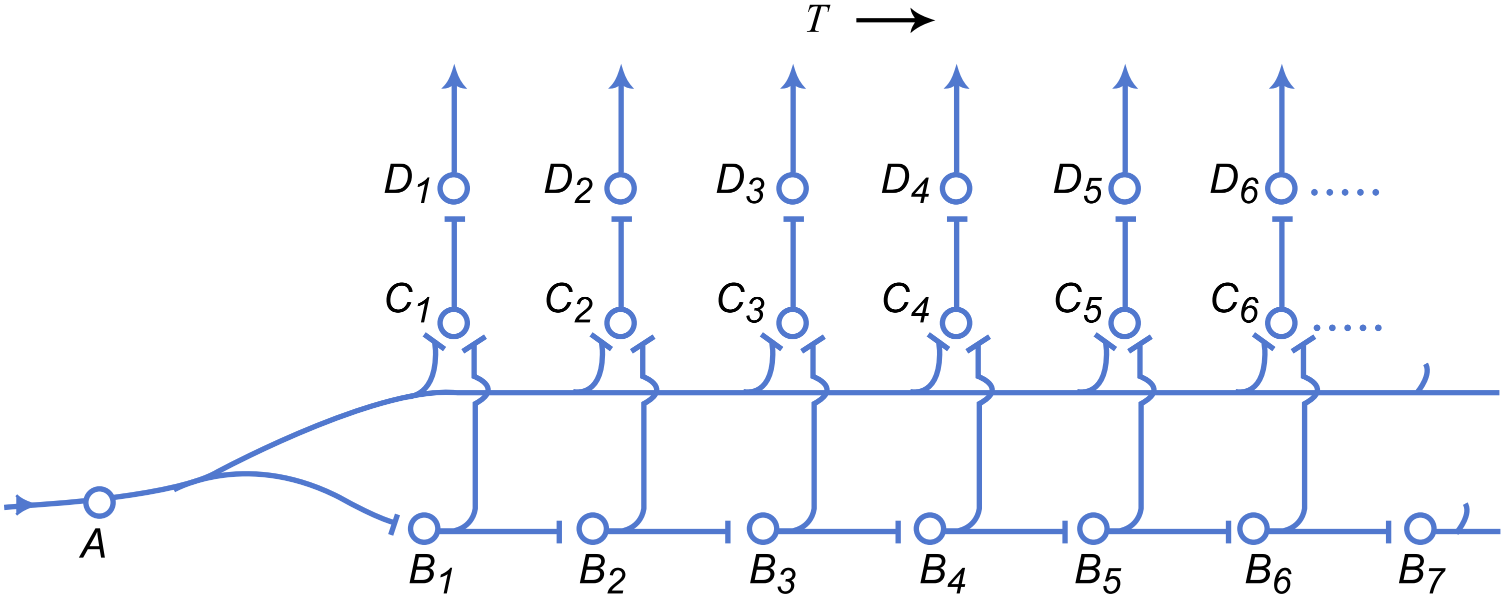
(Image removed due to copyright considerations.)

Licklider’s binaural triplex model

J.C.R. Licklider “Three AuditoryTheories” In *Psychology: A Study of a Science*. Vol. 1, S. Koch, ed., McGraw-Hill, 1959, pp. 41-144.

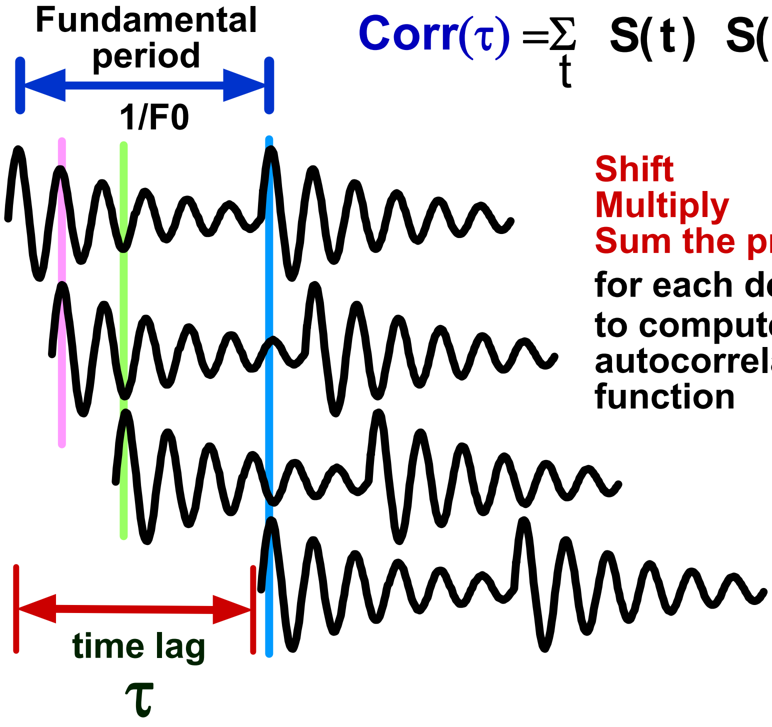
Delay lines

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



Autocorrelation and interspike intervals

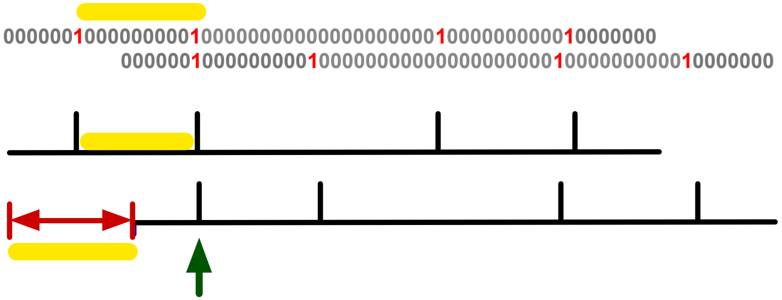
Autocorrelation functions



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

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

Autocorrelations of spike trains = Histograms of all-order intervals



Please see Figure 6.16A-D in Lyon, Richard, and Shihab Shamma. "Auditory Representations of Timbre and Pitch." In *Auditory Computation*. Edited by R. R. Fay. New York: Springer Verlag. 1996.

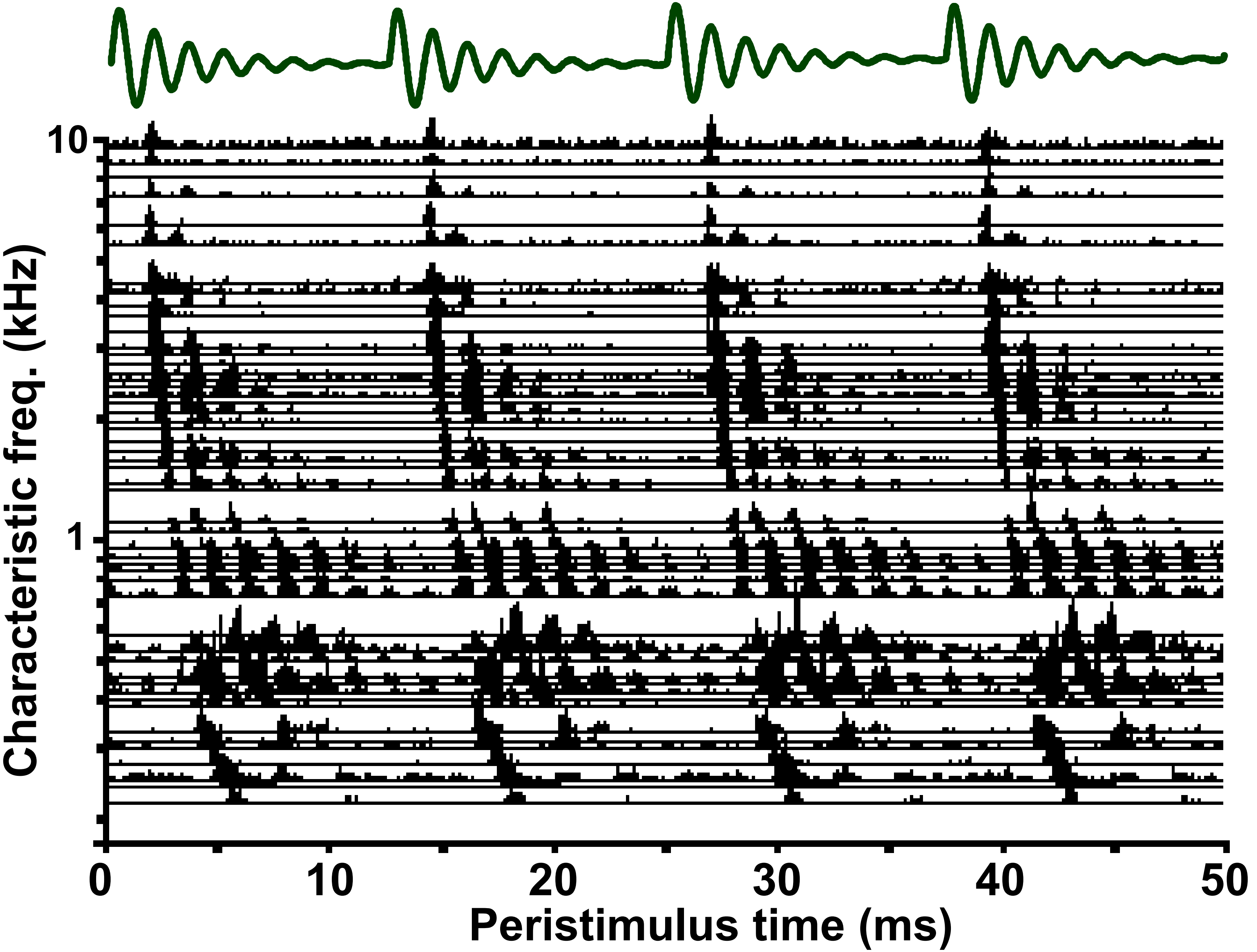
Correlograms: interval-place displays (Slaney & Lyon)

Please see Slaney, Malcolm, and Richard F. Lyon. "On the Importance of Time -A Temporal Representation of Sound." In *Visual Representations of Speech Signals*. Edited by M. Crawford. New York: John Wiley. 1993.

Correlograms

Please see Slaney, Malcolm, and Richard F. Lyon. "On the Importance of Time -A Temporal Representation of Sound." In *Visual Representations of Speech Signals*. Edited by M. Crawford. New York: John Wiley. 1993.

Auditory nerve



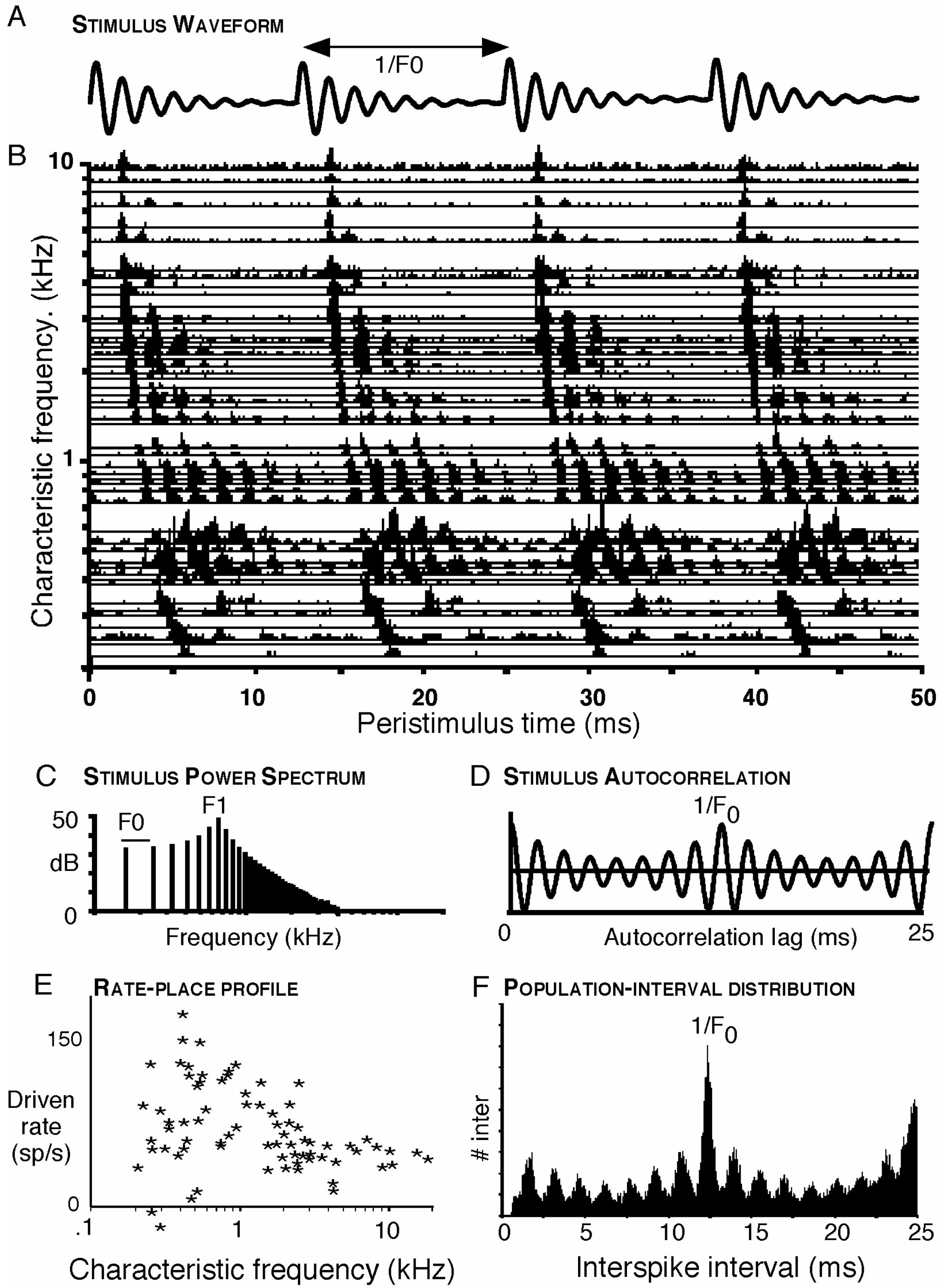
Temporal coding in the auditory nerve

Work with Bertrand Delgutte
Cariani & Delgutte (1996ab)

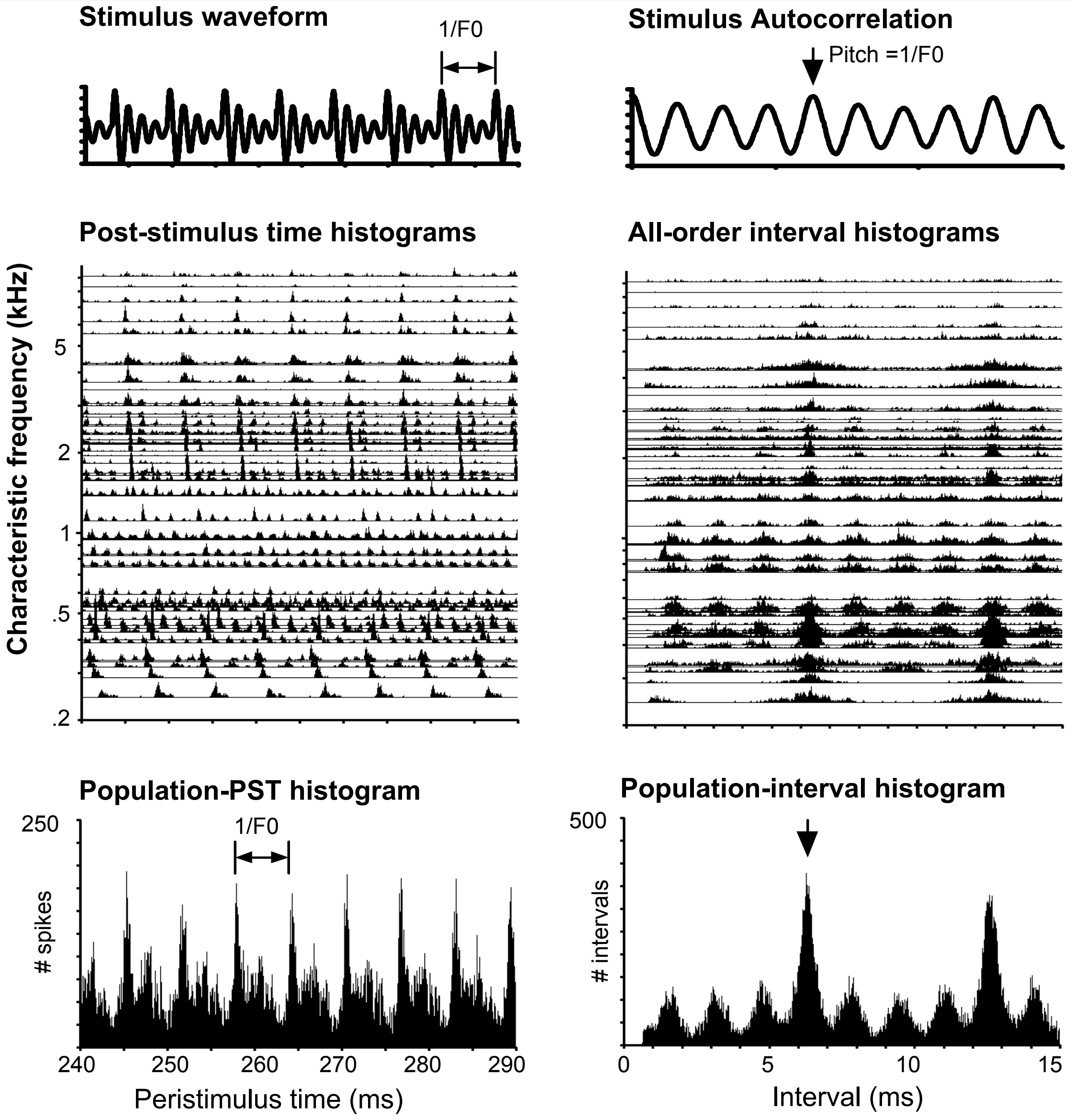
Dial-anesthetized cats.
100 presentations/fiber
60 dB SPL

Population-interval distributions are compiled by summing together intervals from all auditory nerve fibers.

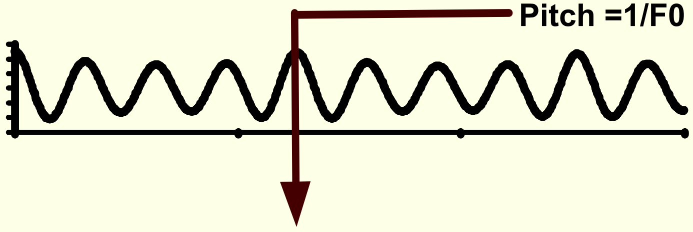
The most common intervals present in the auditory nerve are invariably related to the pitches heard at the fundamentals of harmonic complexes.



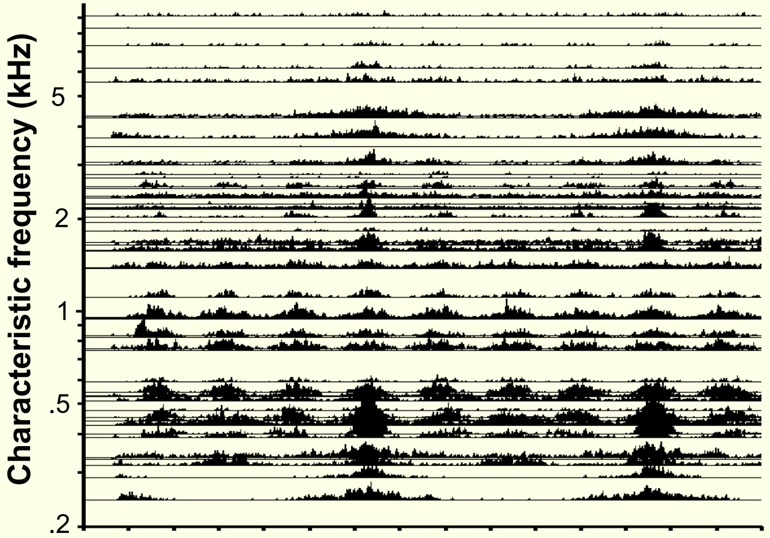
The population-interval distribution of the auditory nerve



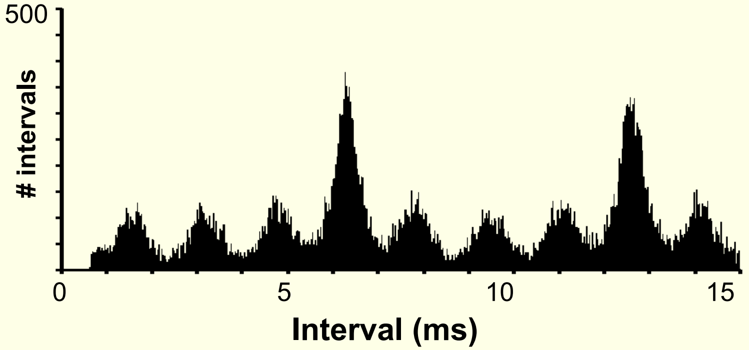
Stimulus Autocorrelation



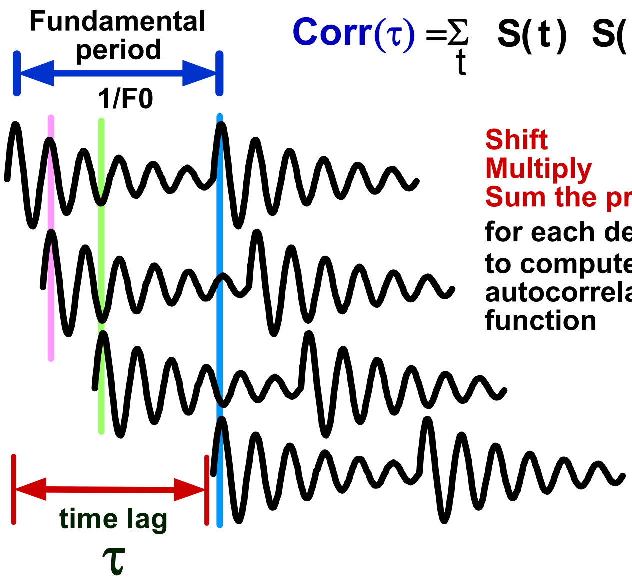
All-order interval histograms



Population-interval histogram



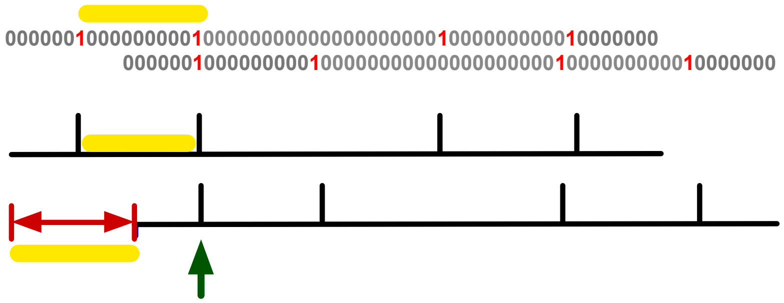
Autocorrelation functions



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

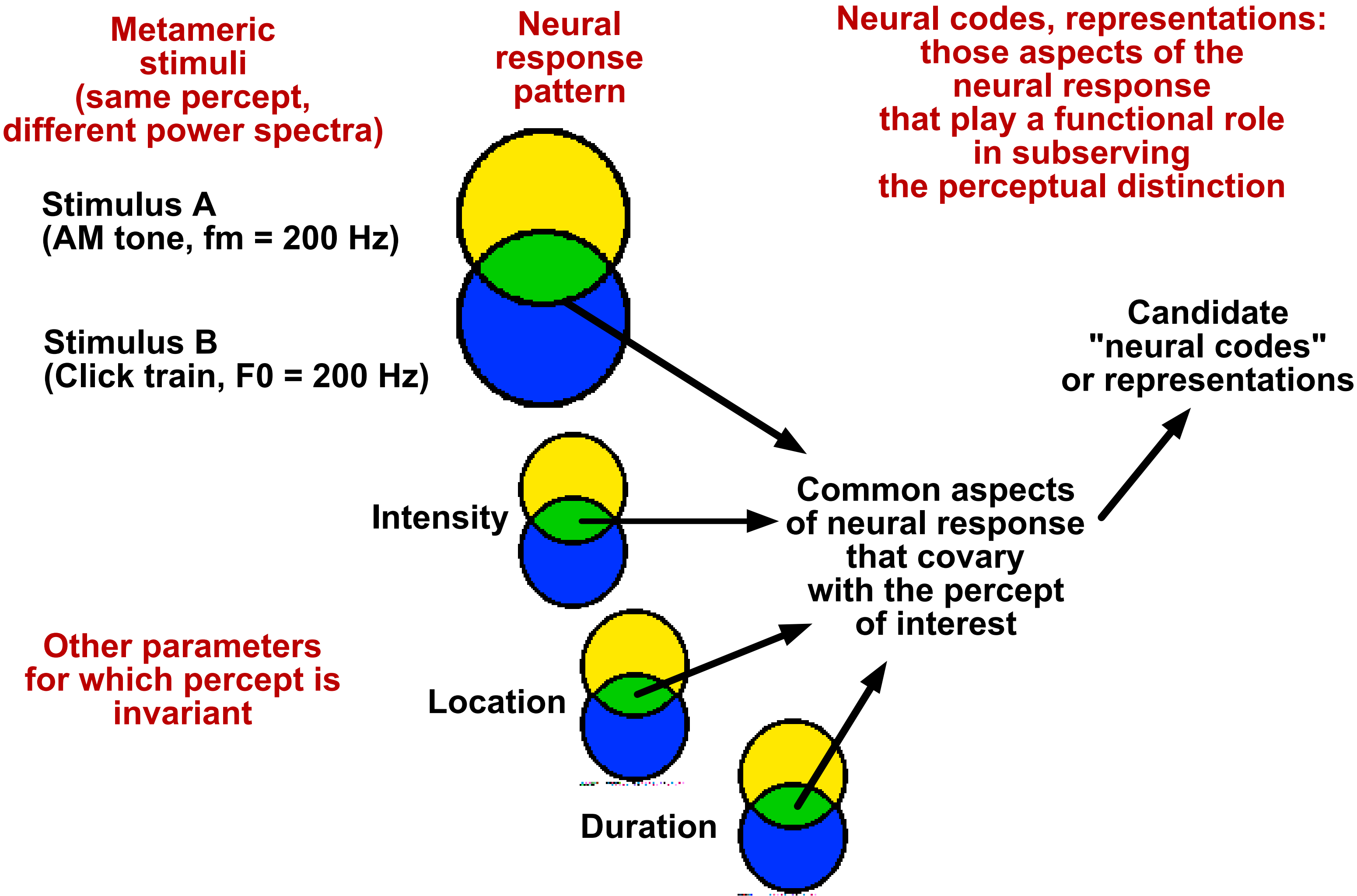
Shift
Multiply
Sum the products
for each delay τ
to compute
autocorrelation
function

Autocorrelations of spike trains = Histograms of all-order intervals



Percept-driven search for neural codes:

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



Metameric stimuli (same percept, different power spectra)

Neural response pattern

Neural codes, representations: those aspects of the neural response that play a functional role in subserving the perceptual distinction

Stimulus A
(AM tone, fm = 200 Hz)

Stimulus B
(Click train, F0 = 200 Hz)

Candidate "neural codes" or representations

Intensity

Common aspects of neural response that covary with the percept of interest

Other parameters for which percept is invariant

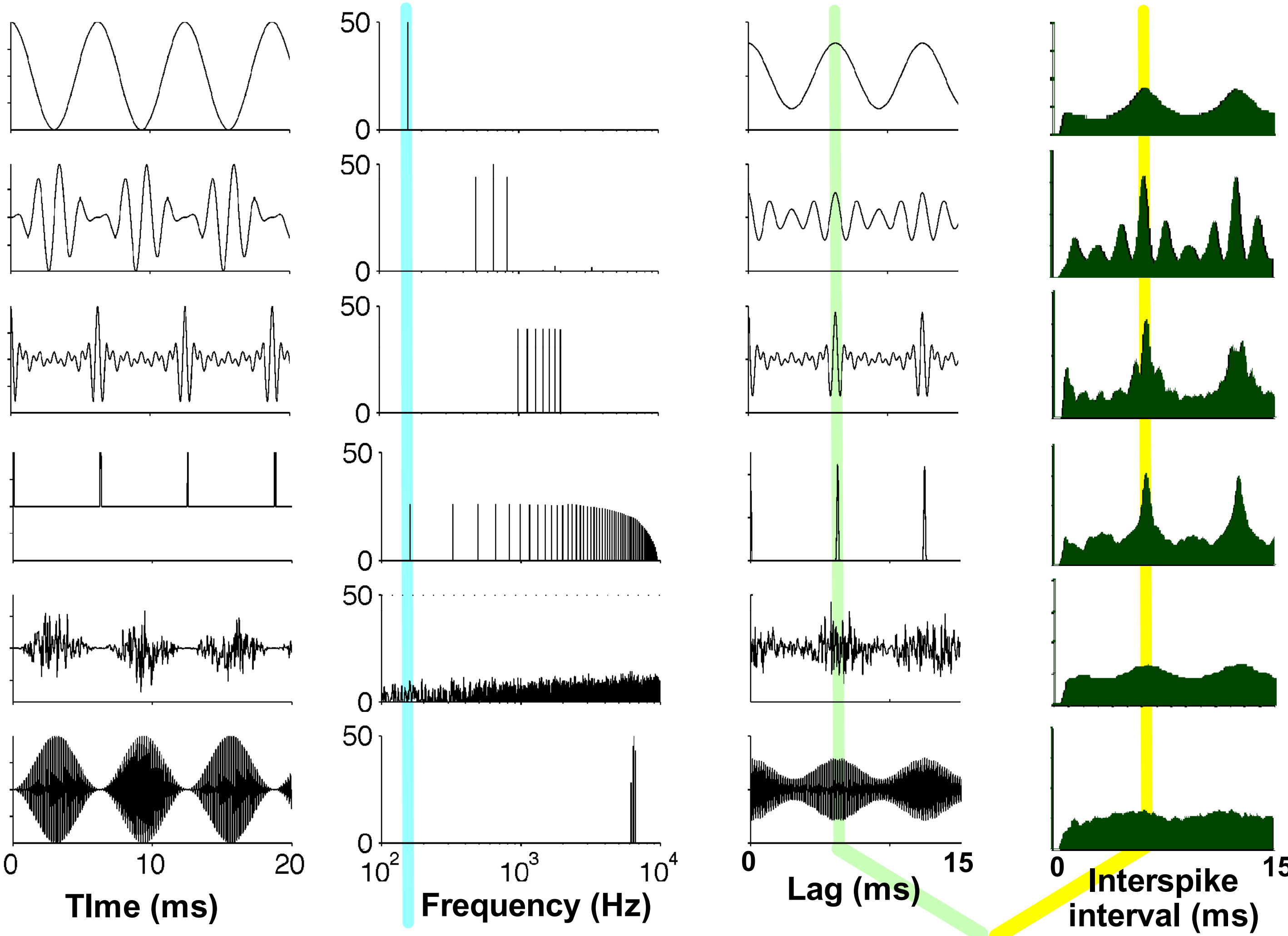
Location

Duration

Pitch equivalence

Six stimuli that produce a low pitch at 160 Hz

Waveform **Power spectrum** **Autocorrelation** **Population interval distribution**



Pure tone
160 Hz

AM tone
F_m:160 Hz
F_c:640 Hz

Harms 6-12
F_0 : 160 Hz

Click train
F_0 :160 Hz

AM noise
F_m :160 Hz

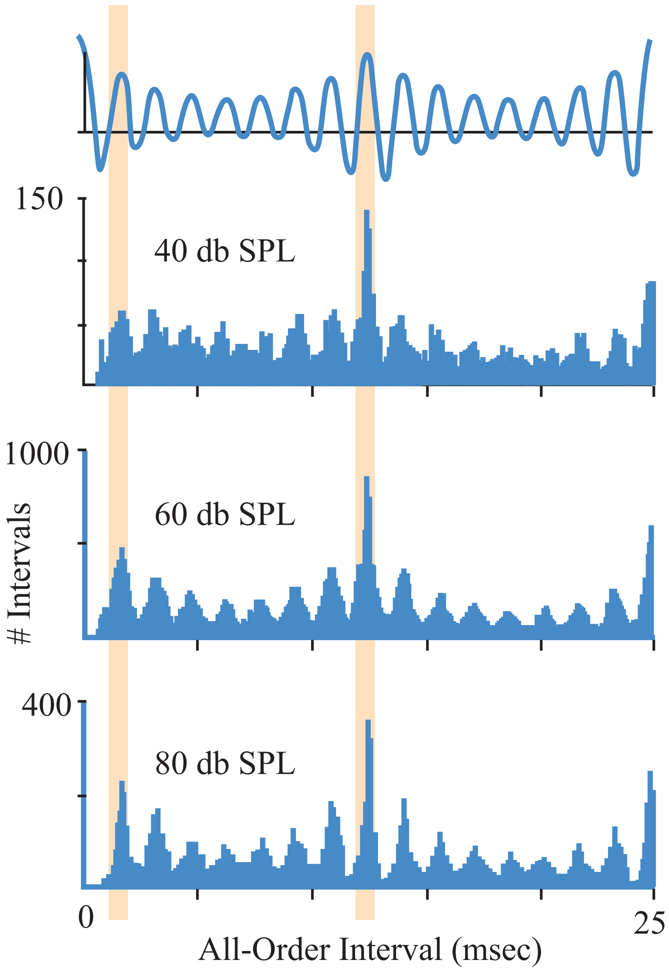
AM tone
F_m :160 Hz
F_c:6400 Hz

STRONG PITCHES

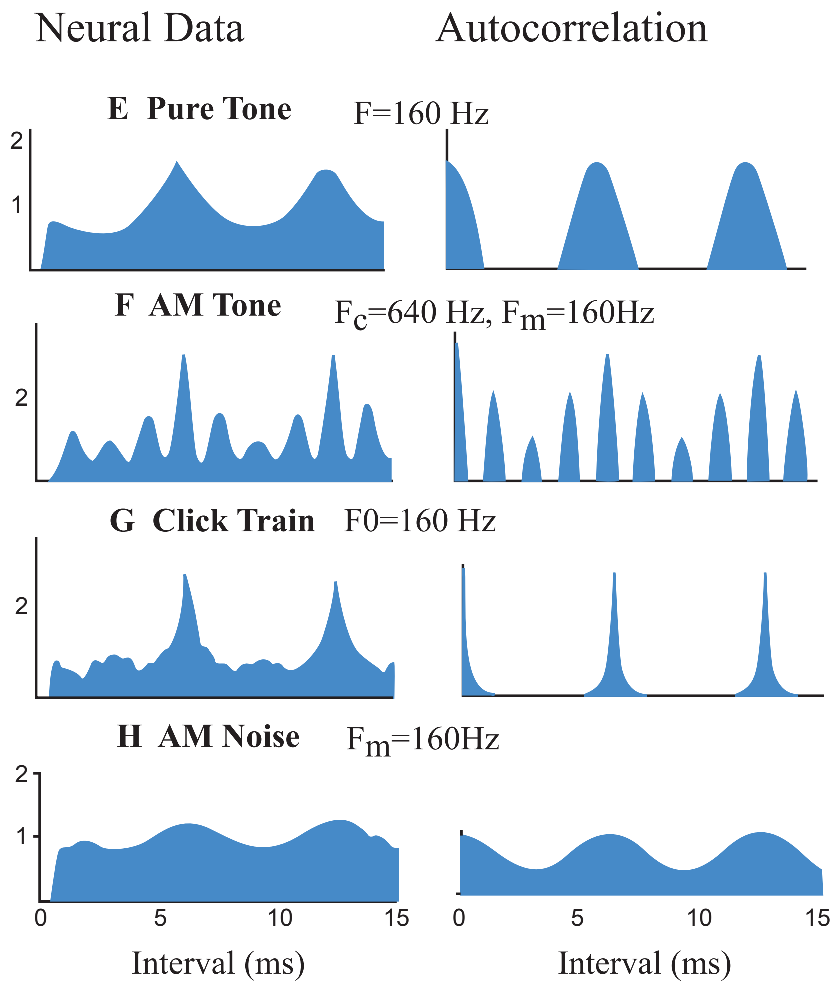
WEAK PITCHES

Pitch frequency

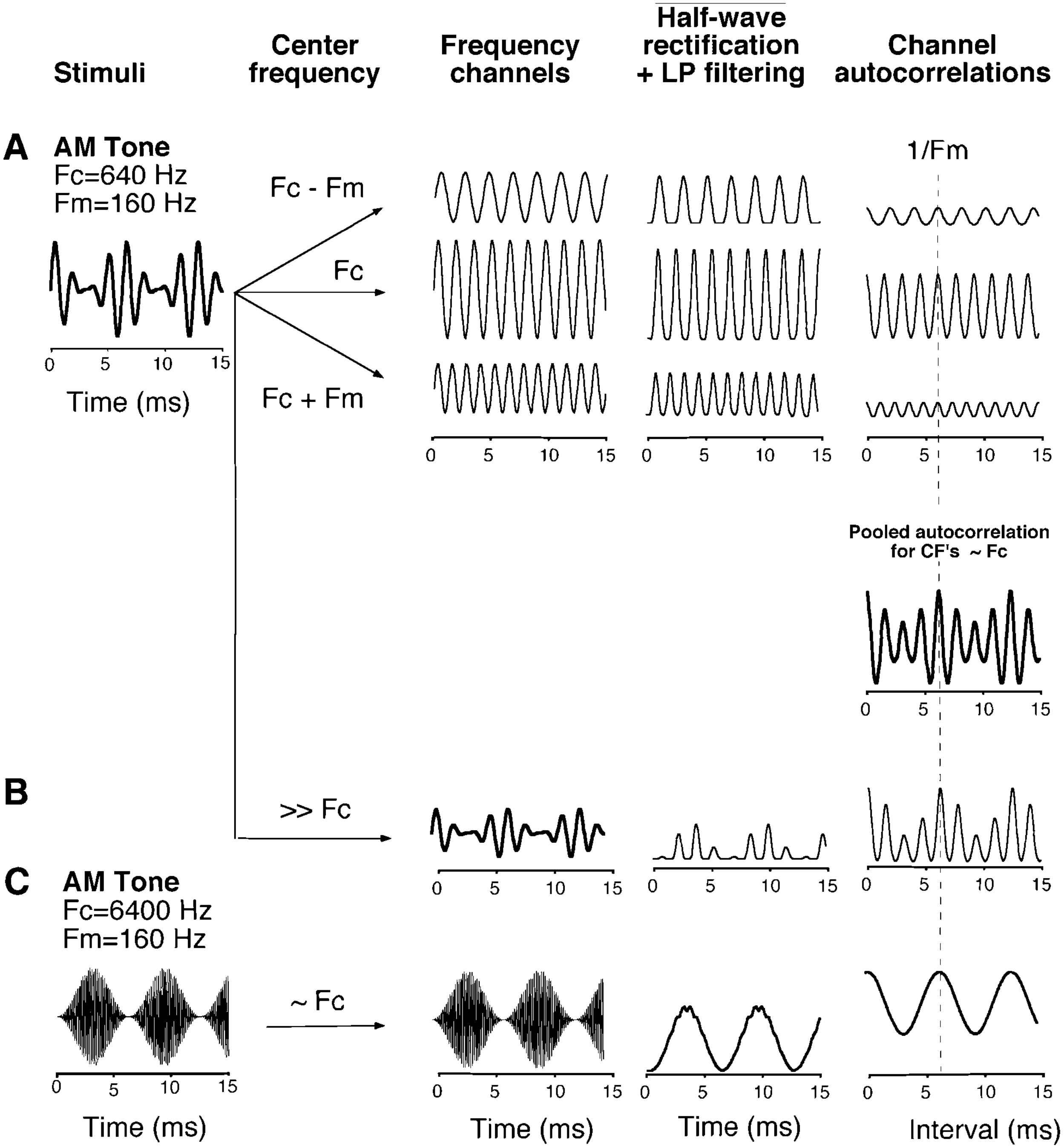
Pitch period



Level-Invariance

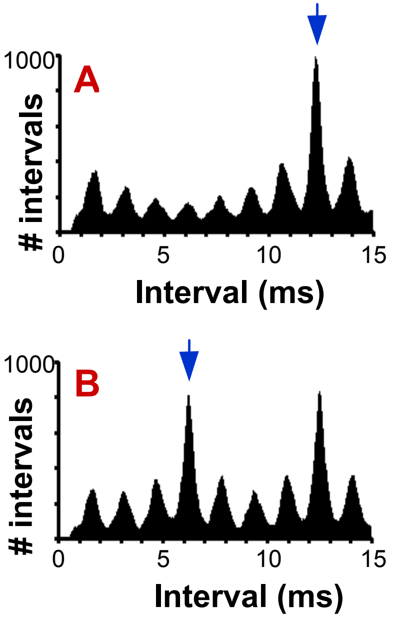


Pitch Equivalence


Stimuli
Center frequency
Frequency channels
Half-wave rectification + LP filtering
Channel autocorrelations
A
AM Tone
Fc=640 Hz
Fm=160 Hz
Time (ms)
Fc - Fm
Fc
Fc + Fm
1/Fm
0
5
10
15
Pooled autocorrelation for CF's ~ Fc
B
>> Fc
C
AM Tone
Fc=6400 Hz
Fm=160 Hz
~ Fc
Time (ms)
Time (ms)
Time (ms)
Interval (ms)

The running population-interval distribution

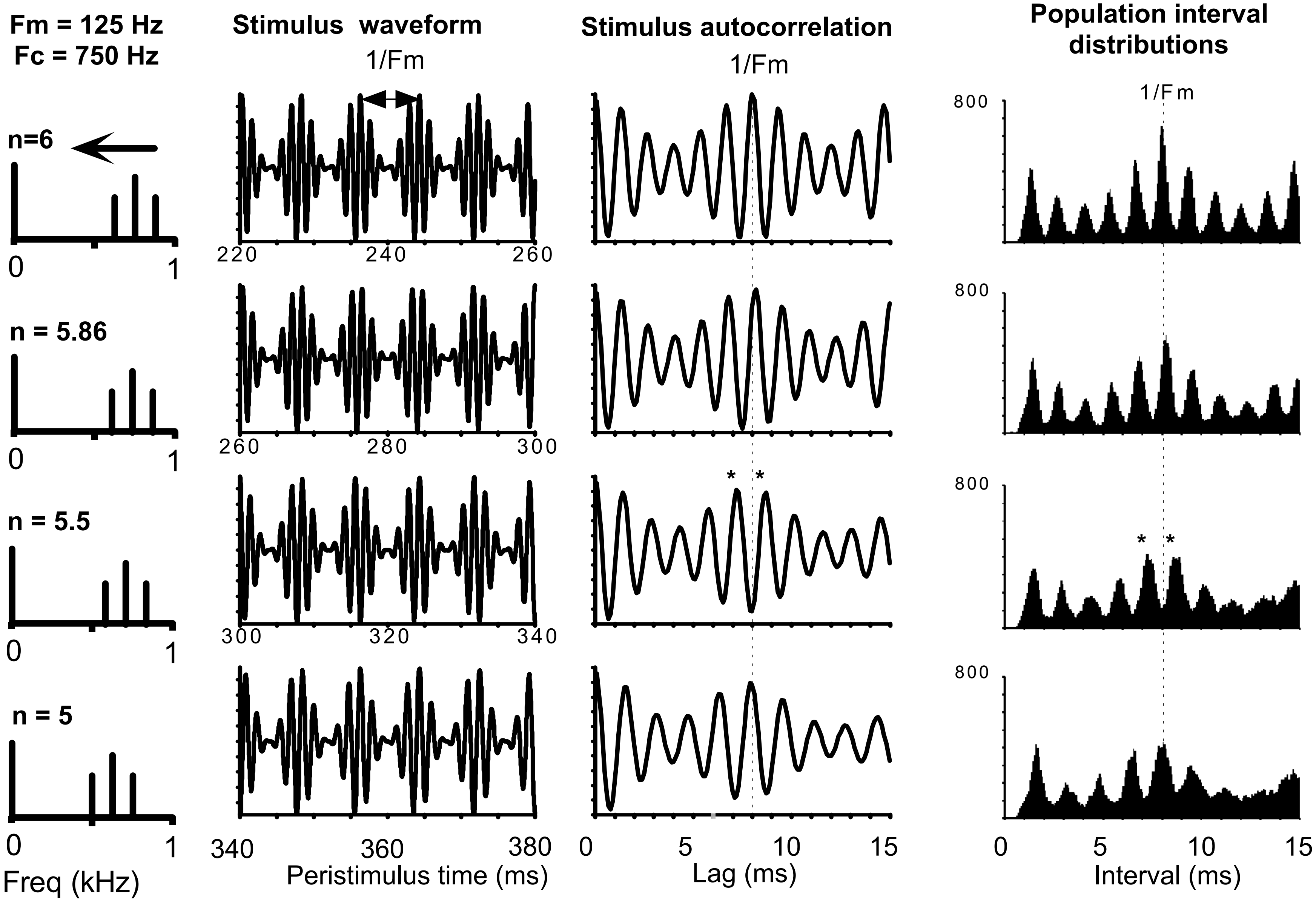
Population interval histograms (cross sections)



Pitch height and pitch chroma

Please see Figure 1, 2, and 7 in Roger N. Shepard. "Geometrical Approximations to the Structure of Musical Pitch."*Psychological Review* 89 no. 4 (1982): 305-322.

Pitch shift of inharmonic complex tones

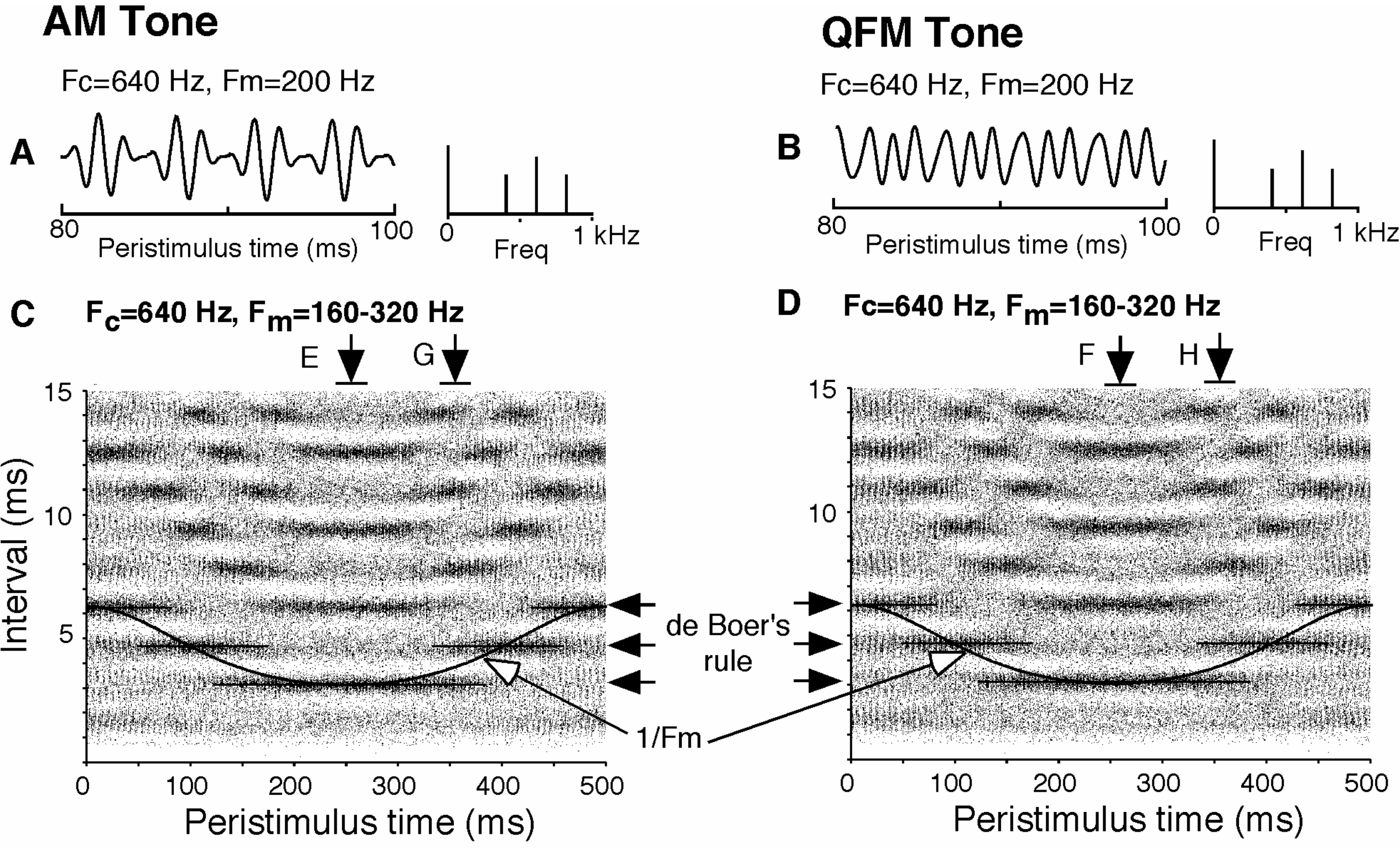

Fm = 125 Hz
Fc = 750 Hz

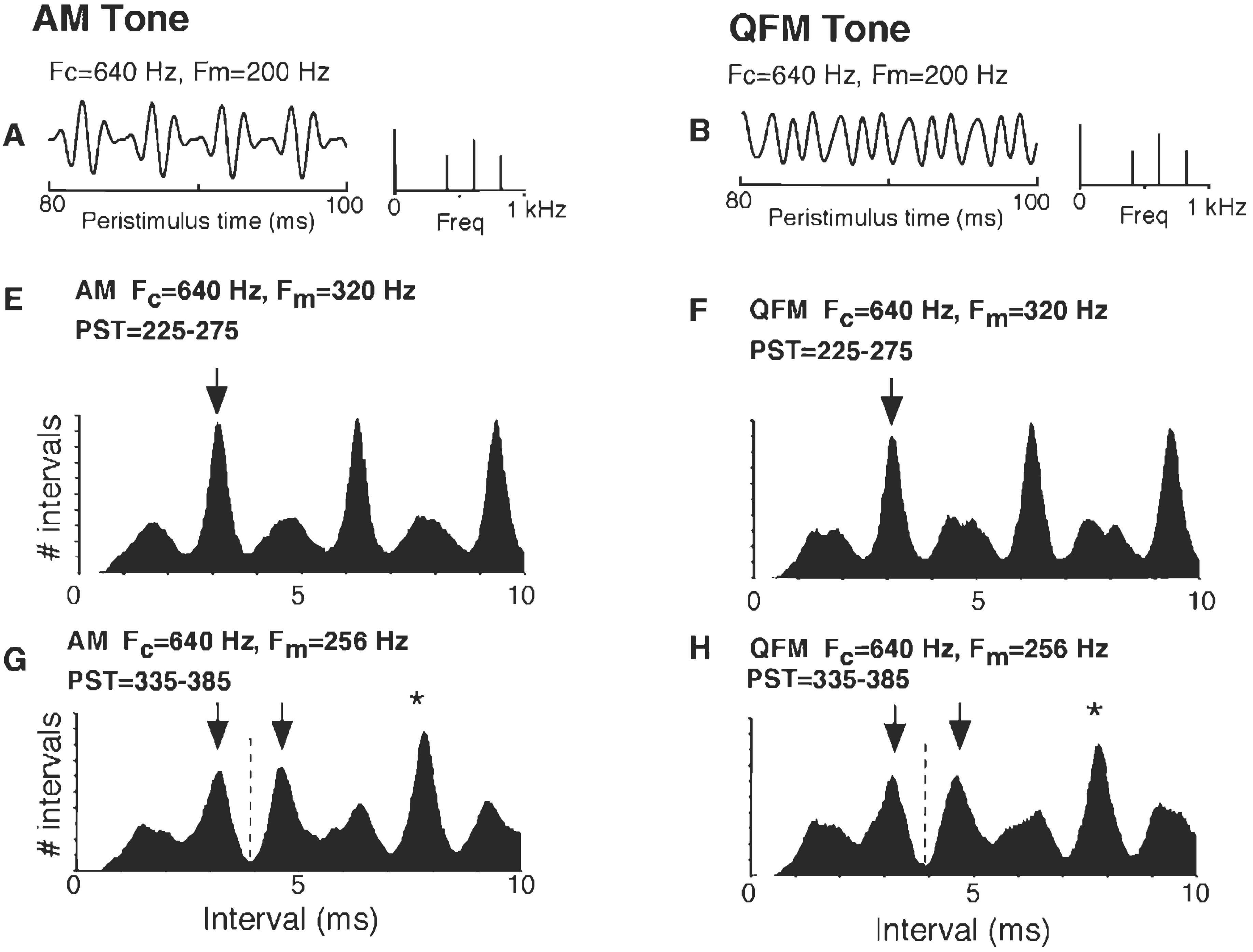
Stimulus waveform

Stimulus autocorrelation

Population interval distributions

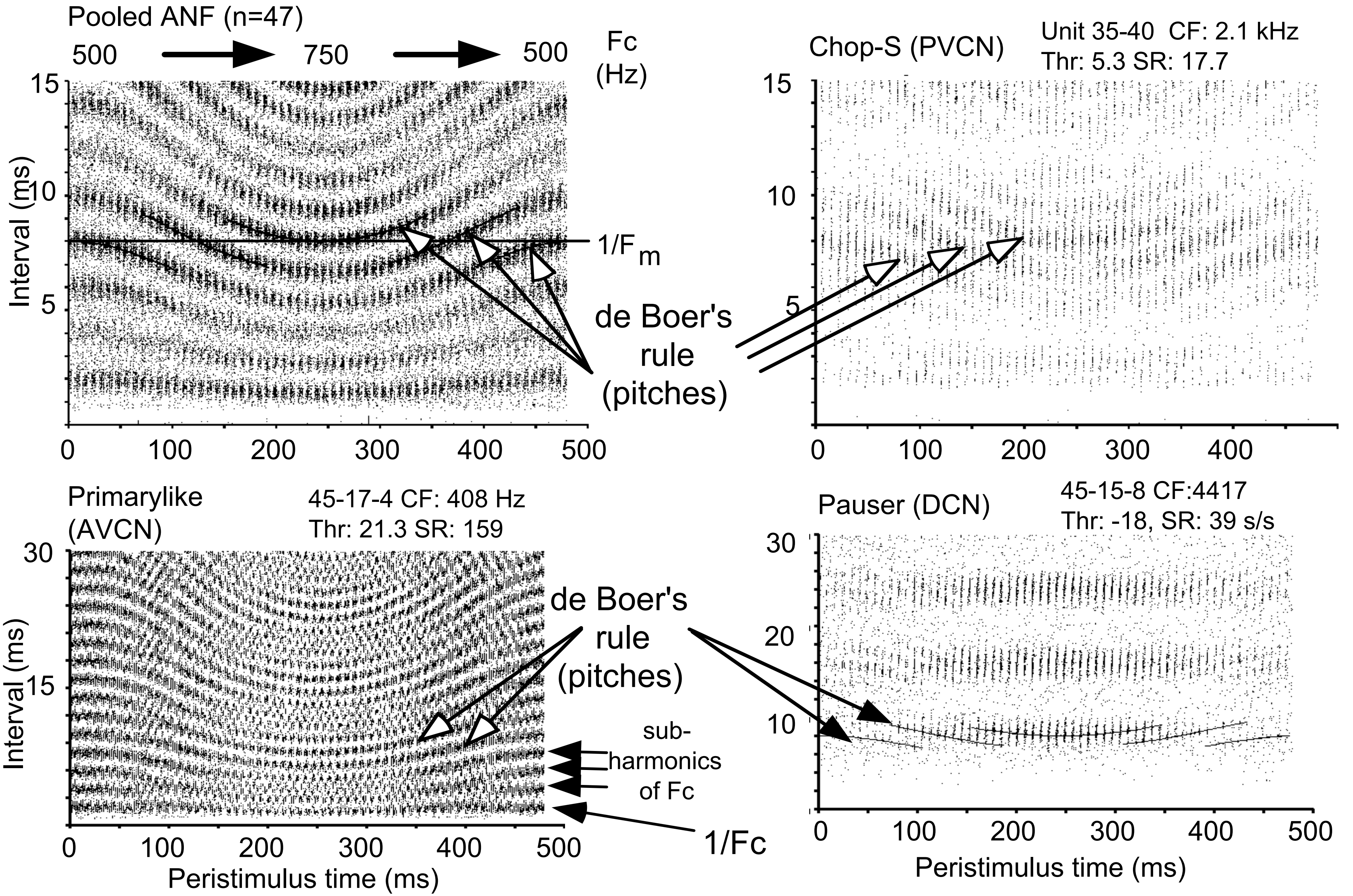
Pitch shift of inharmonic complex tones
Phase-invariant nature of all-order interval code



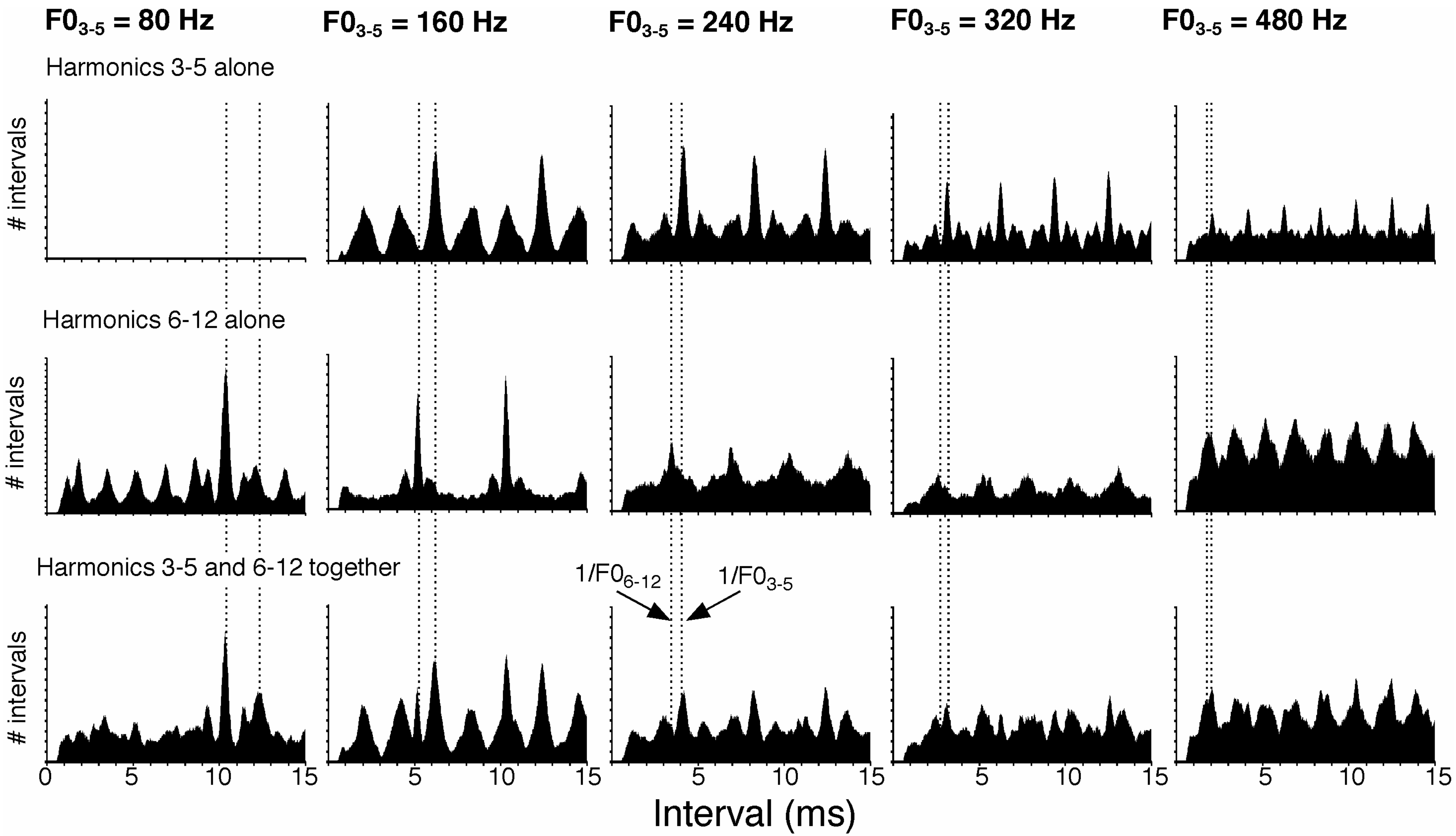

AM Tone
Fc=640 Hz, Fm=200 Hz
A
80
Peristimulus time (ms)
100
0
Freq
1 kHz
QFM Tone
Fc=640 Hz, Fm=200 Hz
B
80
Peristimulus time (ms)
100
0
Freq
1 kHz
E
AM F_c=640 Hz, F_m=320 Hz
PST=225-275
# intervals
0
5
10
F
QFM F_c=640 Hz, F_m=320 Hz
PST=225-275
0
5
10
G
AM F_c=640 Hz, F_m=256 Hz
PST=335-385
*
intervals
0
5
10
Interval (ms)
H
QFM F_c=640 Hz, F_m=256 Hz
PST=335-385
*
0
5
10
Interval (ms)

Cochlear nucleus IV: Pitch shift

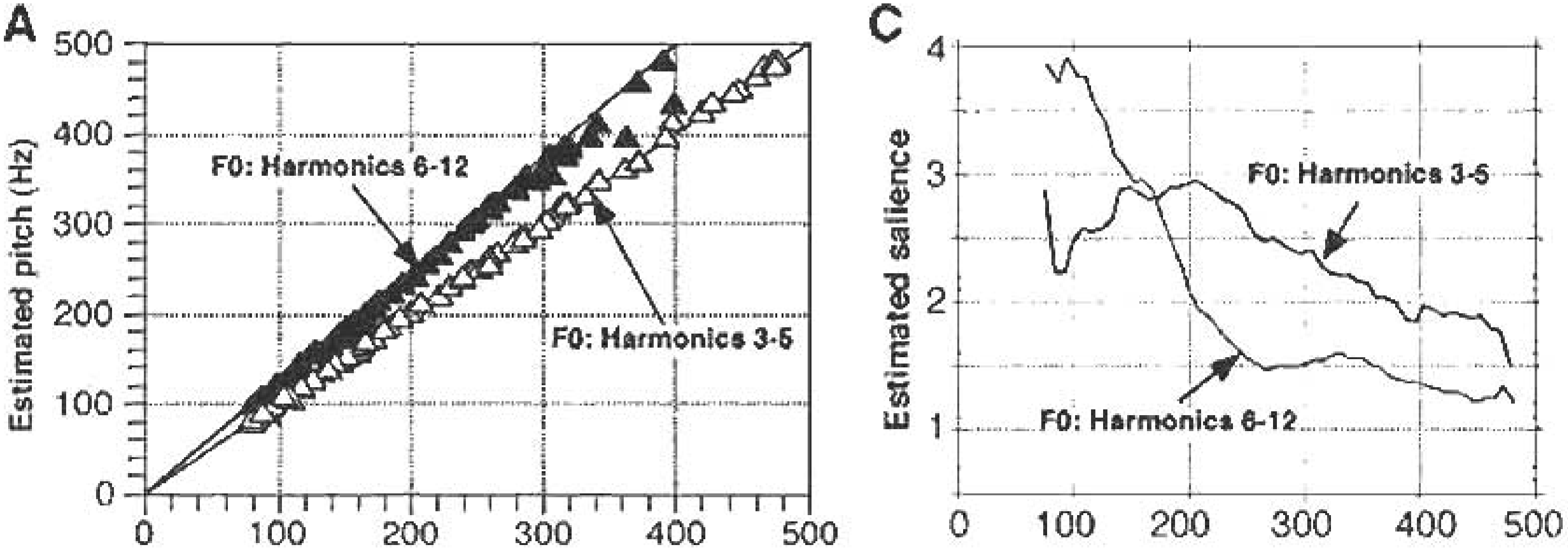
Variable-Fc AM tone F_m = 125 Hz F_c = 500-750 Hz Pitch ~ de Boer's rule



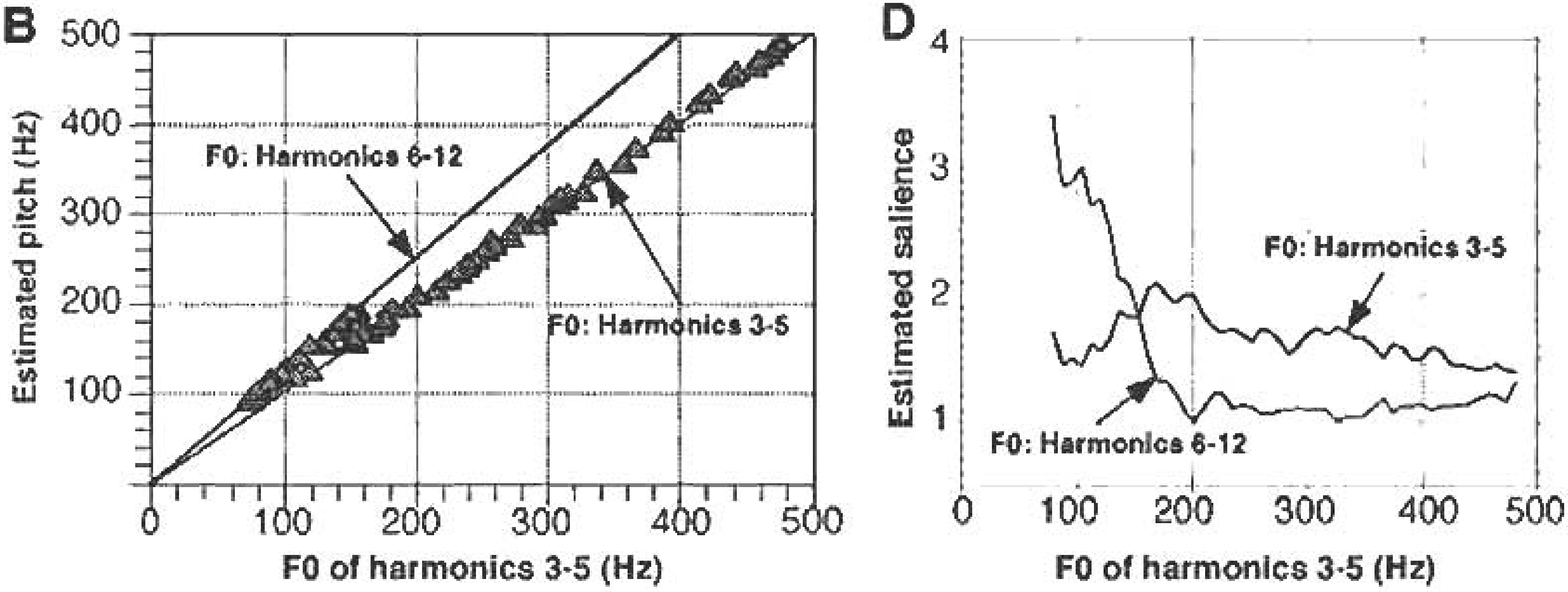
Dominance region for pitch (harmonics 3-5 or partials 500-1500 Hz)



Harmonics 3-5 and 6-12 presented separately



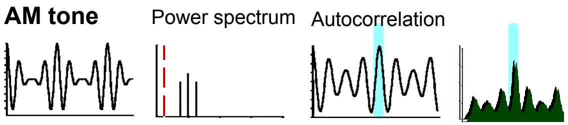
Harmonics 3-5 and 6-12 presented concurrently



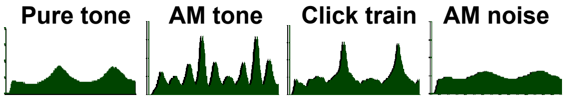
Summary

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

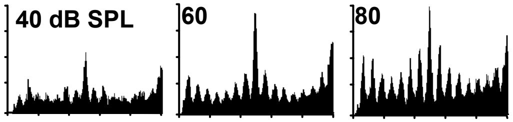
Pitch of the "missing fundamental"



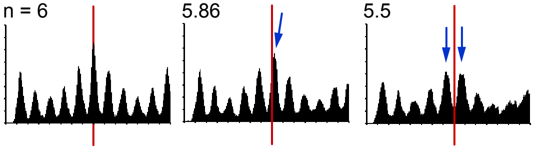
Pitch Equivalence



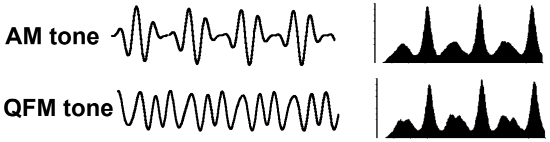
Level invariance



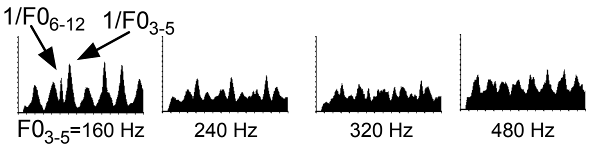
Pitch shift of inharmonic AM tones



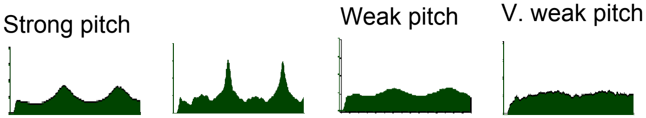
Phase invariance



Dominance region



Pitch salience



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)

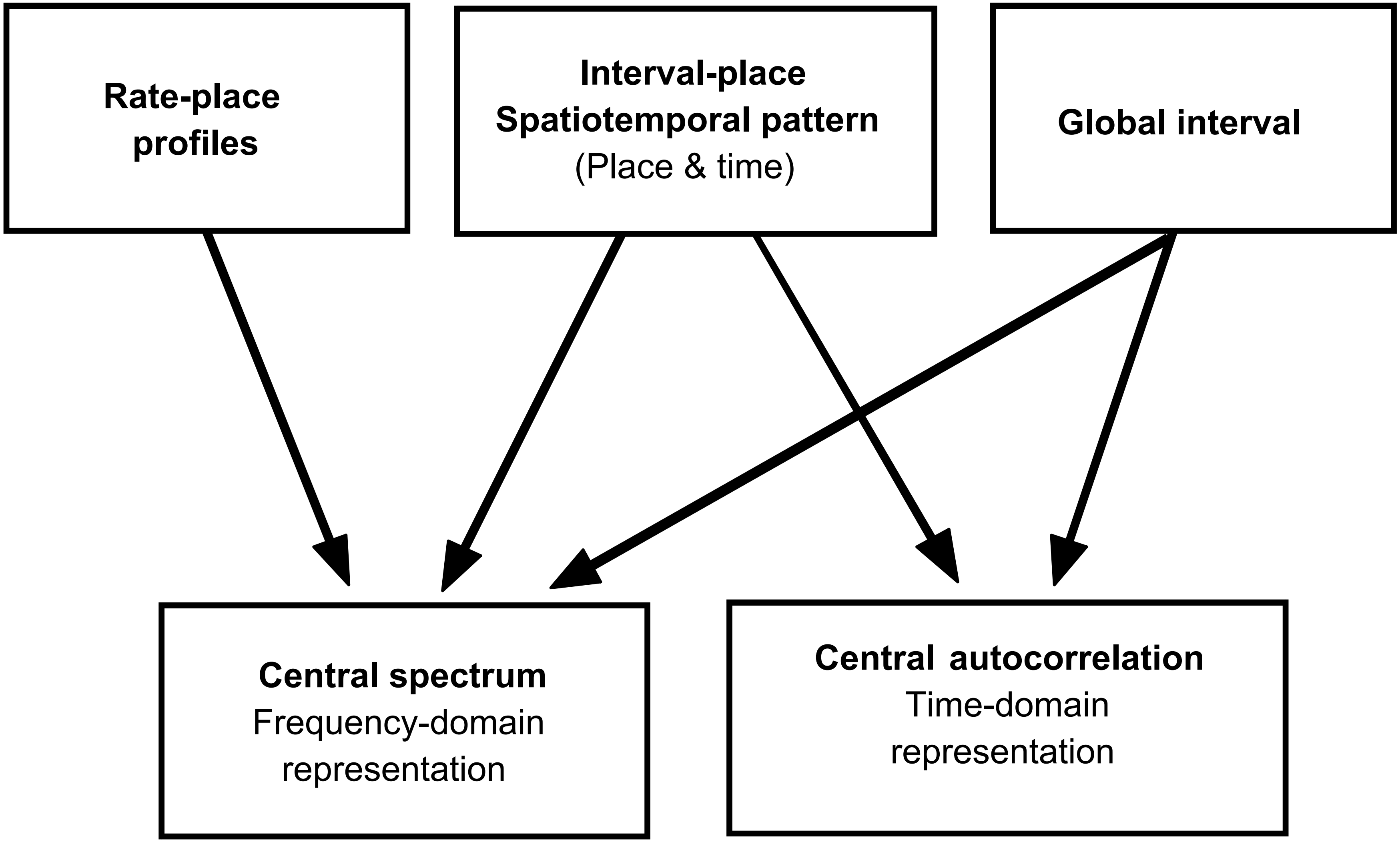
Explain well existence region for F0 (albeit with limits on max interval durations)

Do not explain low pitches of unresolved harmonics

Interval analyzers require precise delays & short coincidence windows

Little or no neural evidence for required analyzers

Physiological and functional representations



Rate-place profiles

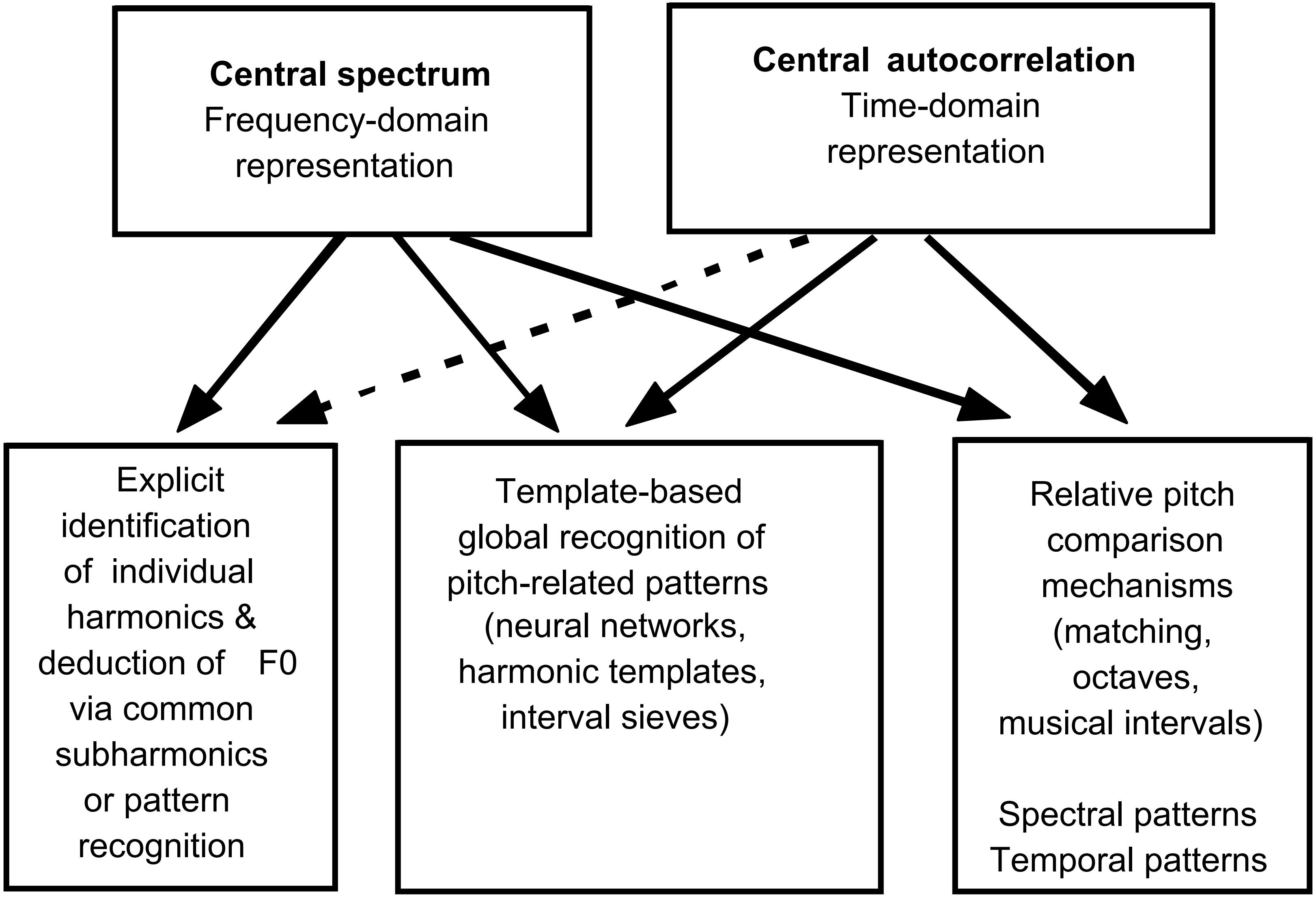
Interval-place
Spatiotemporal pattern
(Place & time)

Global interval

Central spectrum
Frequency-domain representation

Central autocorrelation
Time-domain representation

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



Central spectrum
Frequency-domain
representation

Central autocorrelation
Time-domain
representation

Explicit
identification
of individual
harmonics &
deduction of F0
via common
subharmonics
or pattern
recognition

Template-based
global recognition of
pitch-related patterns
(neural networks,
harmonic templates,
interval sieves)

Relative pitch
comparison
mechanisms
(matching,
octaves,
musical intervals)

Spectral patterns
Temporal patterns

Reading/assignment for next meeting

- **Tuesday. Feb. 24**
- **Pitch mechanisms, continued**
- **Perception of timbre**

- **Reading:**

Moore Chapter 8, pp. 269-273

Chapter in Deutsch by Risset & Wessel on Timbre

(first sections up to p. 113-118)

Handel, chapter on Identification

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

OCTAVE SIMILARITY

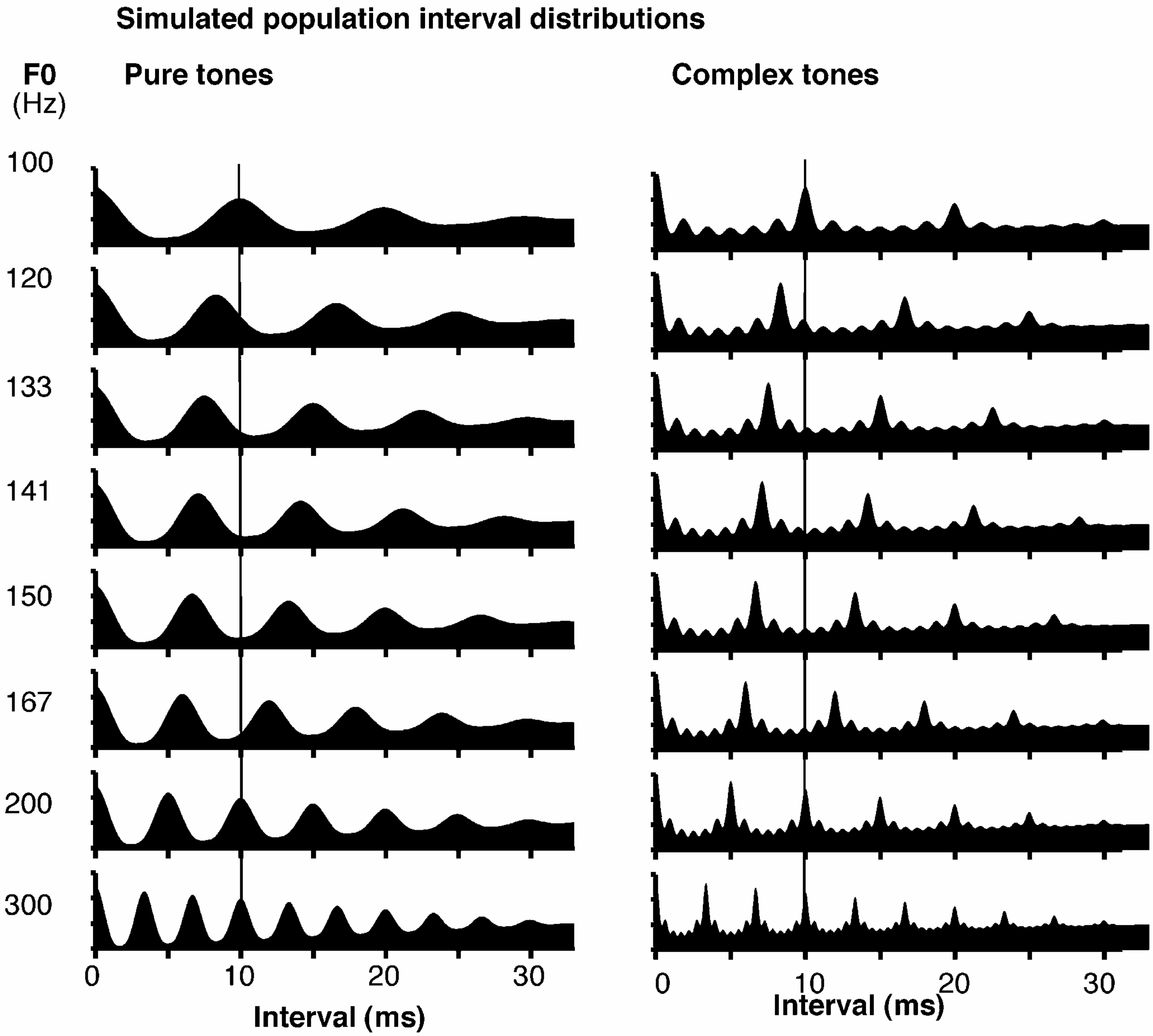


Figure 4. Similarities between population-interval representations associated with different fundamental frequencies. Simulated population-interval distributions for pure tones (left) and complex tones (right) consisting of harmonics 1-6.

Octave similarity

Please see Cariani, Peter. *Journal of New Music Research* 30 no. 2 (2001): 107-135.

Pitch height and pitch chroma

Please see Figure 1, 2, and 7,in Roger N. Shepard. "Geometrical Approximations to the Structure of Musical Pitch." *Psychological Review* 89 no. 4 (1982): 305-322.

Please see Cariani, Peter. *Journal of New Music Research*. 30 no. 2 (2001): 107-135.

Musical tonal relations

Please see Roger N. Shepard. "Geometrical Approximations to the Structure of Musical Pitch." *Psychological Review* 89 no. 4 (1982): 305-322.

Please see Cariani, Peter. *Journal of New Music Research*. 30 no. 2 (2001): 107-135.