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

HST 725

Music Perception & Cognition

Lecture 3

What we hear:

Basic dimensions of auditory experience

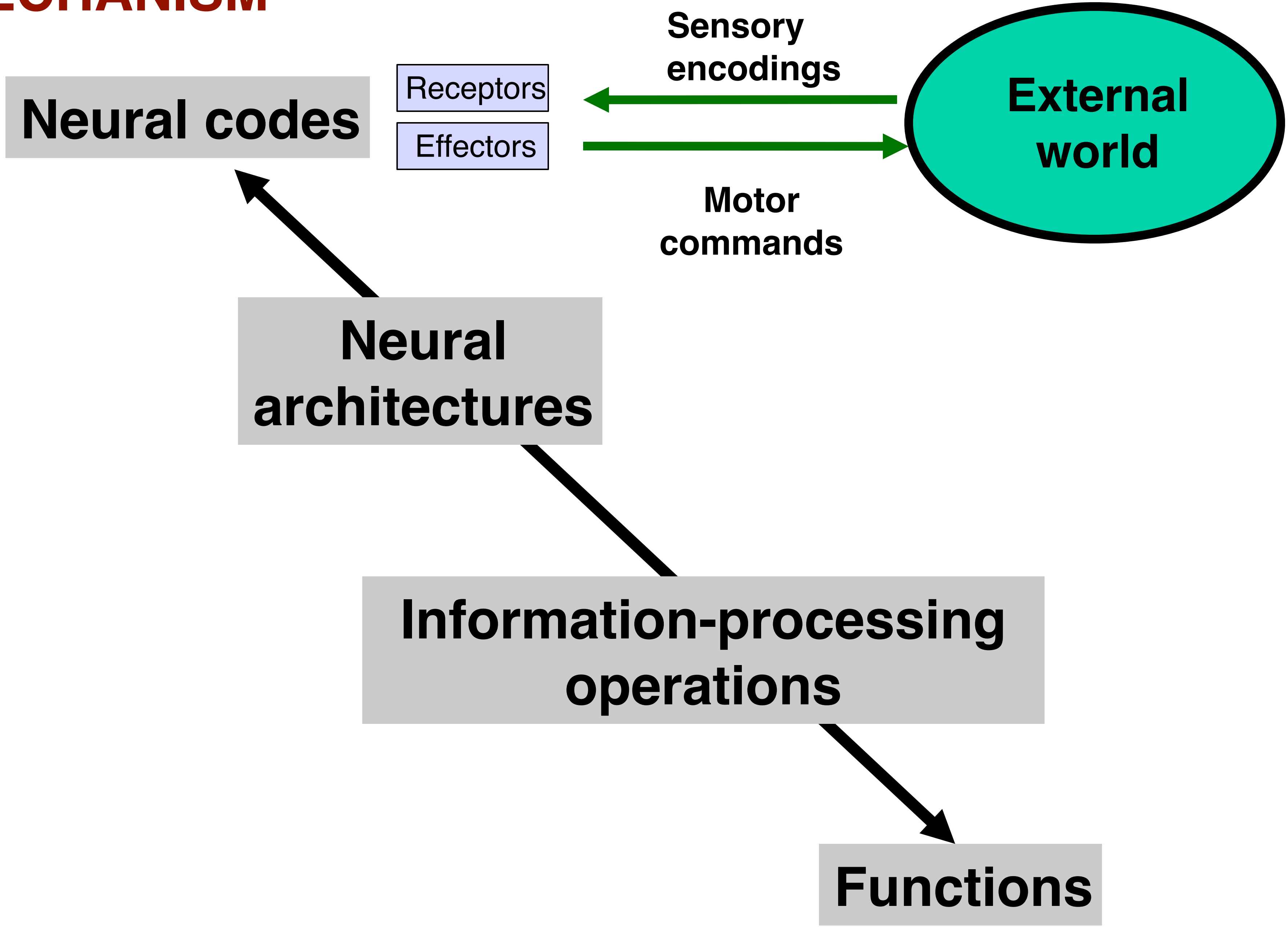
www.cariani.com

Tuesday, February 10, 2009

What we hear: dimensions of auditory experience

- Hearing: ecological functions (distant warning, communication, prey detection; works in the dark)
- Detection, discrimination, recognition, reliability, scene analysis
- Operating range: thresholds, ceilings, & frequency limits
- Independent dimensions of hearing & general properties
  - Pitch
  - Timbre (sound quality)
  - Loudness
  - Duration
  - Location
  - Distance and Size
- Perception of isolated pure tones
- Interactions of sounds: beatings, maskings, fusions
- Masking (tones vs. tones, tones in noise)
- Fusion of sounds & the auditory "scene":
  - how many objects/sources/voices/streams?
- Representation of periodicity and spectrum

MECHANISM


Sensory encodings
Neural codes
Receptors
Effectors
External world
Motor commands
Neural architectures
Information-processing operations
Functions

Hearing: ecological functions

- Distant warning of predators approaching
- Identification of predators
- Localization/tracking of prey
- Con-specific communication

Mating/competition

Cooperation (info. sharing)

Territory

- Navigation in the dark
- General recognition of sounds

http://www.pbs.org/wgbh/nova/wolves/

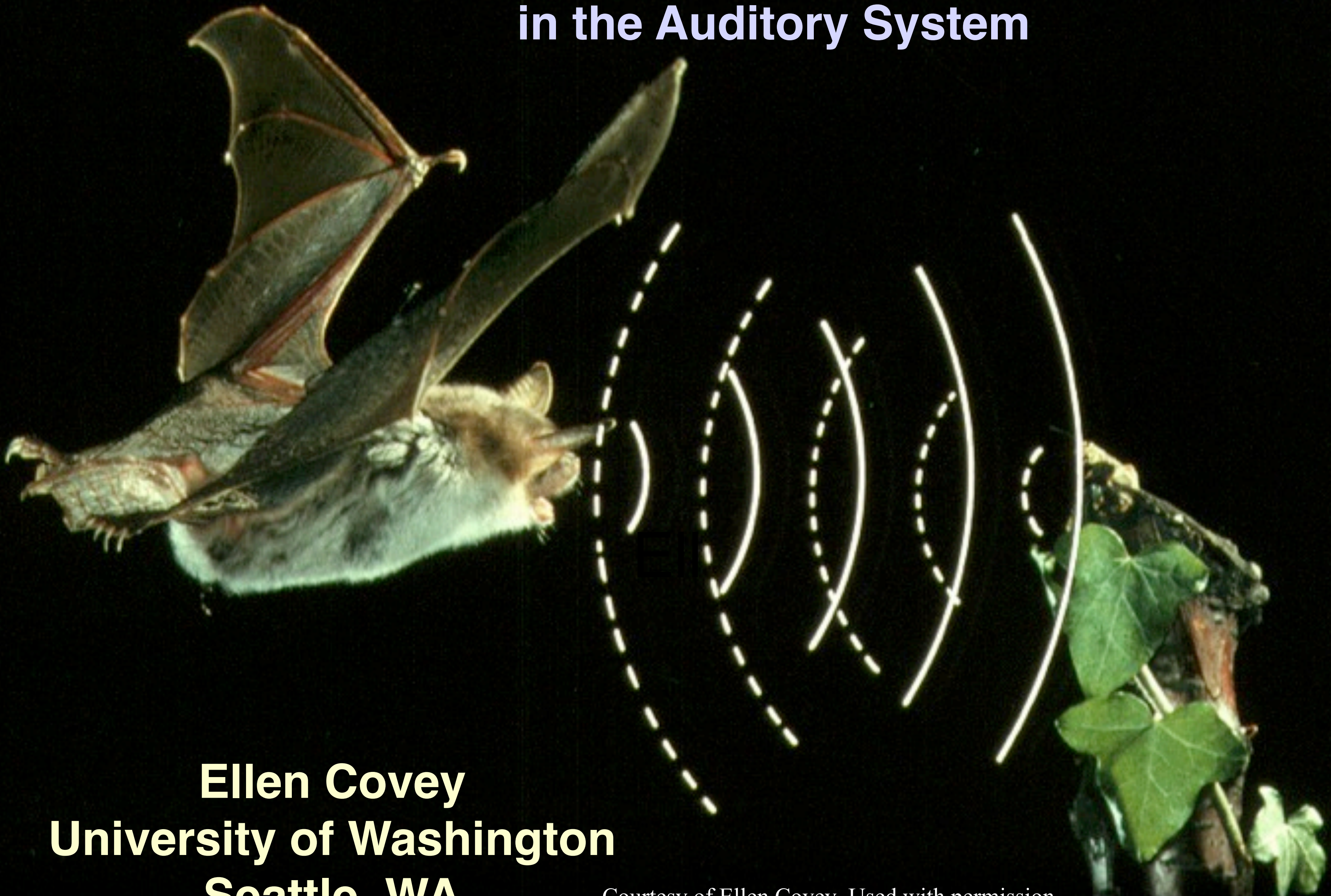
http://www.pbs.org/lifeofbirds/songs/index.html

http://www.batsnorthwest.org/

Corynorhinus townsendii - Townsend's Big-eared Bat

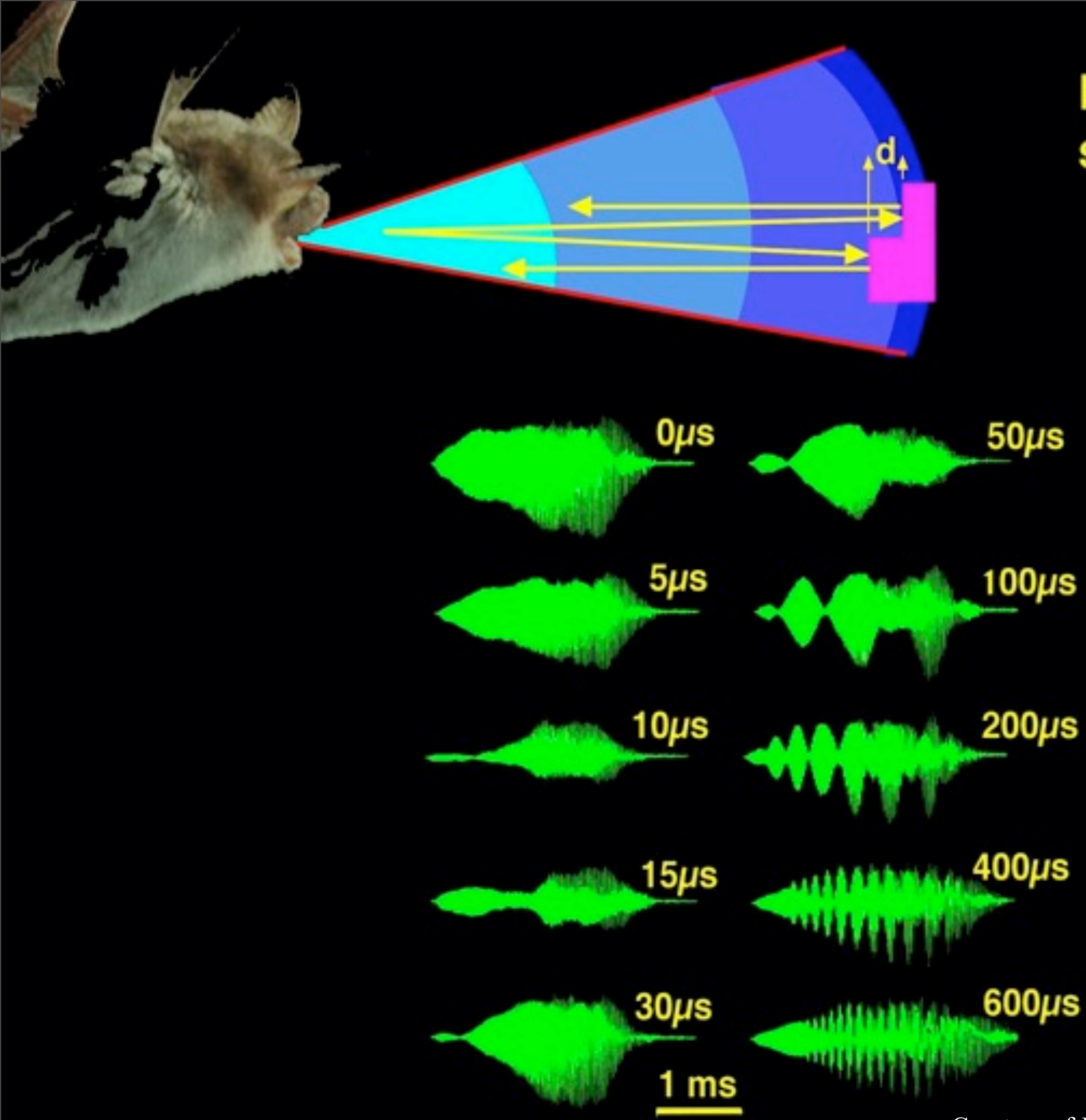
Photo: US Fish and Wildlife Service.

The Effect of Context on Information Processing in the Auditory System



Ellen Covey
University of Washington
Seattle, WA

Echoes from a 2-surface object



Jim Simmons, Brown

The auditory scene: basic dimensions

Temporal organization
- **Events**
- **Notes**
- **Temporal patterns of events**

Organization of sounds
- **Voices, instruments**
- **Streams**
- **Objects**
- **Sources**

Attributes of sounds
- **Loudness** (intensity)
- **Pitch** (dominant periodicity)
- **Timbre** (spectrum)
- **Duration**
- **Location** (bearing, range)

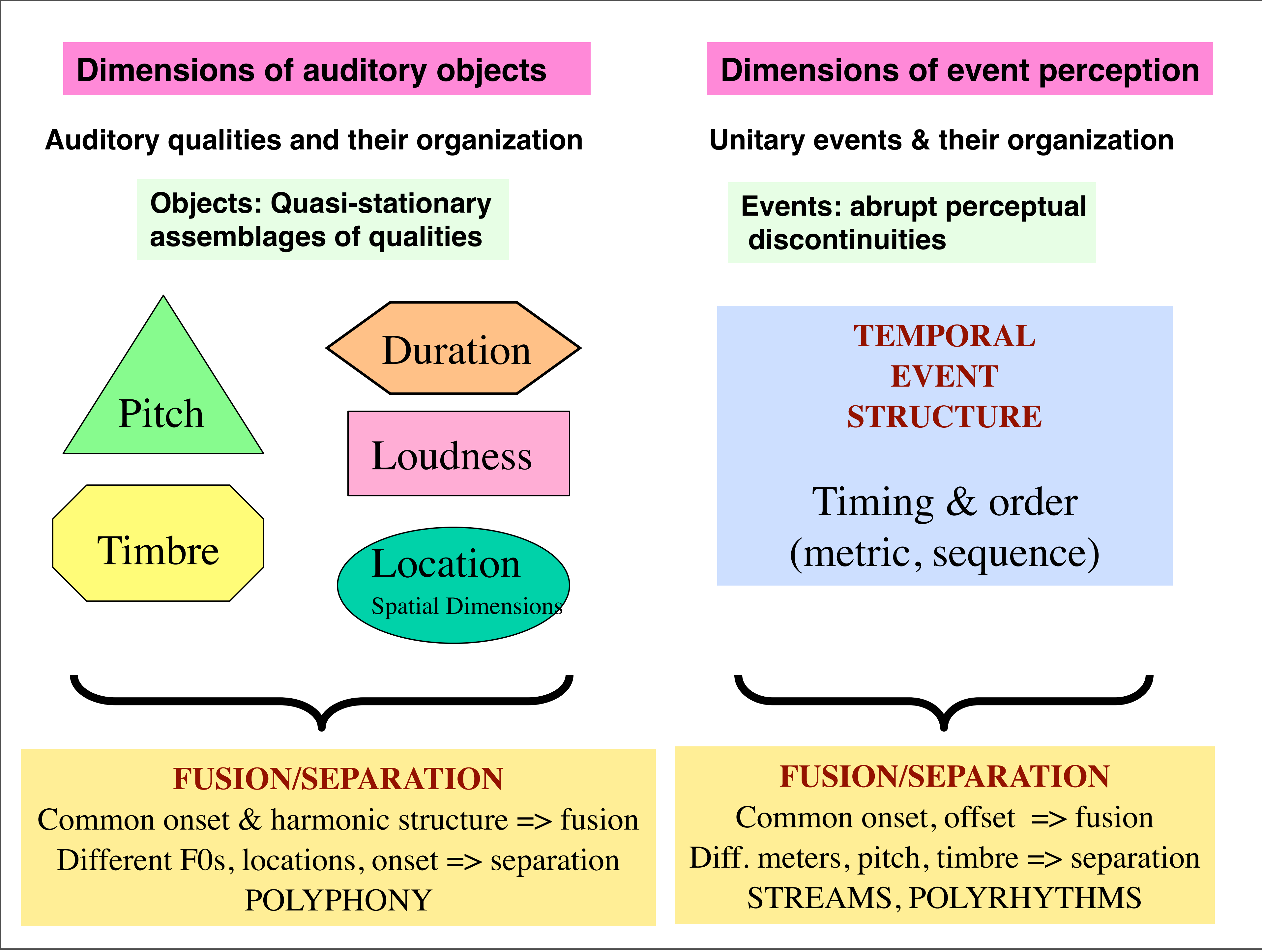
Paul Cezanne. "Apples, Peaches, Pears and Grapes."
Courtesy of the iBilio.org WebMuseum.

Auditory qualities in music perception & cognition

- **Pitch** Melody, harmony, consonance
- **Timbre** Instrument voices
- **Loudness** Dynamics
- **Organization** Fusions, objects. How many voices?

- **Rhythm** Temporal organization of events
- **Longer pattern** Repetition, sequence

- **Mnemonics** Familiarity, novelty
- **Hedonics** Pleasant/unpleasant
- **Affect** Emotional associations, meanings
- **Semantics** Cognitive associations/expectations


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

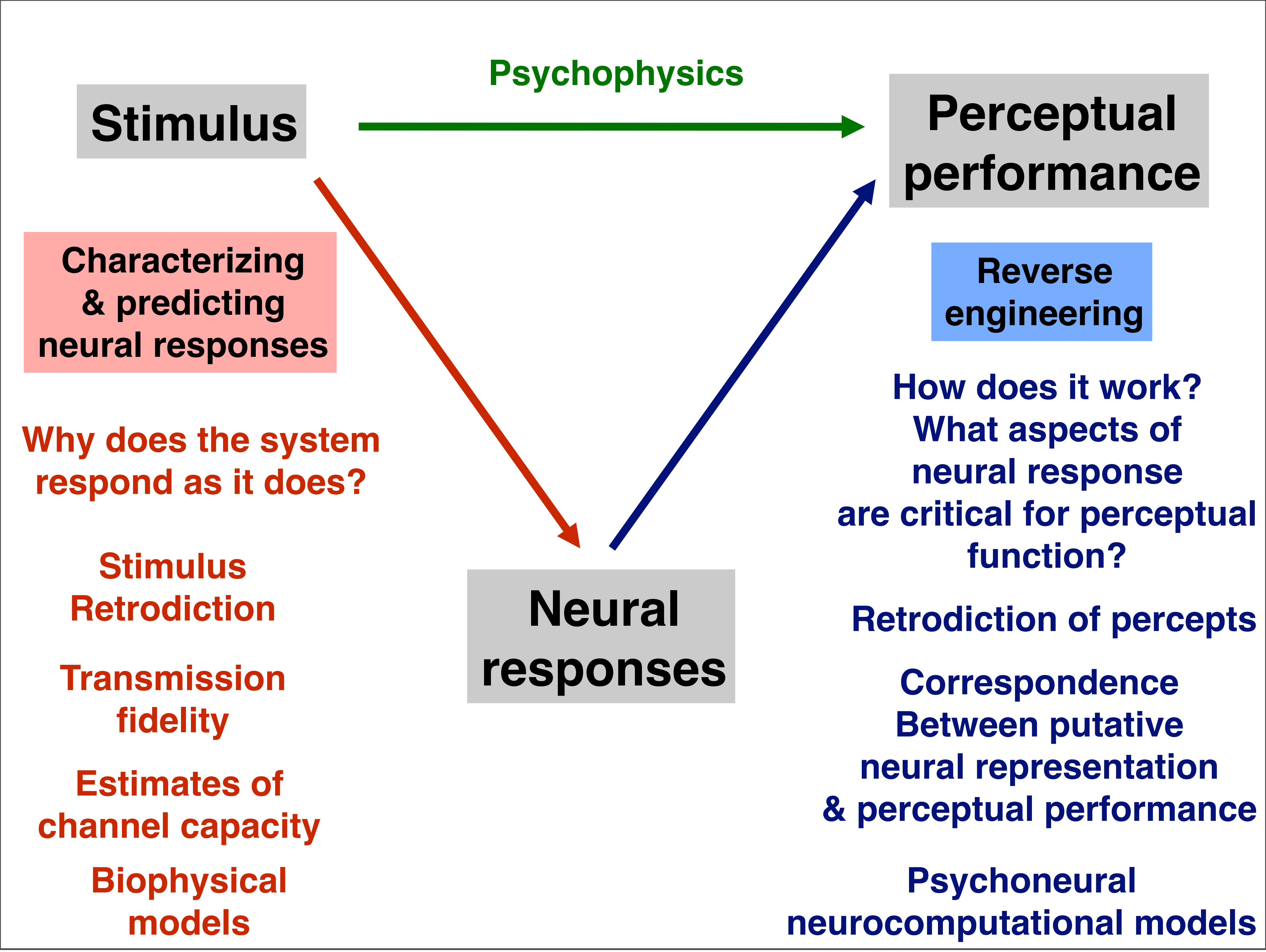
Visual scene

Line
Shape
Texture
Lightness
Color
Transparency

Objects
Apparent distance
Apparent size
etc.

Photo removed due to copyright restrictions. Cover of *LIFE Magazine*, Nov. 23, 1936.
View in Google Books.

Margaret Bourke-White
Fort Peck Dam 1936


Psychophysics
Stimulus
Perceptual performance
Characterizing & predicting neural responses
Reverse engineering
Why does the system respond as it does?
Stimulus Retrodiction
Transmission fidelity
Estimates of channel capacity
Biophysical models
Neural responses
How does it work? What aspects of neural response are critical for perceptual function?
Retrodiction of percepts
Correspondence Between putative neural representation & perceptual performance
Psychoneural neurocomputational models

Perceptual functions

Subjective vs. objective measures
Subjective measures
Magnitude estimation
Objective measures
Detection: capability of distinguishing the presence or absence of a stimulus (or some aspect of a stimulus, e.g. AM detection)
Threshold: the value of a stimulus parameter at which a stimulus can be reliably detected
Sensation level (SL): sound level re: threshold

Discrimination: capability of distinguishing between two stimuli
Difference limen: the change in a stimulus parameter required for reliable discrimination, just-noticeable-difference (jnd)
Weber fraction: Difference limen expressed as proportional change (e.g. $\Delta f/f$)

Matching task: subject changes parameter that matches two stimuli
Two-alternative forced choice (2AFC)
Ranking tasks

Recognition: correct identification of a particular stimulus
Masking: impairment of ability to detect a signal in the presence of other signals

Vibrations create compressions and expansions of air

Sound waves are alternating local changes in pressure

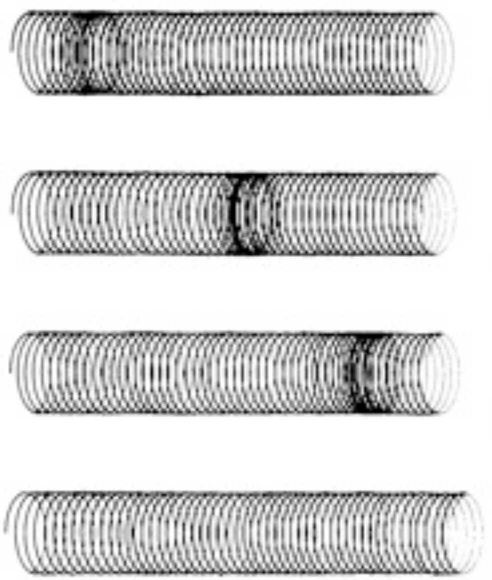
These changes propagate through space as “longitudinal” waves

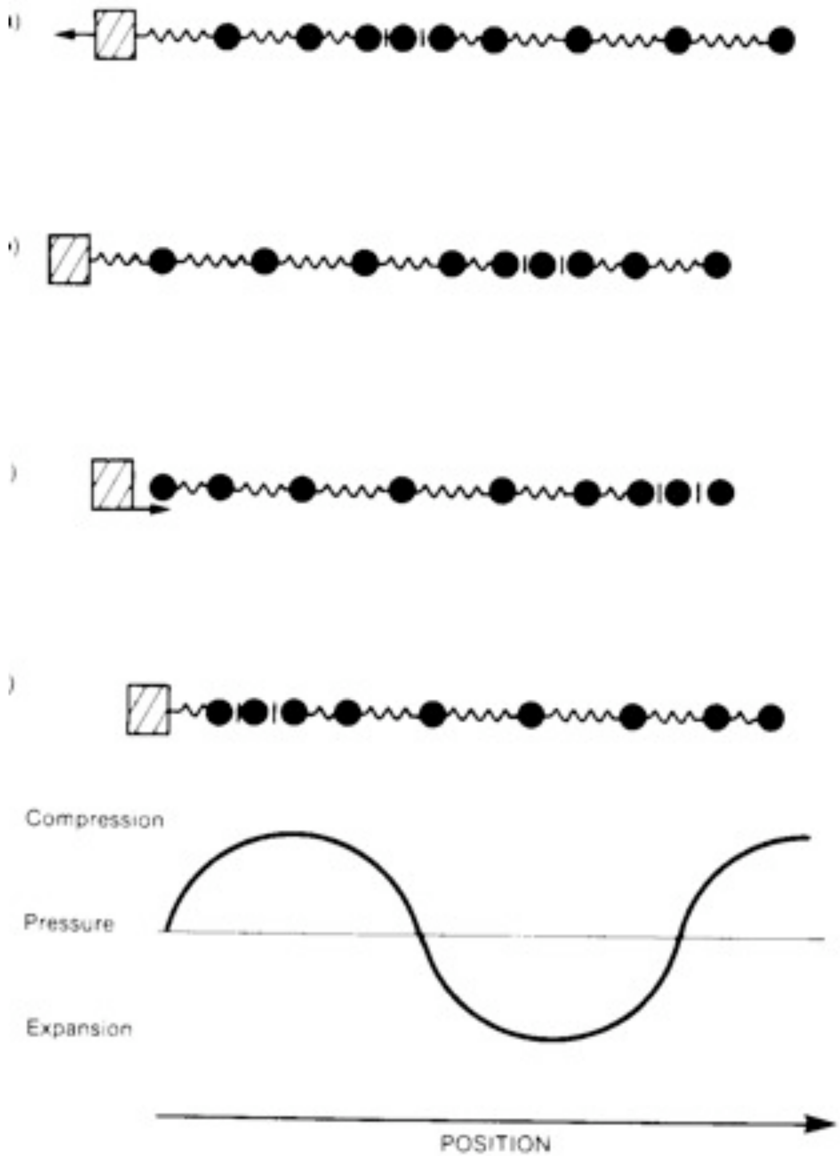
Condensation phase (compression):

pressure increases

Rarefaction (expansion) phase:

pressure decreases





Source: Handel, S. *Listening: an Introduction to the Perception of Auditory Events*. Cambridge, MA: MIT Press, 1989. Courtesy of MIT Press. Used with permission.

Handel

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.

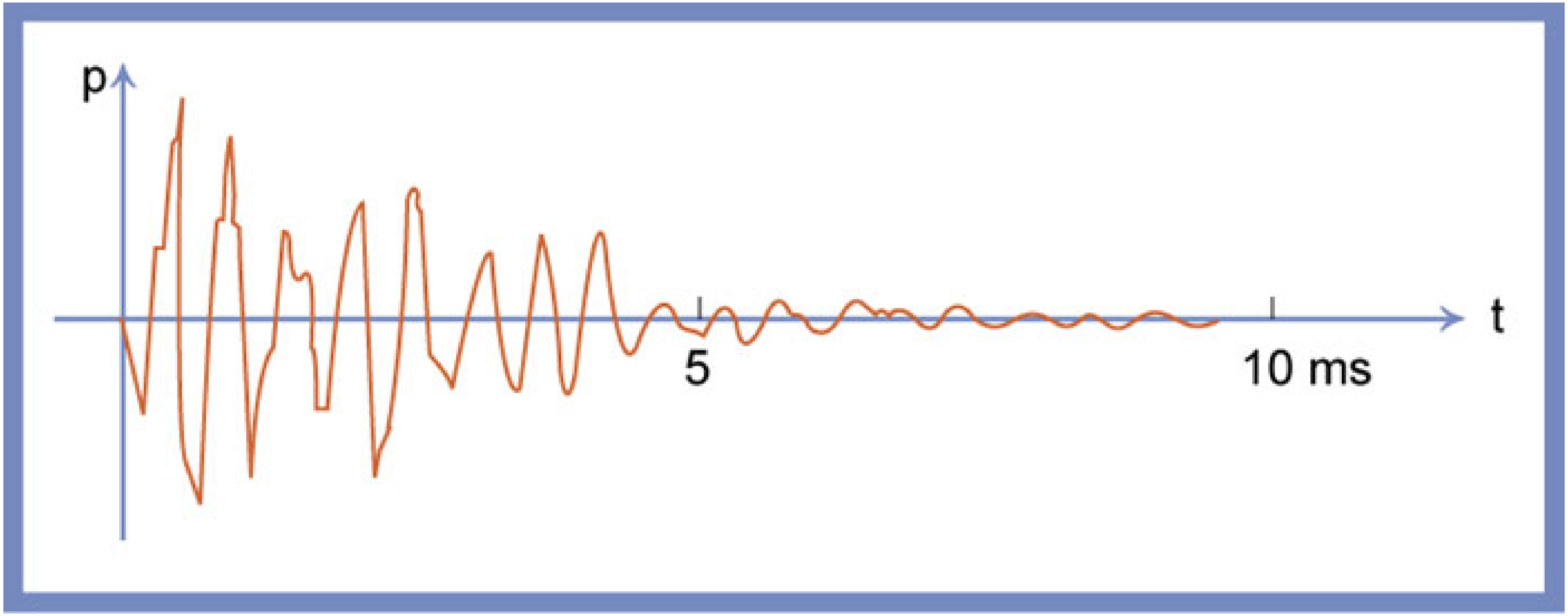
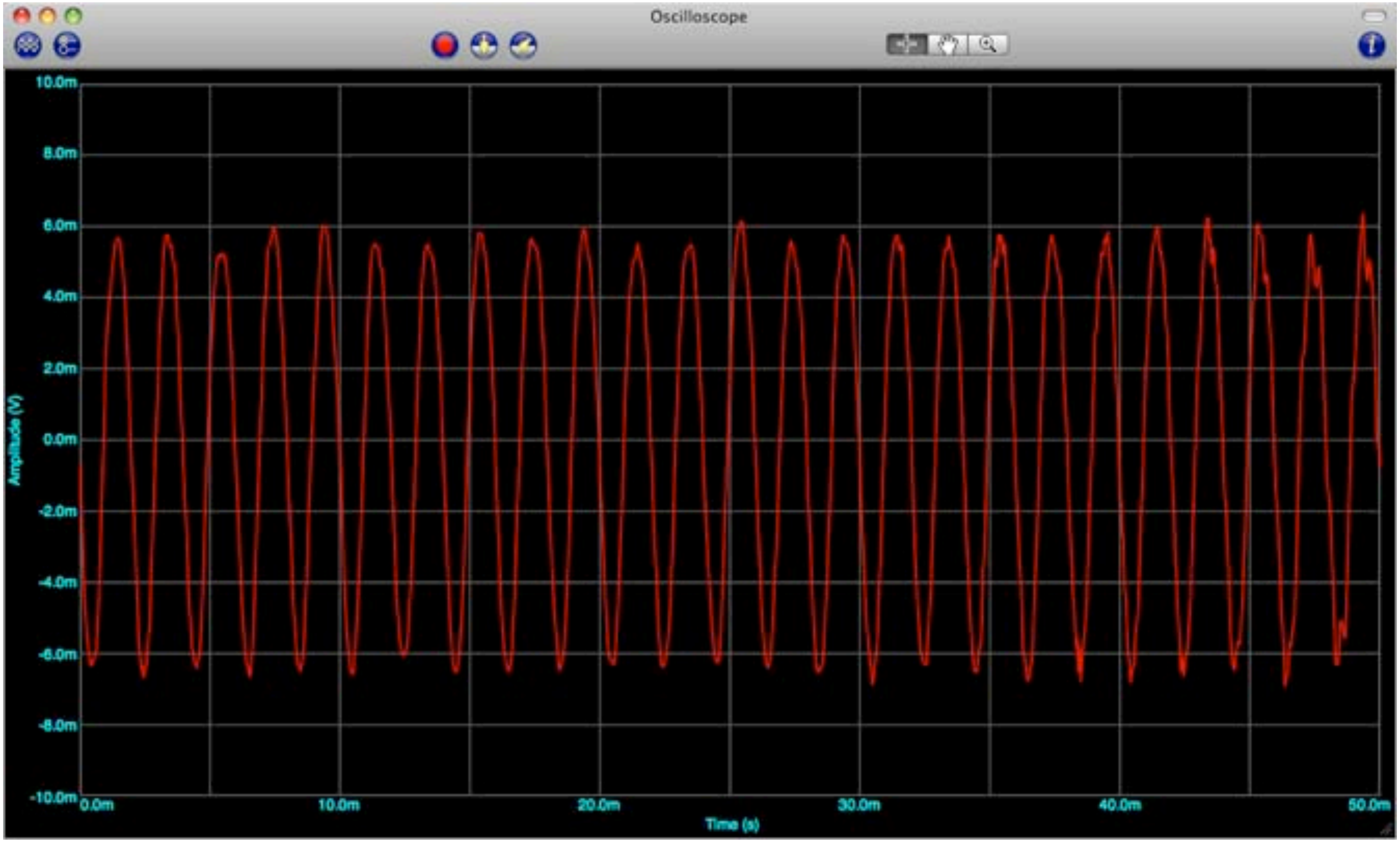


Figure by MIT OpenCourseWare.

Oscilloscope demonstration

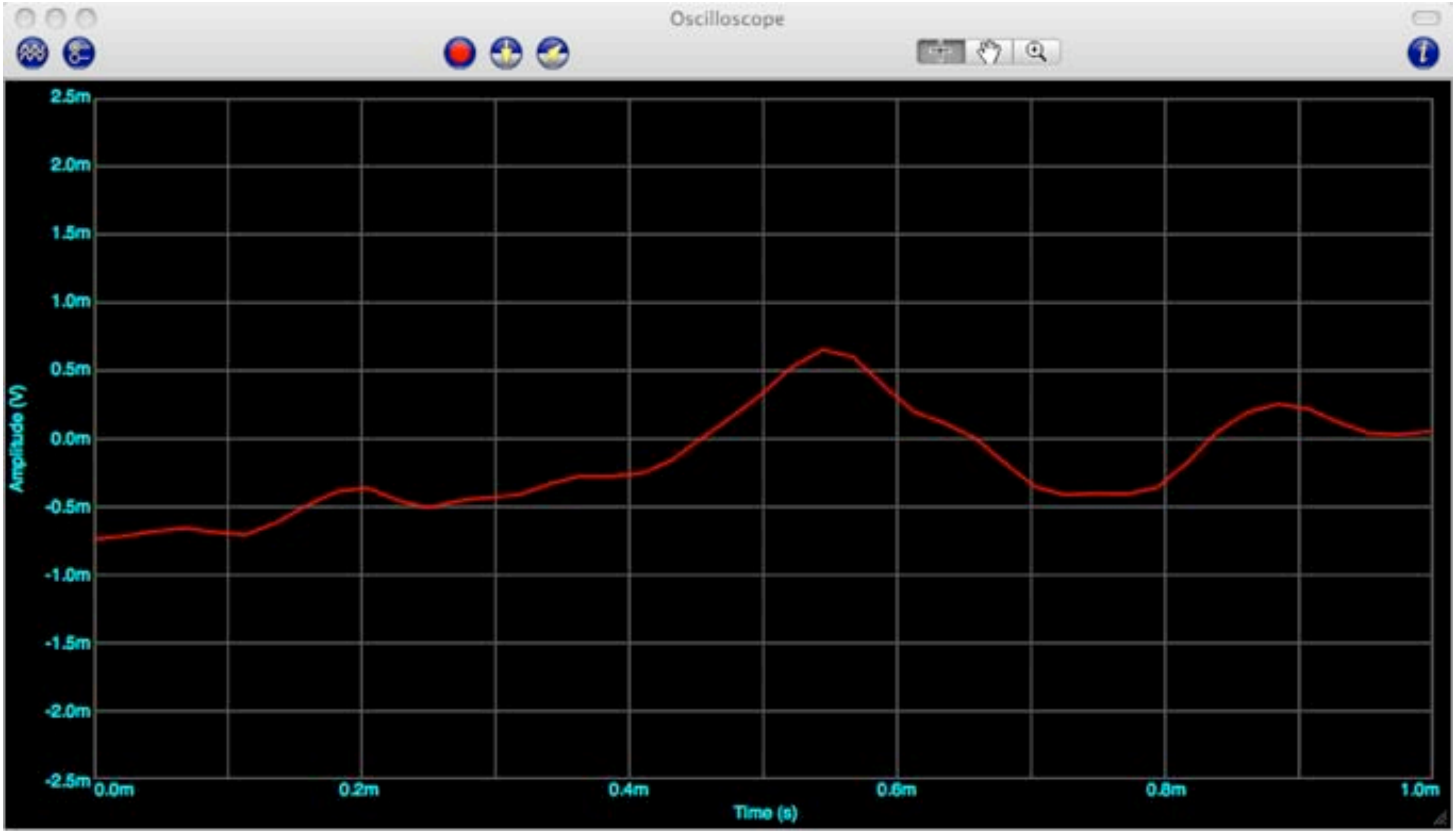
Waveforms plot pressure as a function of time, i.e. a "time-series" of amplitudes.

Waveforms are complete descriptions of sounds.



Sampling rate (samples/second)

Nowadays sounds are usually converted to strings of numbers that indicate sound pressure or voltage at each of many equally spaced points in time. THe number of samples collected per sec is the sampling rate.



From sound to numbers

CD quality sound
16 bits = 2^{16} = 64k voltage levels,
sampling rate @ 44,100 samples/sec

sound pressure changes

↓ **microphone**

electrical voltage changes

↓ **digitizer (analog to digital converter)**

numerical values, time-series

The **upper limit of human hearing** is **~ 20,000 cycles/sec** (Hertz, Hz).
There is a theorem in signal processing mathematics that the highest frequency that can be represented is 1/2 the sampling rate (called the Nyquist frequency).
This is why sound for CDs is sampled at 44.1 kHz.
In theory, this is the point where all sound distinctiona we can hear is captured.
MP3's compress the description by about 10-fold (we will discuss later).

Sound level basics

- Sound pressure levels are measured relative to an absolute reference
- (re: 20 micro-Pascals, denoted Sound Pressure Level or SPL).
- Since the instantaneous sound pressure fluctuates, the average amplitude of the pressure waveform is measured using root-mean-square RMS. (Moore, pp. 9-12)
- Rms(x) = sqrt(mean(sum(x_t^2)))
  - Where x_t is the amplitude of the waveform at each instant t in the sample
  - Because the dynamic range of audible sound is so great, magnitudes are expressed in a logarithmic scale, decibels (dB).
- A decibel of amplitude expresses the ratio of two amplitudes (rms pressures, P1 and P_reference) and is given by the equation:

dB = 20 * log10(P1/P_reference)

20 dB = 10 fold change in rms level

Decibel scale for relative amplitudes (levels) (rules of thumb)

20 dB = fold change amplitude

10 dB = 3+ fold change

6 dB = 2 fold change amplitude

3 dB = 1.4 fold change

2 dB = 1.26 fold change (26 %)

1 dB = 1.12 fold change (12%)

0 dB = 1 fold change (no change)

-6 dB = 1/2

-20 dB = 1/10 fold change

Dynamic range

0 dB SPL is set at 20 microPascals

60 dB SPL is therefore a 1000 fold change in RMS over 0 dB

A typical background sound level is 50-60 dB SPL.

Dynamic range describes the range of sound pressure levels.

The auditory system registers sounds from 20 dB to >> 120 dB SPL

The auditory system has a dynamic range in excess of 100 dB (!) or a factor of 10^5 = 100,000 in amplitude.

It is quite remarkable that musical sounds remain recognizable over most of this range. This a fundamental aspect of hearing that all auditory theories must address -- how auditory percepts remain largely invariant over this huge range (perceptual constancy).

Hearing has a huge dynamic range!

Hearing has a huge dynamic range!

The **dynamic range of human hearing** is the ratio of the sound pressure level of the softest sound that can be heard to the loudest one that can be tolerated without pain.

This dynamic range is **> 100,000** (> 100 dB or 10^5 fold), and is roughly comparable to the 65,536 amplitude steps that are afforded by 16-bit digitization.

CD quality sound
16 bits = 2^{16} = 64k voltage levels, sampling rate @ 44,100 samples/sec

Typical sound levels in music

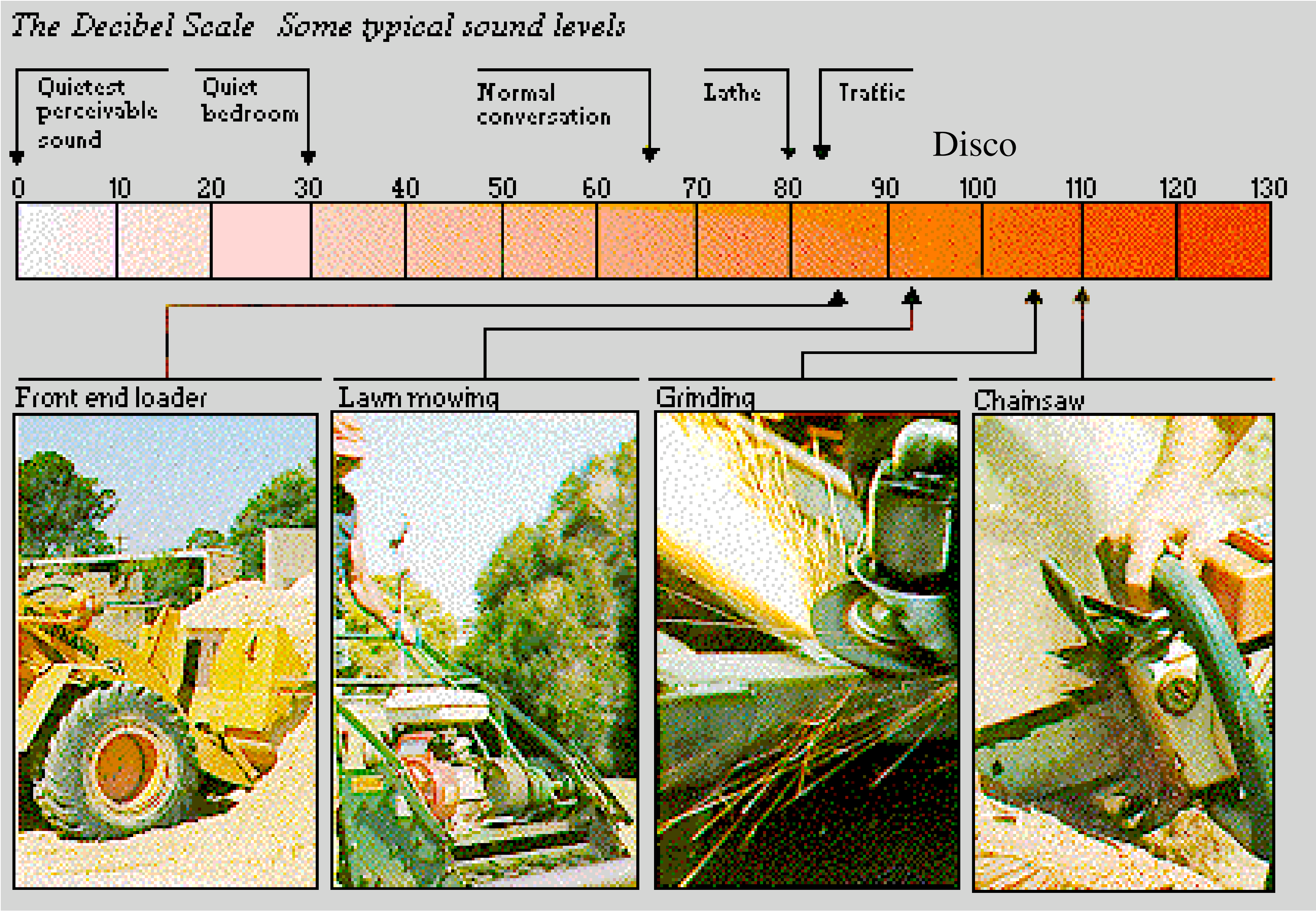
On origins of music dynamics notation
http://www.wikipedia.org/wiki/Pianissimo

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



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

Demonstrations

- **Demonstrations using waveform generator**
- Relative invariance of pitch & timbre with level
- Loudness matching
- Pure tone frequency limits
- Localization

Loudness

Dimension of perception that changes with sound intensity (level)

Intensity ~ power;

Level~amplitude

Demonstration using waveform generator

Masking demonstrations

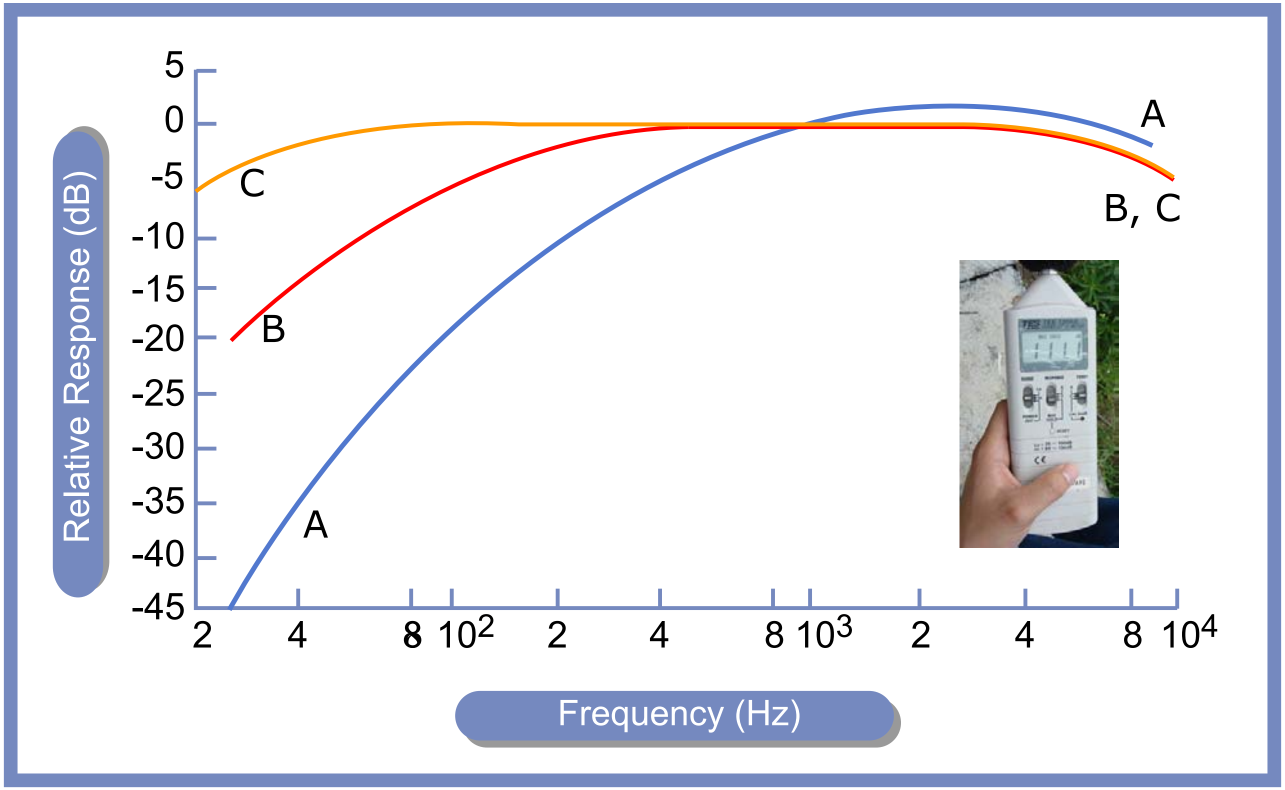
Magnitude estimation

Loudness matching

Sound level meters and frequency weightings

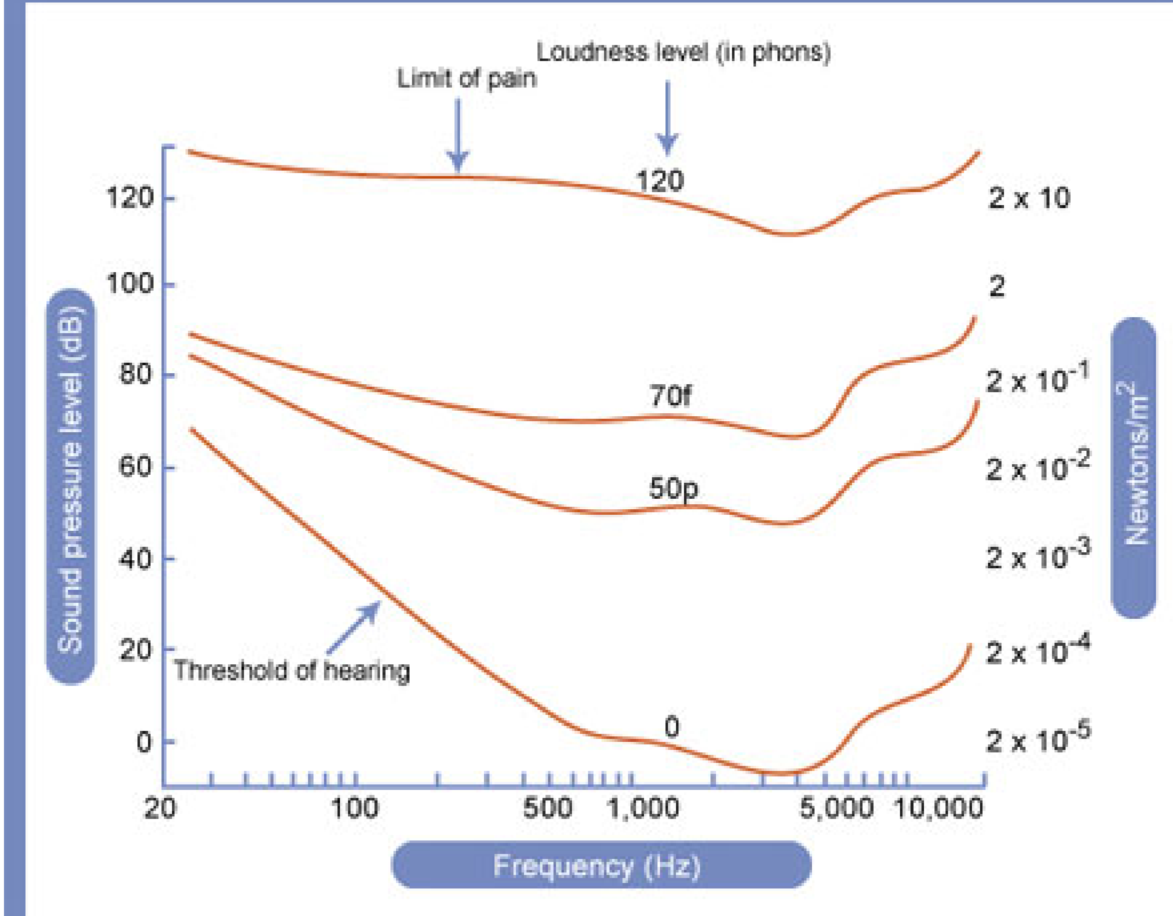
A-weighting: perceived loudness

C-weighting: flat



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

Loudness as a function of pure tone level & frequency

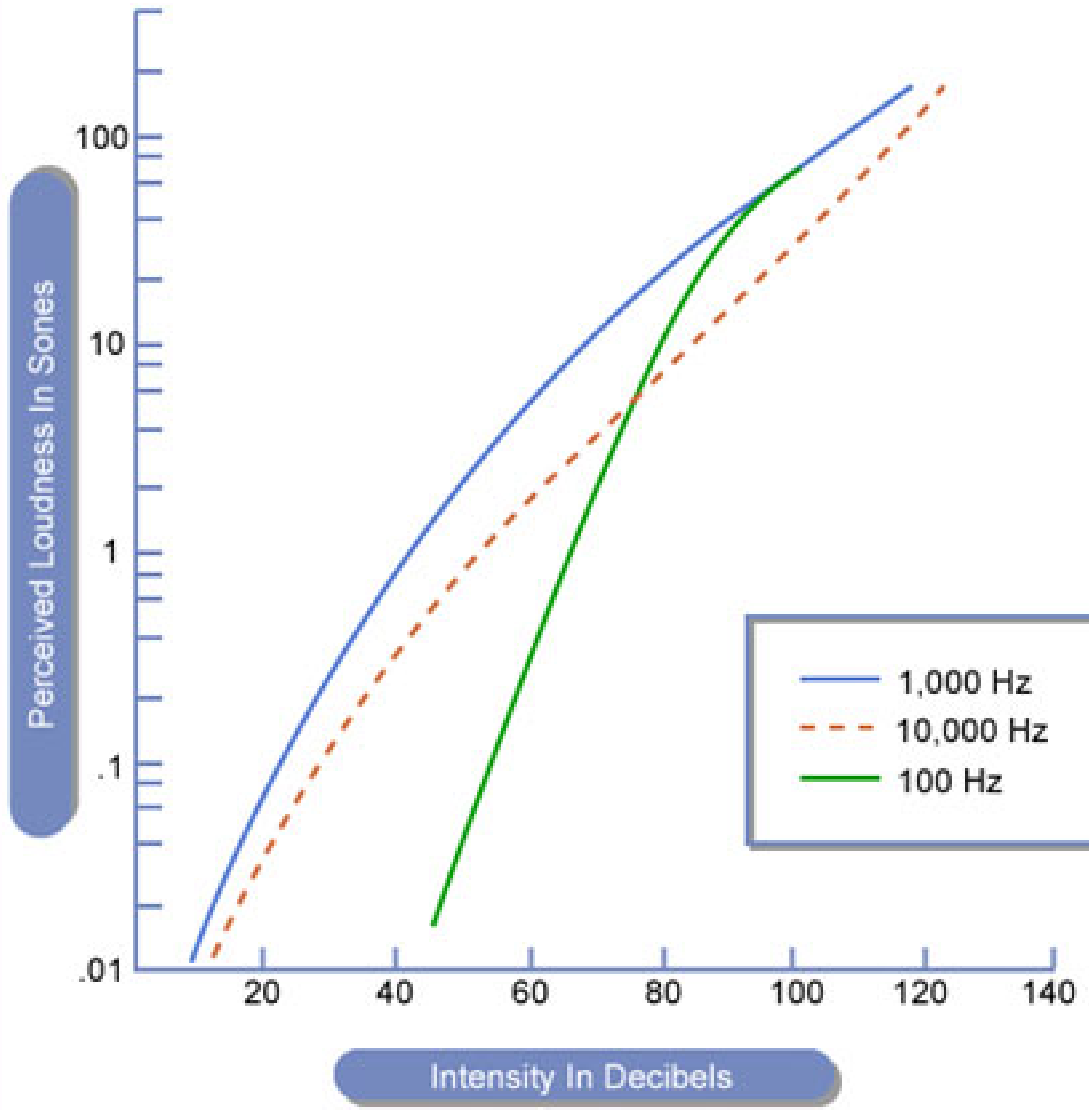


Constant-loudness curves for persons with acute hearing. All sinusoidal sounds whose levels lie on a single curve (an *isophon*) are equally loud. A particular loudness-level curve is designated as a loudness level of number of *phons*. The number of phons is equal to the number of decibels only at the frequency 1,000 Hz.

Figure by MIT OpenCourseWare.

Absolute detection thresholds on the order of
1 part in a million,
Δ pressure ~1/1,000,000 atm
(Troland, 1929)

Loudness perception:



Perceived loudness of tones of various frequencies as a function of physical intensity.

Loudness perception: population percentiles

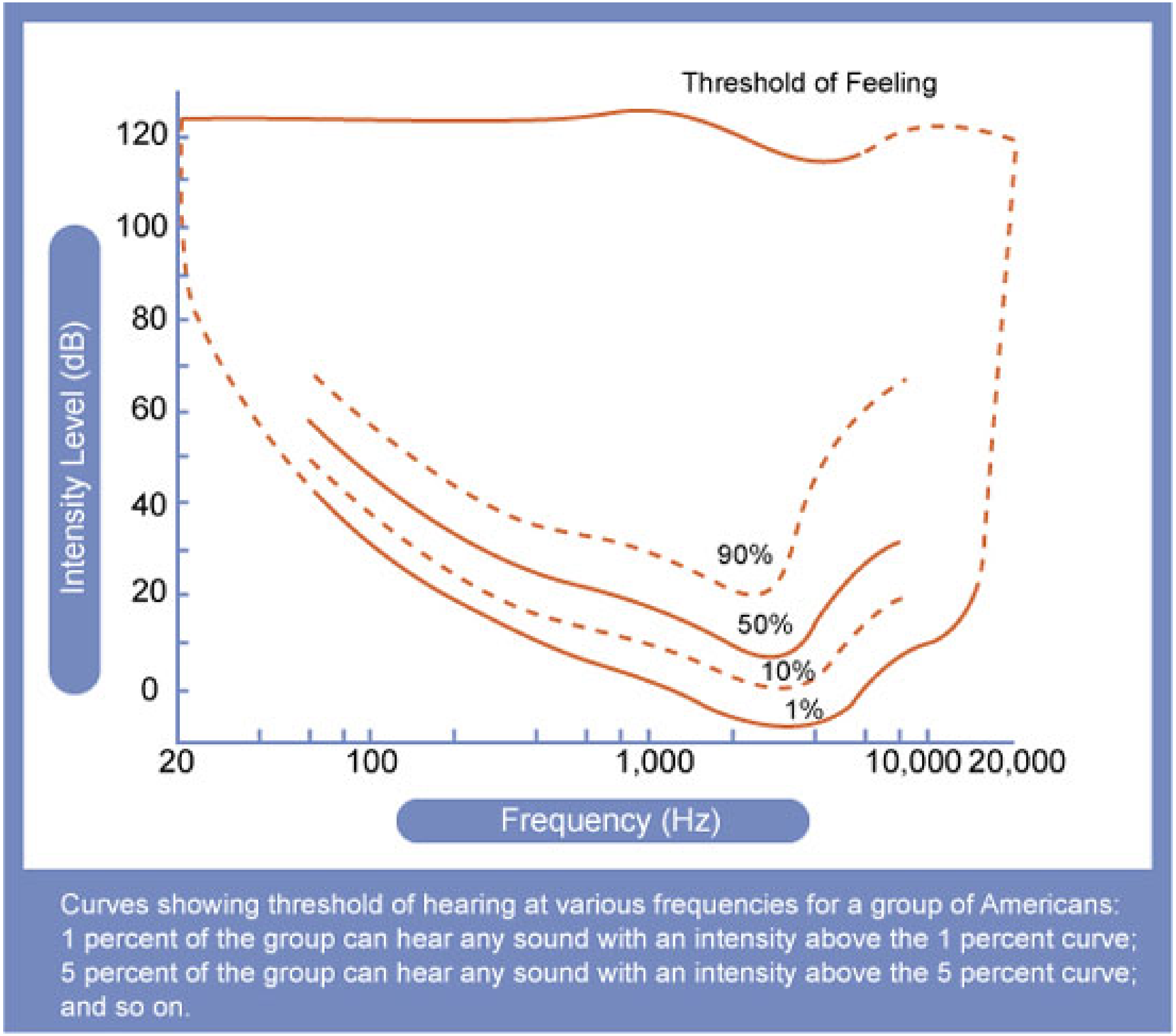
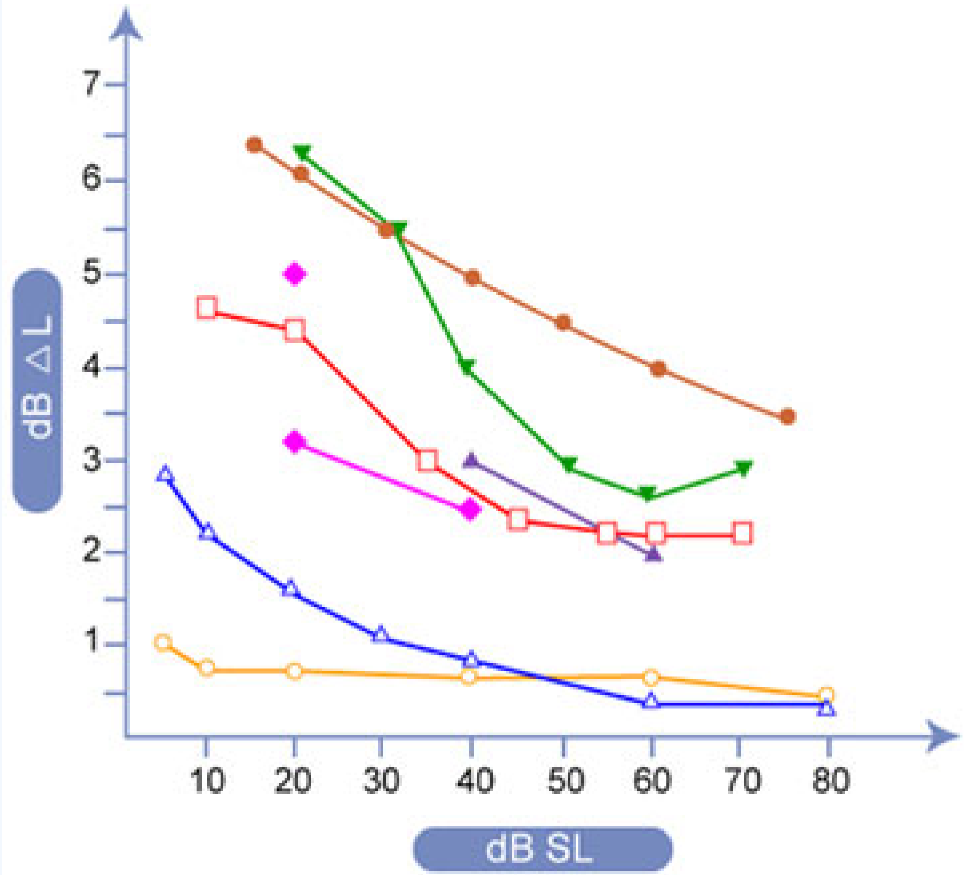


Figure by MIT OpenCourseWare.

Intensity discrimination improves at higher sound levels

Best Weber fraction

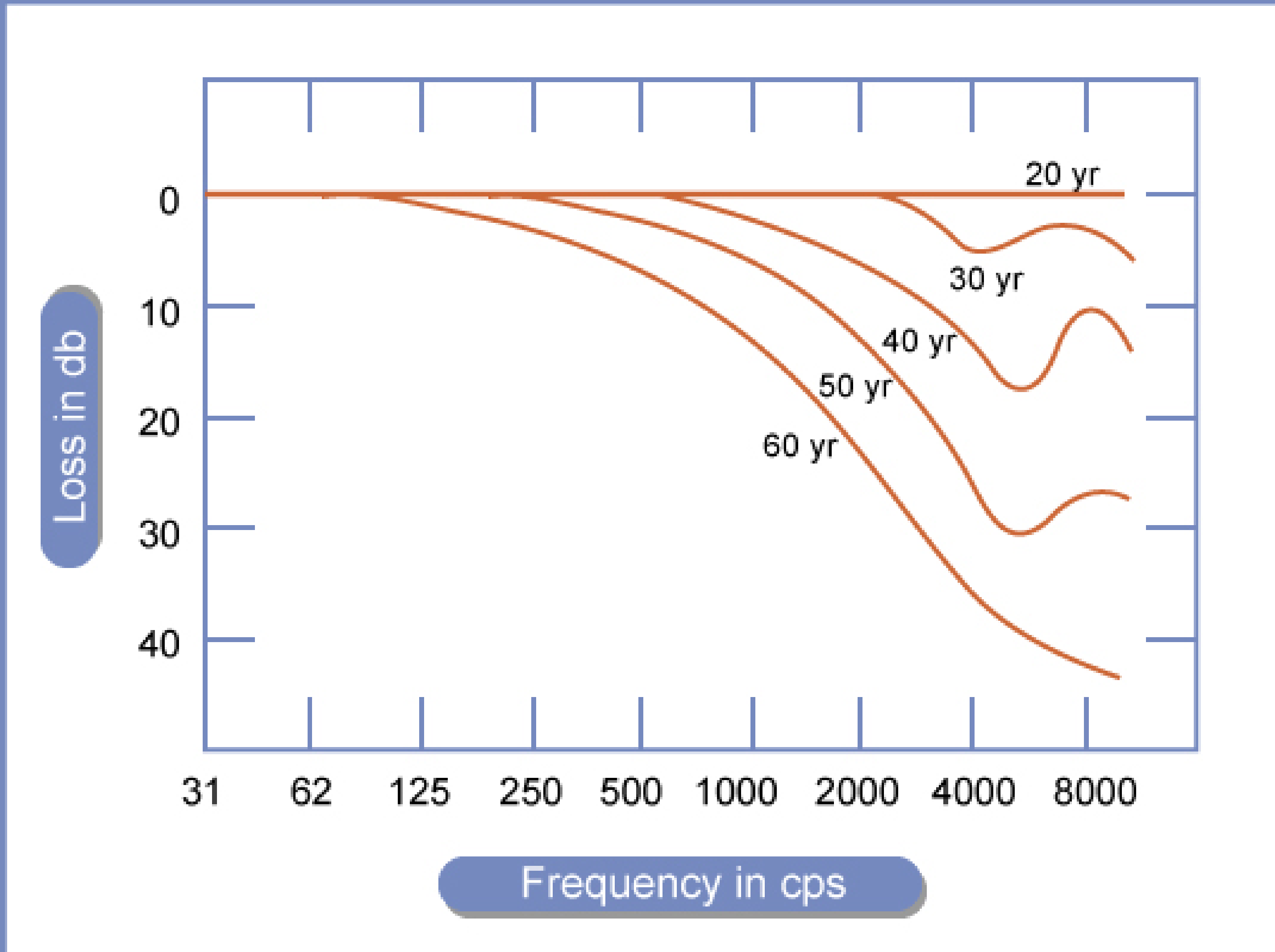


A comparison of just noticeable intensity differences (averaged across frequencies) for various species. Man (open symbols): red (Dimmick & Olson, 1941), orange (experiment I), blue (experiment III, Harris, 1963); cat: purple (Raab & Ades,1946; Elliott & McGee, 1965); rat: pink (Henry, 1938; Hack, 1971); mouse: brown (Ehret, 1975b); parakeet: green (Dooling & Saunders, 1975b).

Figure adapted from cited sources above.

Figure by MIT OpenCourseWare.

Hearing loss with age



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.

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.

Periodicity and spectrum

Periodicity vs. frequency

Longstanding and ongoing dichotomy between formally-equivalent, yet complementary perspectives (de Cheveigne chapter on pitch)

Vibrating strings vs. Helmholtz resonators

Comb filters vs. band-limited filters

Autocorrelation vs. Fourier analysis

and yet another paradigm

Complex oscillator (delay loop)

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.

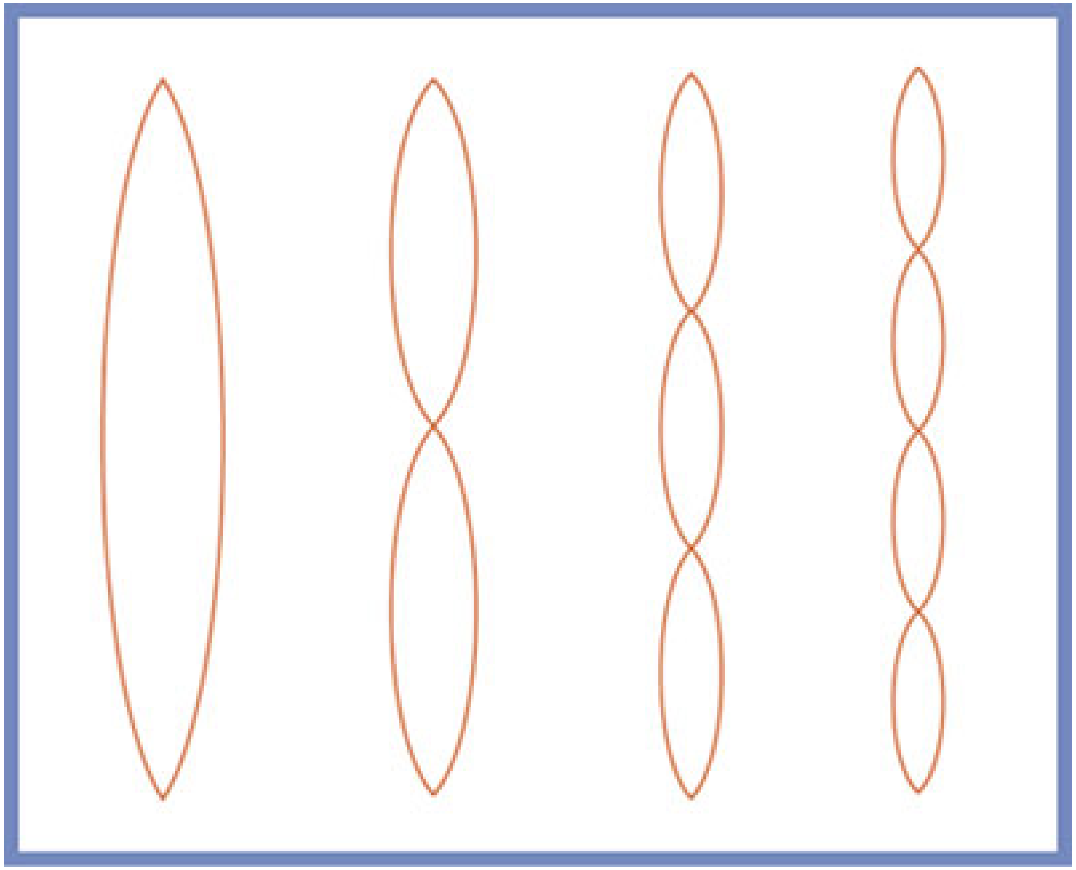


Figure by MIT OpenCourseWare.

Frequency spectra

The Greeks understood simple relationships between vibration rate & pitch.

Experiments with musical instruments and tuning systems were carried out by many people (Galileo's father, Galileo, Saveur, Mersenne, others).

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

George Ohm (1789-1854) postulated that sounds can be decomposed into component sinusoids

Hermann von Helmholtz (1821-1894) postulated that the ear analyzes sound by first breaking sounds into their partials and then doing associative pattern-recognition

Debate between Seebeck, Ohm, & Helmholtz (1844) over periodicity vs. spectral pattern

Foreshadows temporal vs. place codes, autocorrelation vs. Fourier spectrum

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 1 frequency component is called a *pure tone*.

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

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.

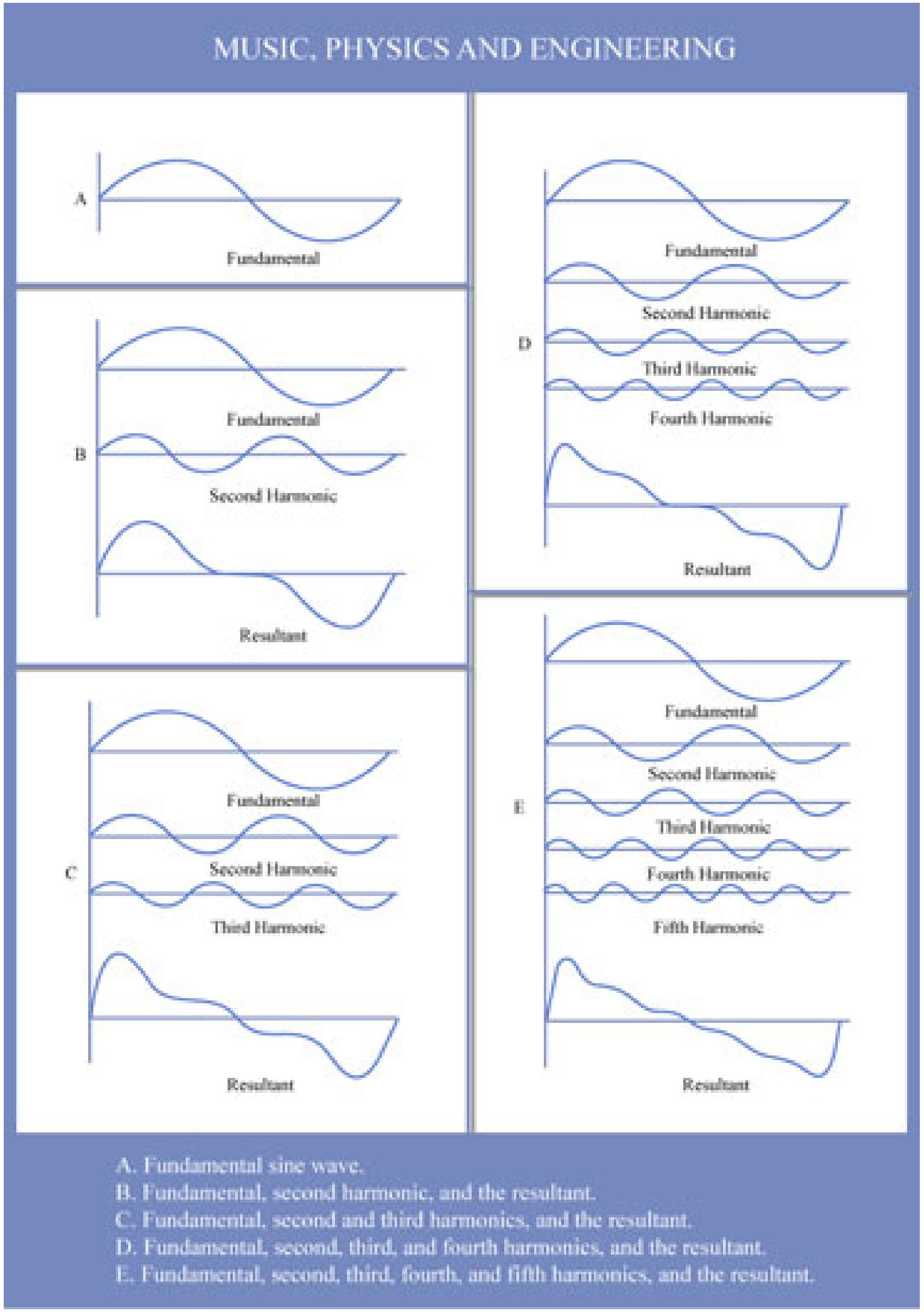


Figure by MIT OpenCourseWare.

Harmonic series

A harmonic series consists 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

Sound quality contrasts

Duration

- Impulsive sounds
- Sustained sounds
  - Stationary vs. nonstationary

Pattern complexity, coherence

- Pitched sounds
  - Time domain: Periodic sound patterns
  - Frequency domain: harmonics
- Inharmonic sounds
  - Combinations of unrelated periodic patterns
  - Complexity: Number of independent patterns
- Noises
  - Aperiodic sound patterns, high complexity

Minimal durations

Graph removed due to copyright restrictions.
Figure 36, comparing "Tone pitch" and "click pitch" responses.
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.

Licklider (1951)
“Basic correlates
of the auditory stimulus”

Periodic vs. aperiodic sounds

- Periodic sound patterns -- “tones”
- Aperiodic sound patterns -- “noise”

Range of pitches of pure & complex tones

- **Pure tone pitches**
  - Range of hearing (~20-20,000 Hz)
  - Range in tonal music (100-4000 Hz)
- **Most (tonal) musical instruments produce harmonic complexes that evoke pitches at their fundamental frequencies (F0's)**
  - Range of F0's in tonal music (30-4000 Hz)
  - Range of missing fundamental (30-1200 Hz)

JND's

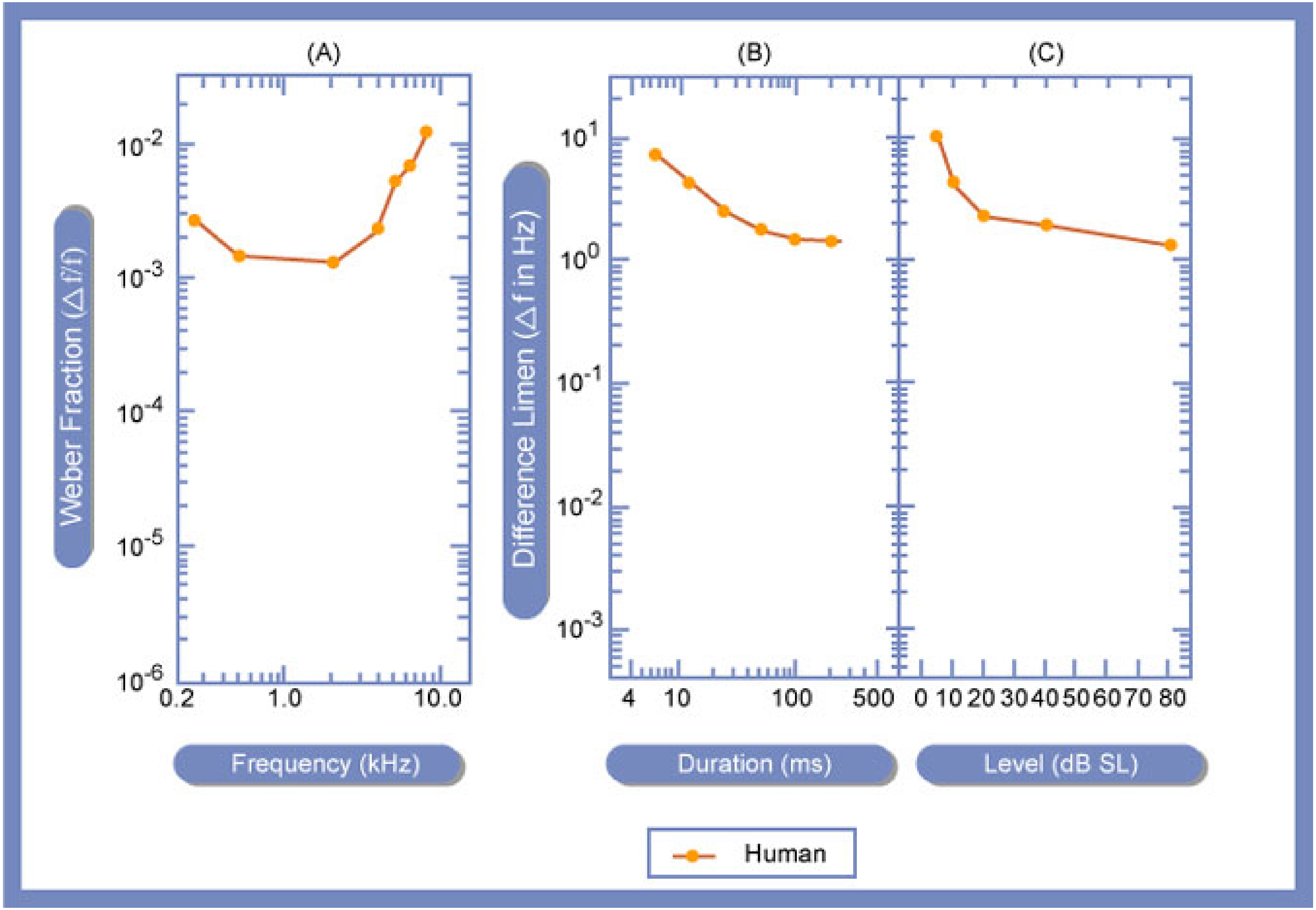


Figure by MIT OpenCourseWare.

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

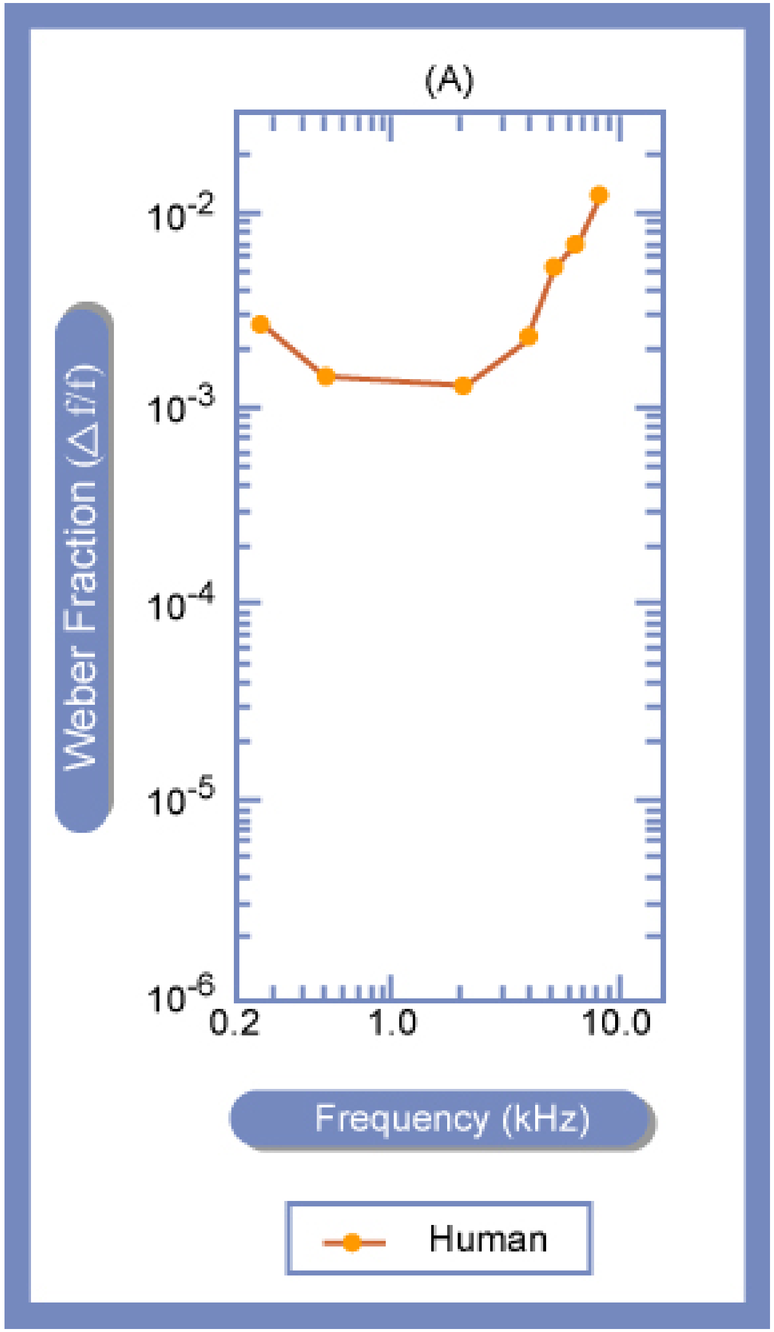


Figure by MIT OpenCourseWare.

Pure tone pitch discrimination improves

at longer tone durations

and

at higher sound pressure levels

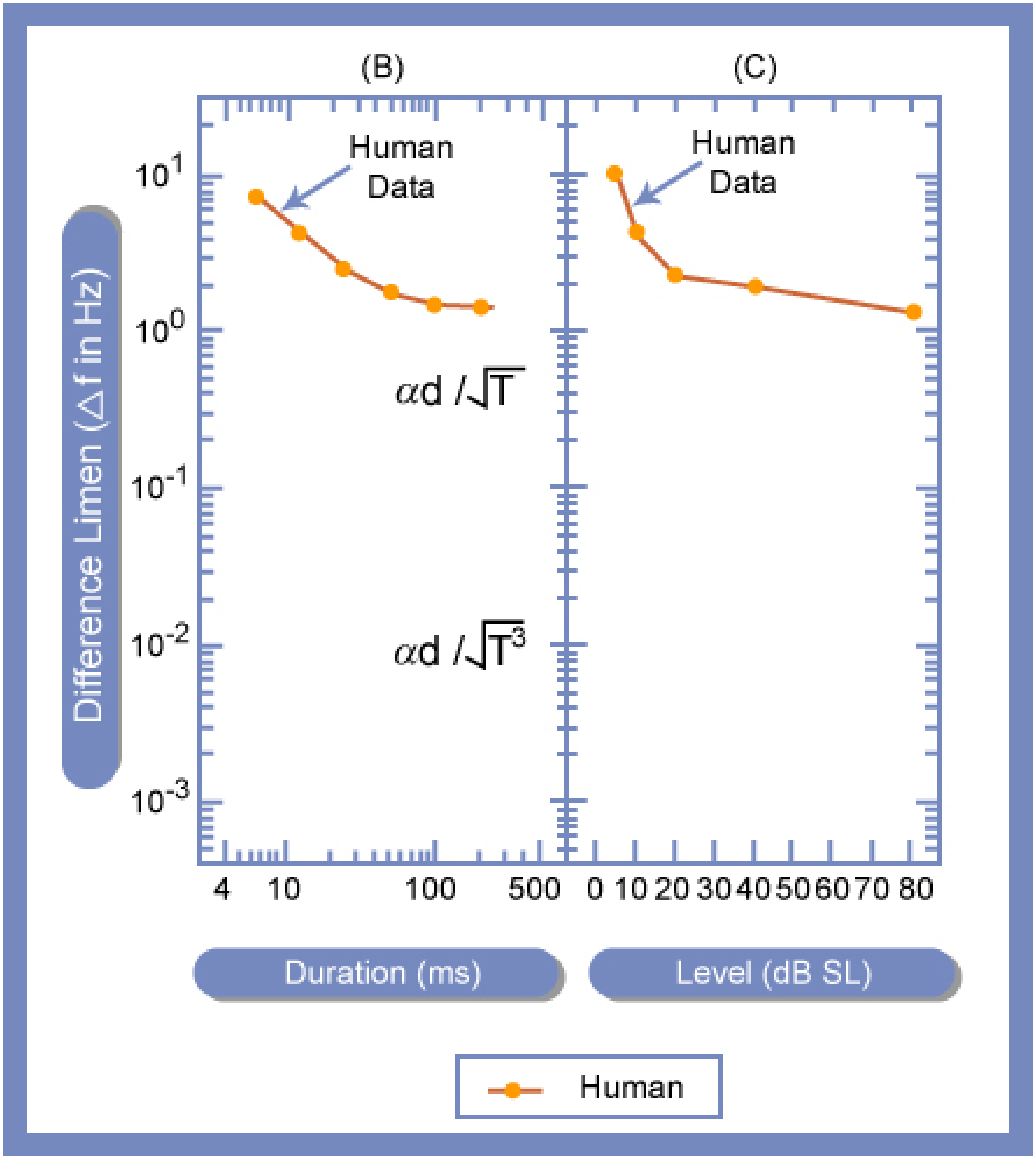
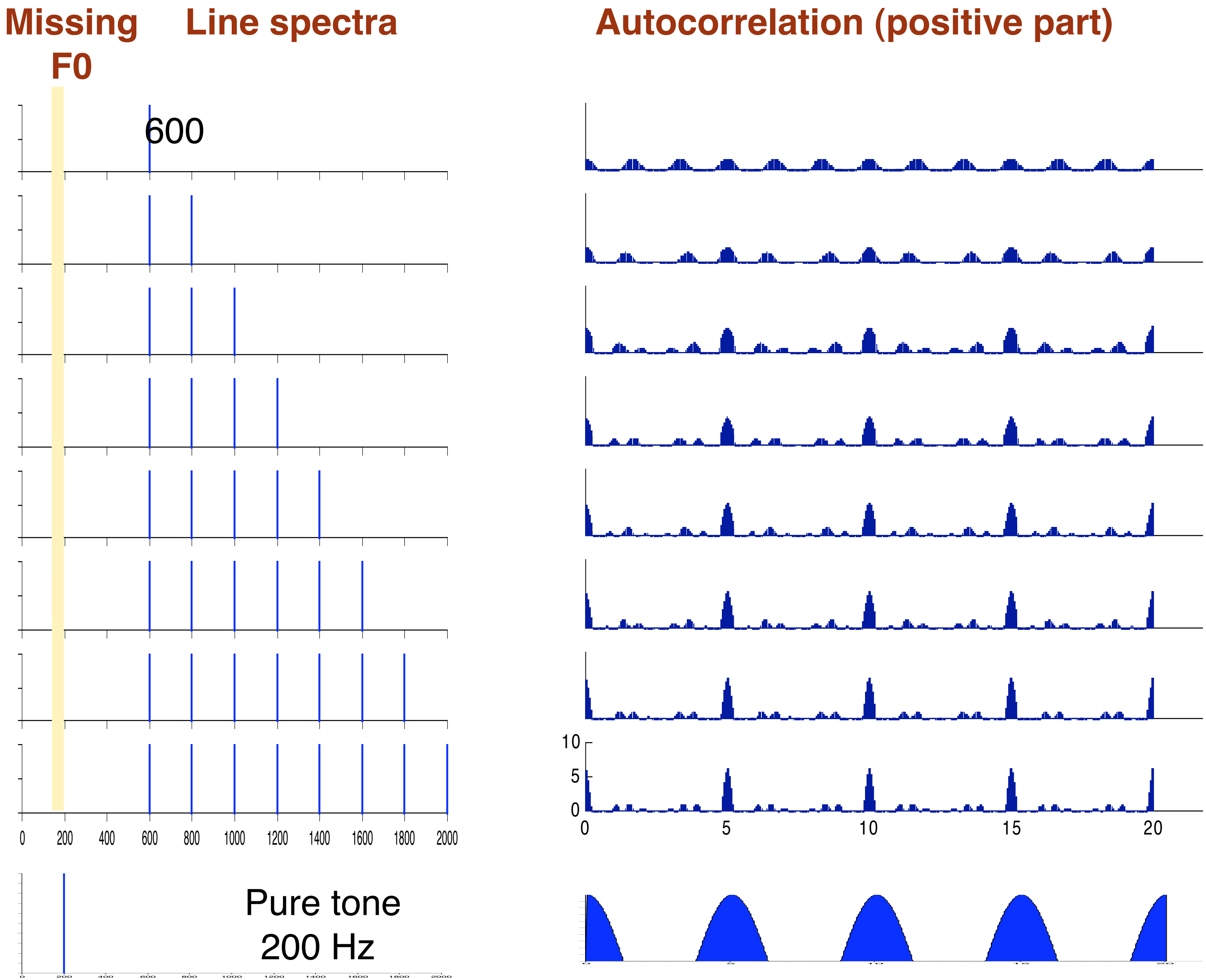
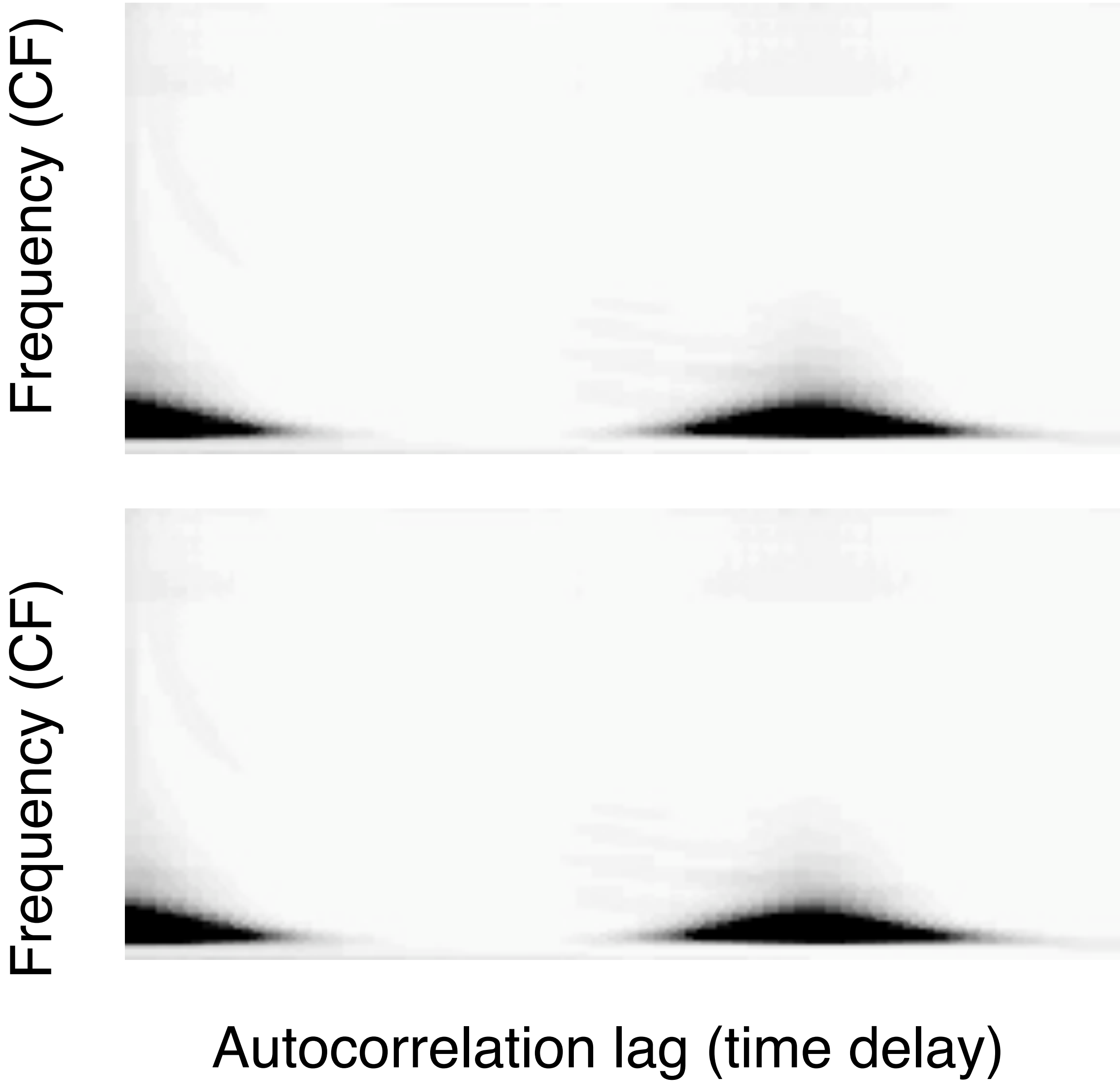


Figure by MIT OpenCourseWare.

Emergent pitch



Correlograms: interval-place displays (Slaney & Lyon)



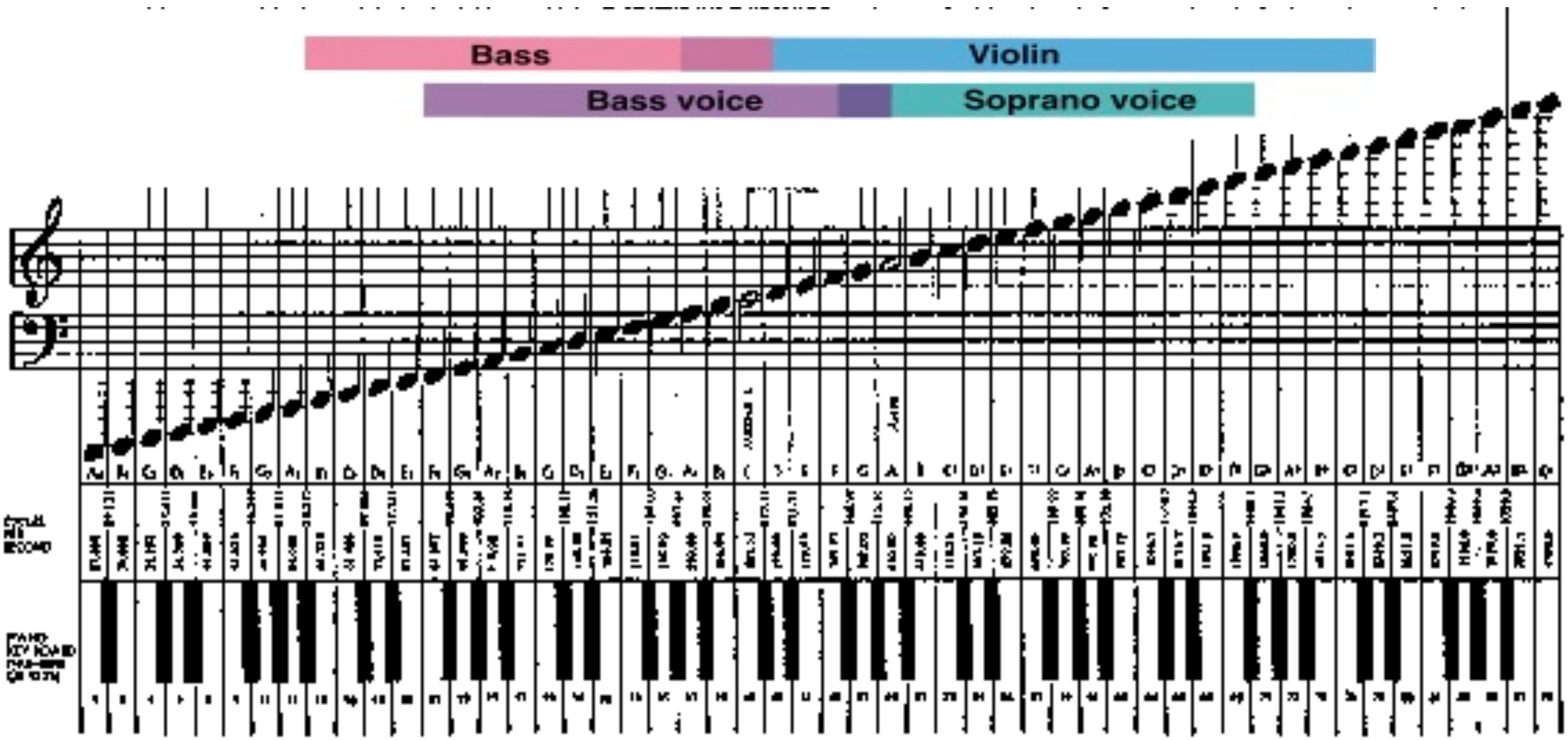
Courtesy of Malcolm Slaney (Research Staff Member of IBM Corporation). Used with permission.

Frequency ranges of (tonal) musical instruments

> 6 kHz

2.5-4 kHz

8k
6
5
4
3
2



27 Hz

110 Hz

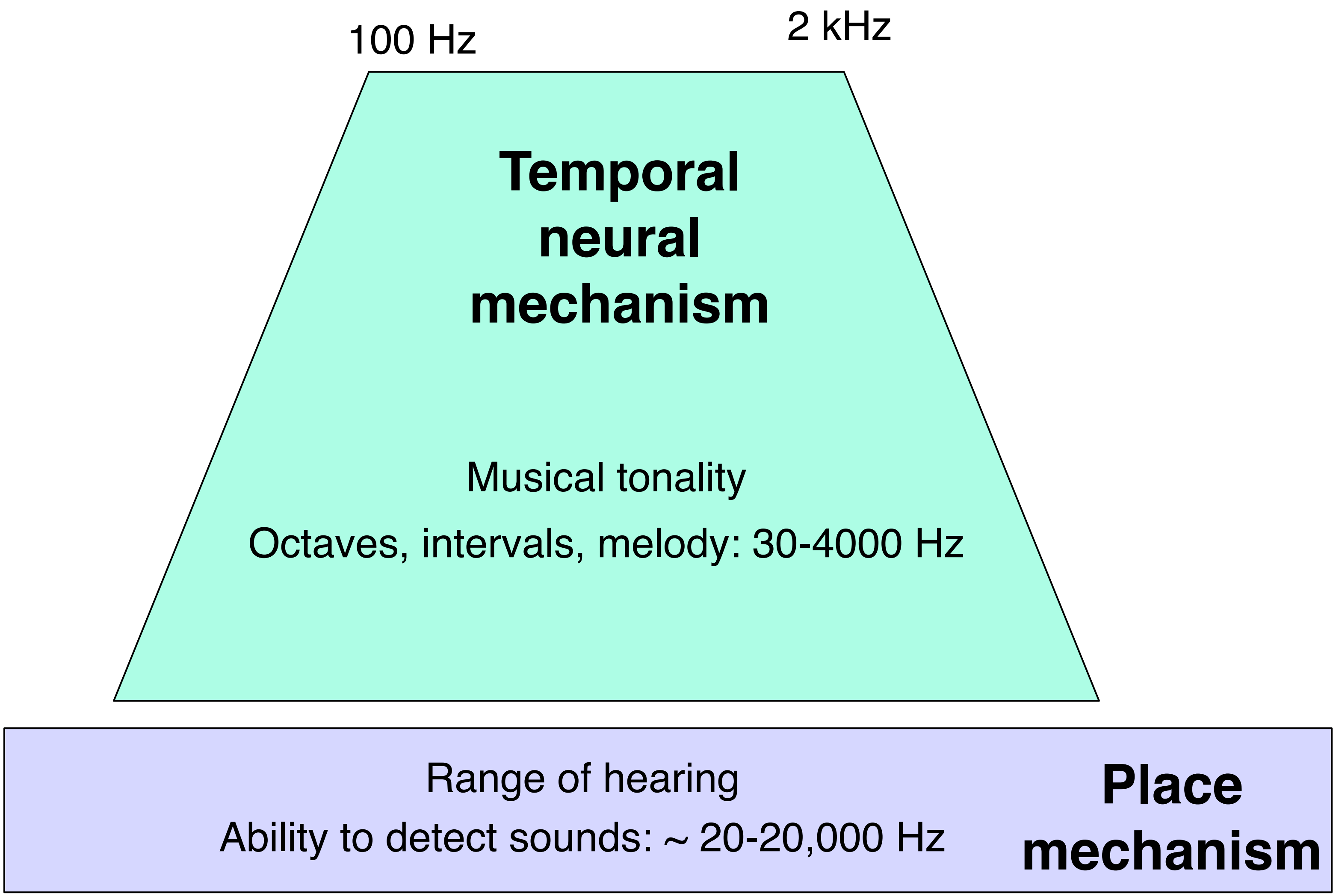
262 Hz

440 Hz

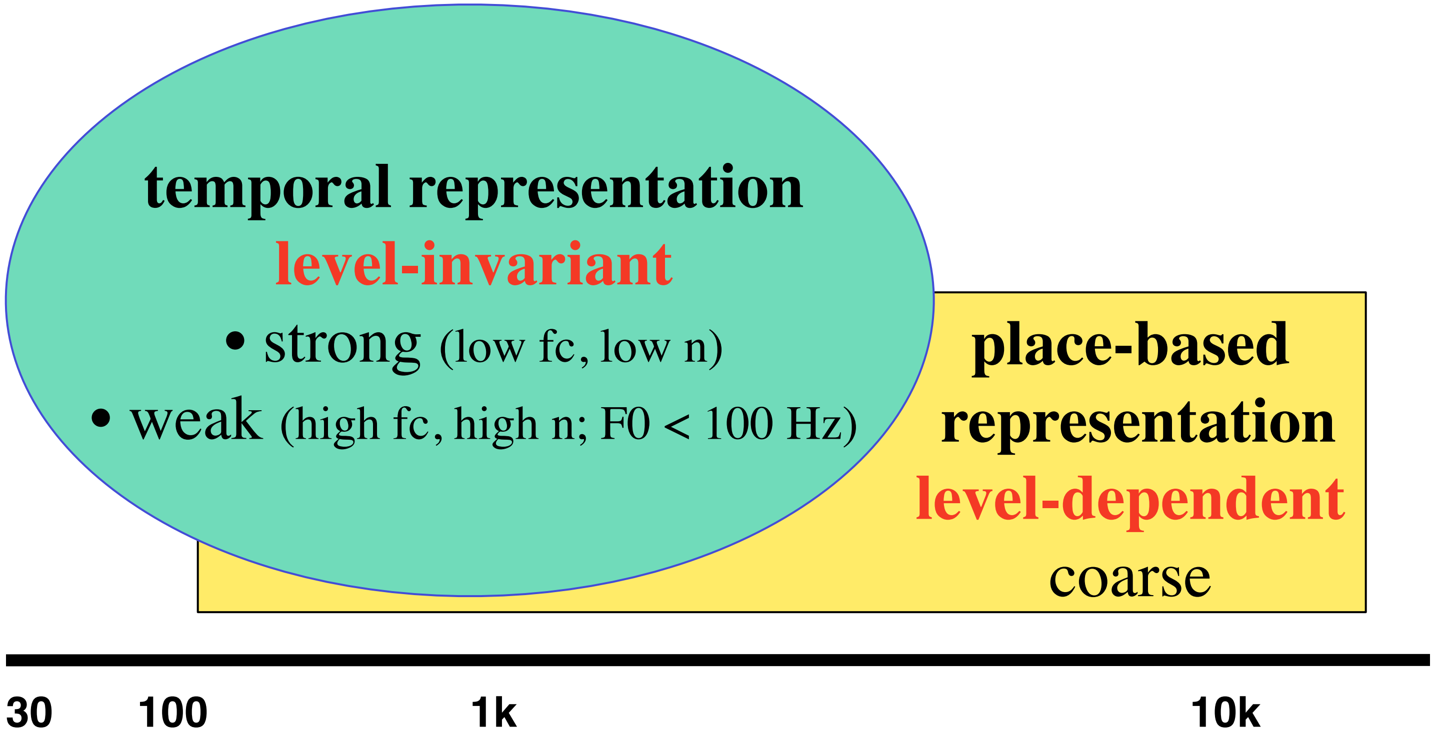
880 Hz

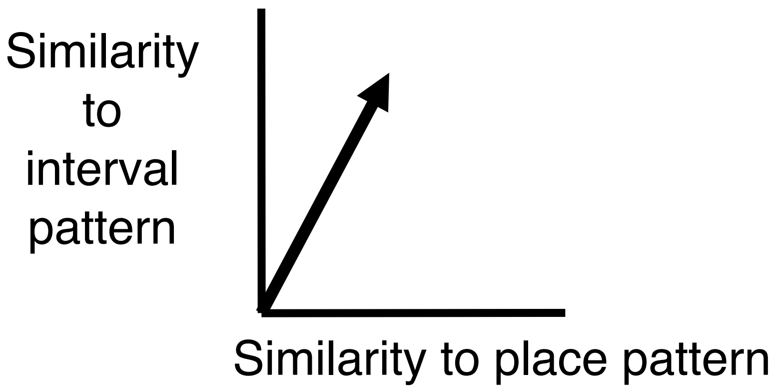
4 kHz

Frequency ranges: hearing vs. musical tonality



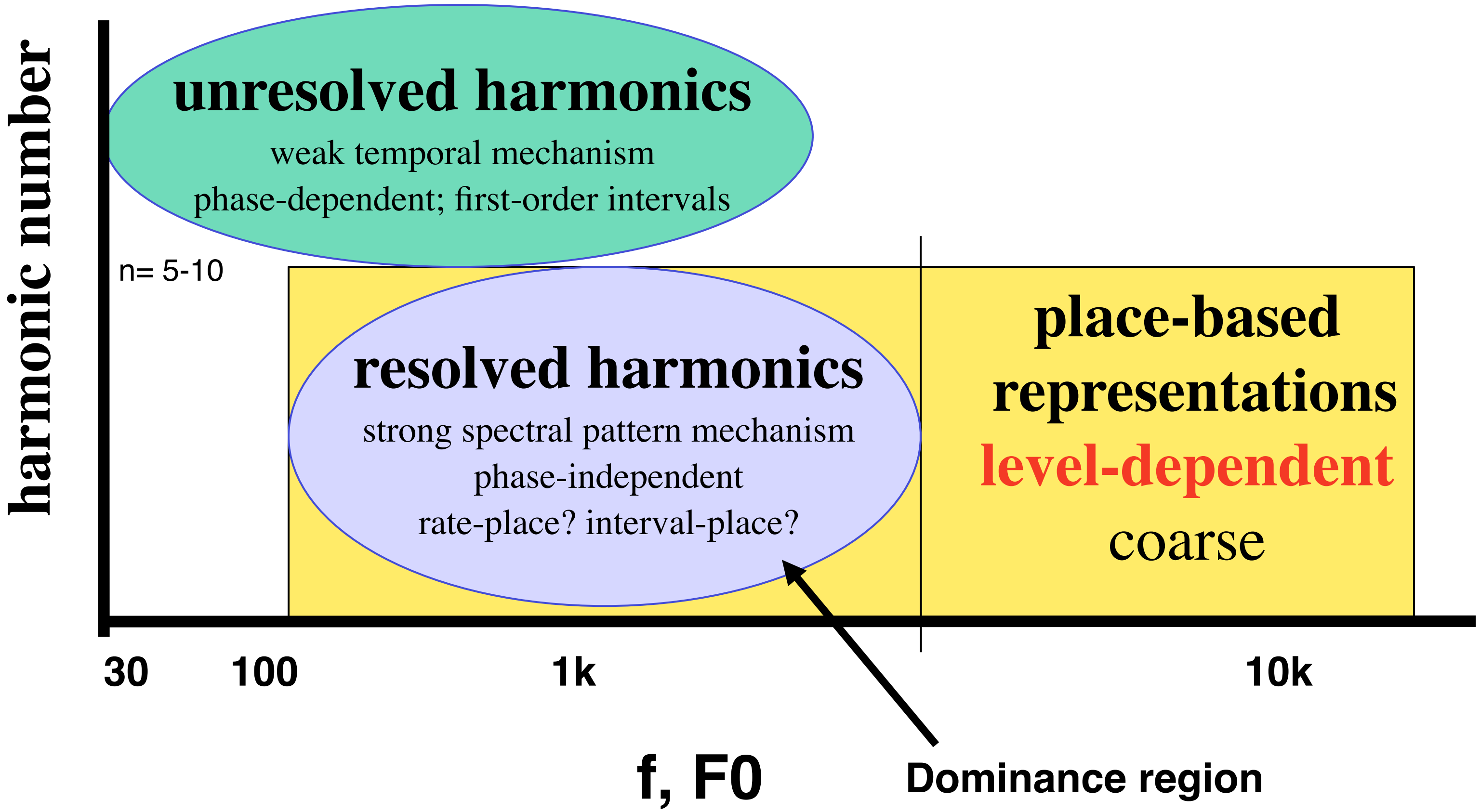
Duplex time-place representations



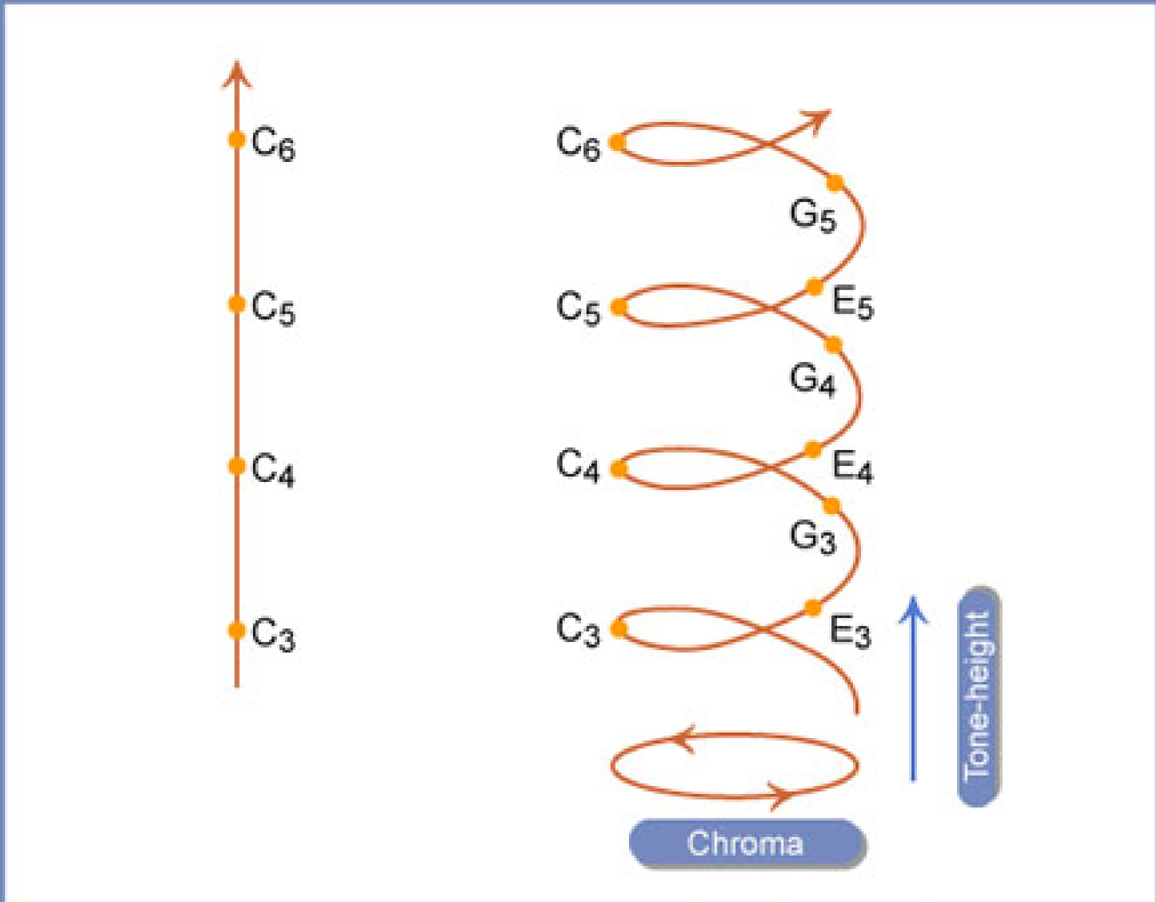


<u>cf. Terhardt's</u>
<u>spectral and virtual pitch</u>

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



Pitch dimensions: height & chroma



Contrast between one-dimensional and two-dimensional models of pitch perception. Notes of a scale played on an ordinary instrument sprial upward around the surface of a cylinder, but computer-generated notes can form a Shepard scale that goes around in circle.

Figure by MIT OpenCourseWare.

Pitch height and pitch chroma

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.

Two codes for “pitch

Place code

Pitch = place of excitation

Pitch height

Absolute

Low vs. high

Existence region
f: .100-20,000 Hz

Time code

Pitch = dominant periodicity

Musical pitch
"Chroma", Tonality

Relational

Musical intervals, tonality
Melodic recognition
& transposition

Existence region
F0: 30-4000 Hz

Note durations in music

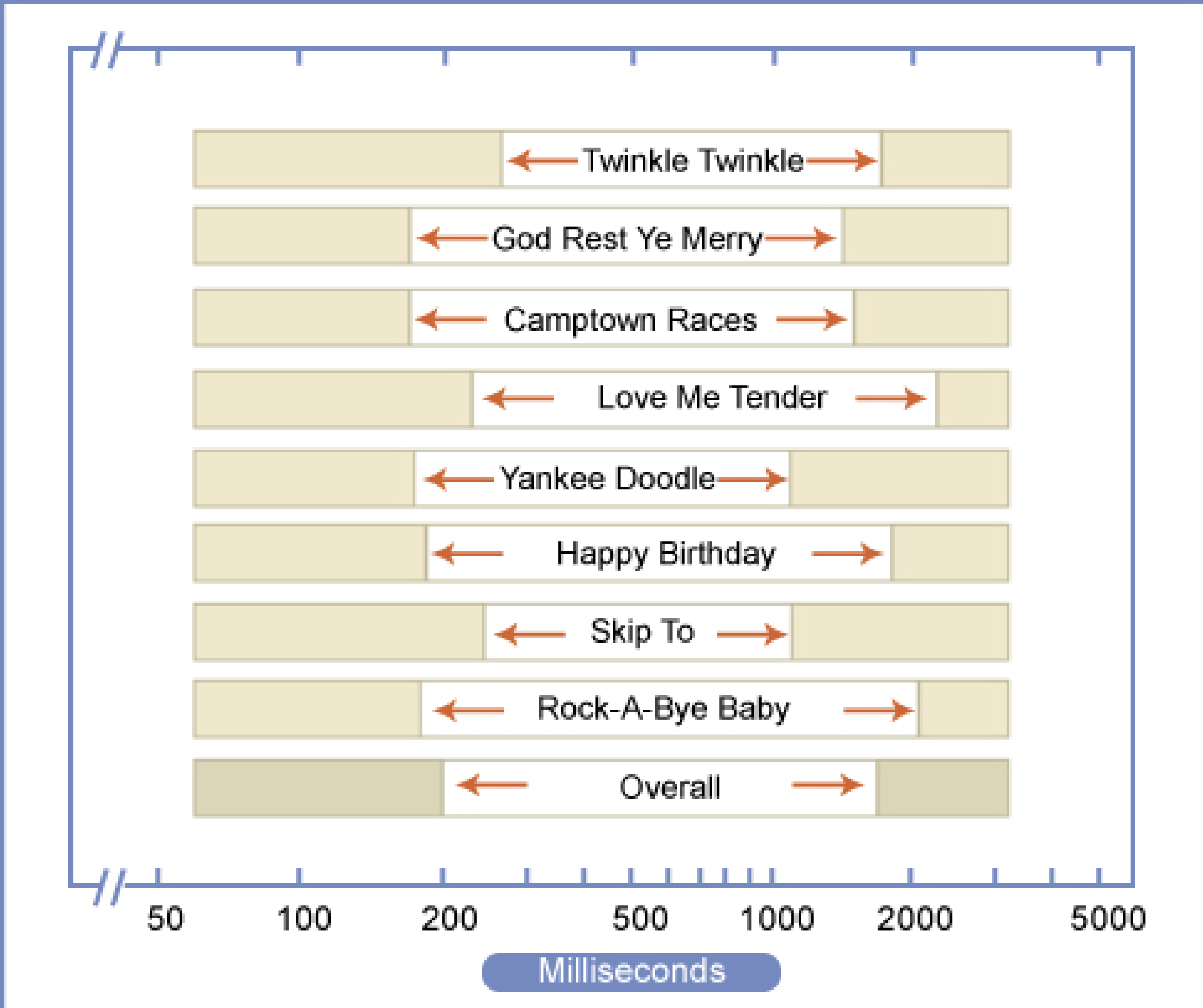


Image adapted from: McAdams, and Bigand. *Thinking in Sound: The Cognitive Psychology of Human Audition.* Oxford University Press, 1993.

Figure by MIT OpenCourseWare.

Timbre: a multidimensional tonal quality

tone texture, tone color distinguishes voices, instruments

(Photo Courtesy of Pam Roth. Used with permission.)

Stationary Aspects

(spectrum)

Dynamic Aspects

Δ spectrum
Δ intensity
Δ pitch
attack
decay

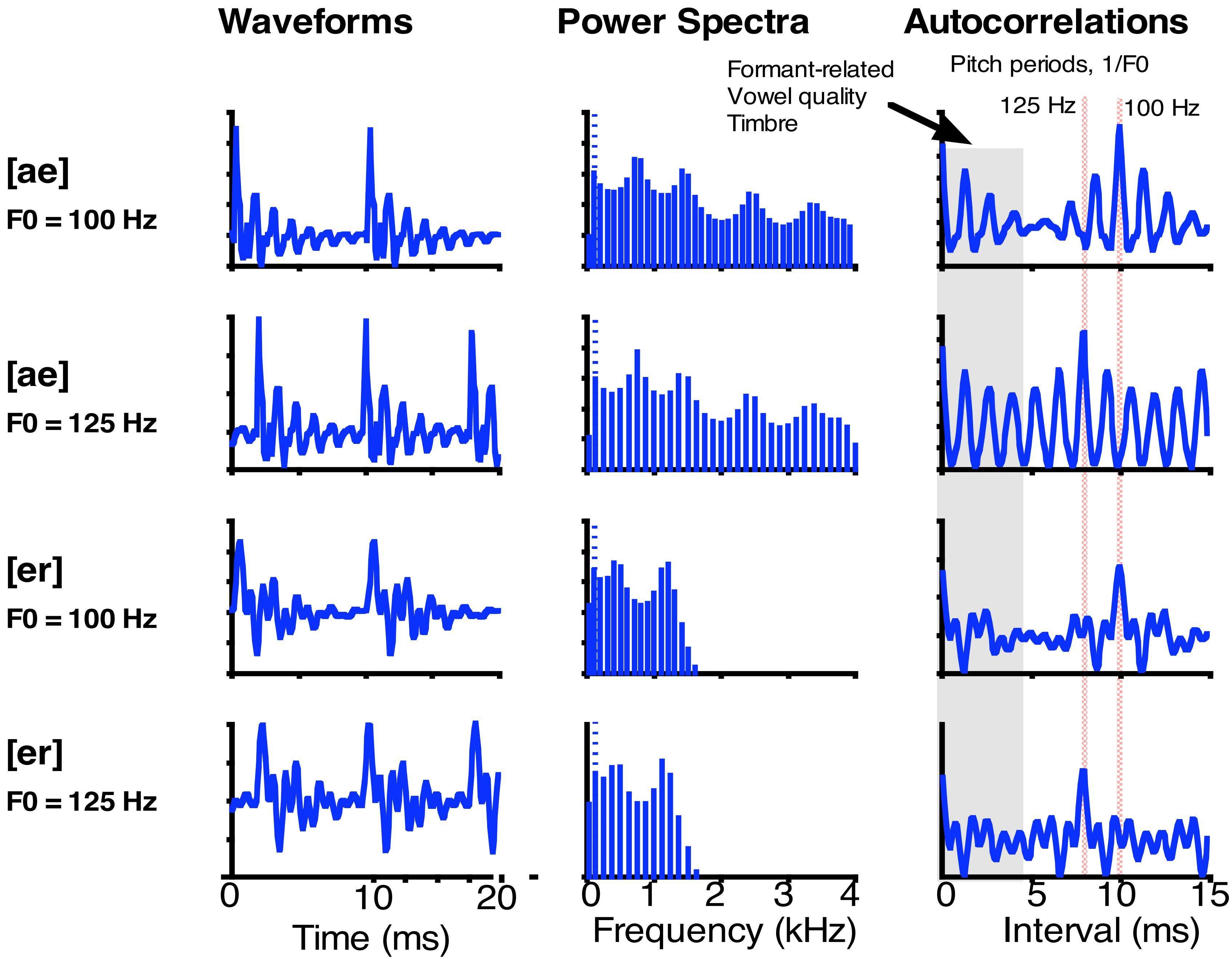
Photo Courtesy of Per-Ake Bystrom.Used with permission.)

Vowels

Photo Courtesy of Miriam Lewis.Used with permission.)

Consonants

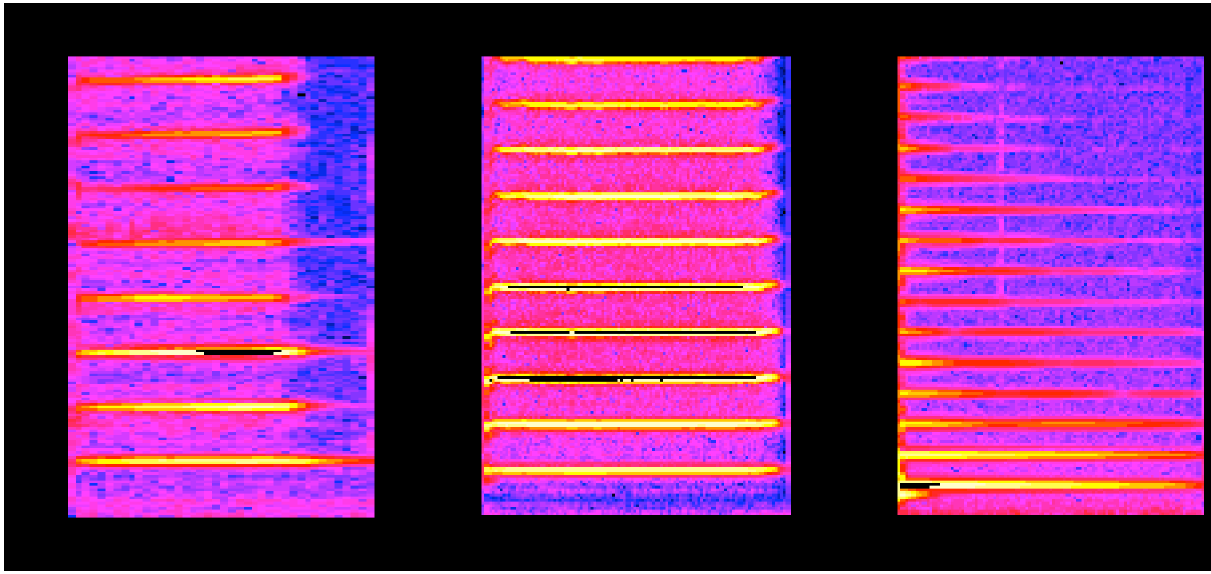
Stationary spectral aspects of timbre



Rafael A. Irizarry's Music and Statistics Demo

Spectrograms
of Harmonic Instruments

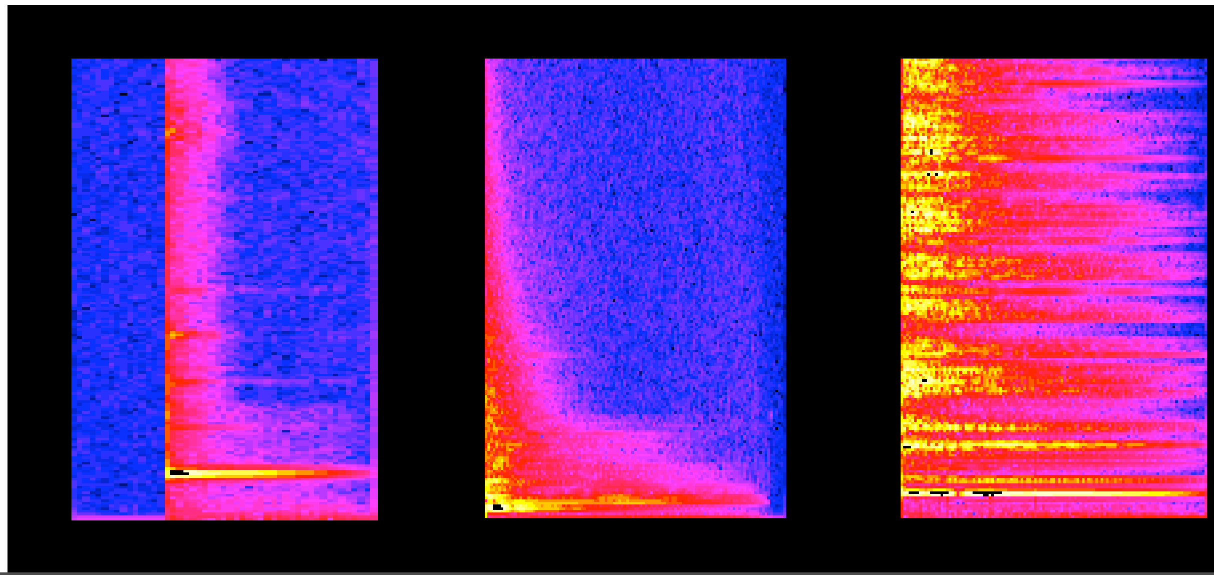
violin, trumpet, guitar



Courtesy of Rafael A. Irizarry. Used with permission.

Non-Harmonic Instruments

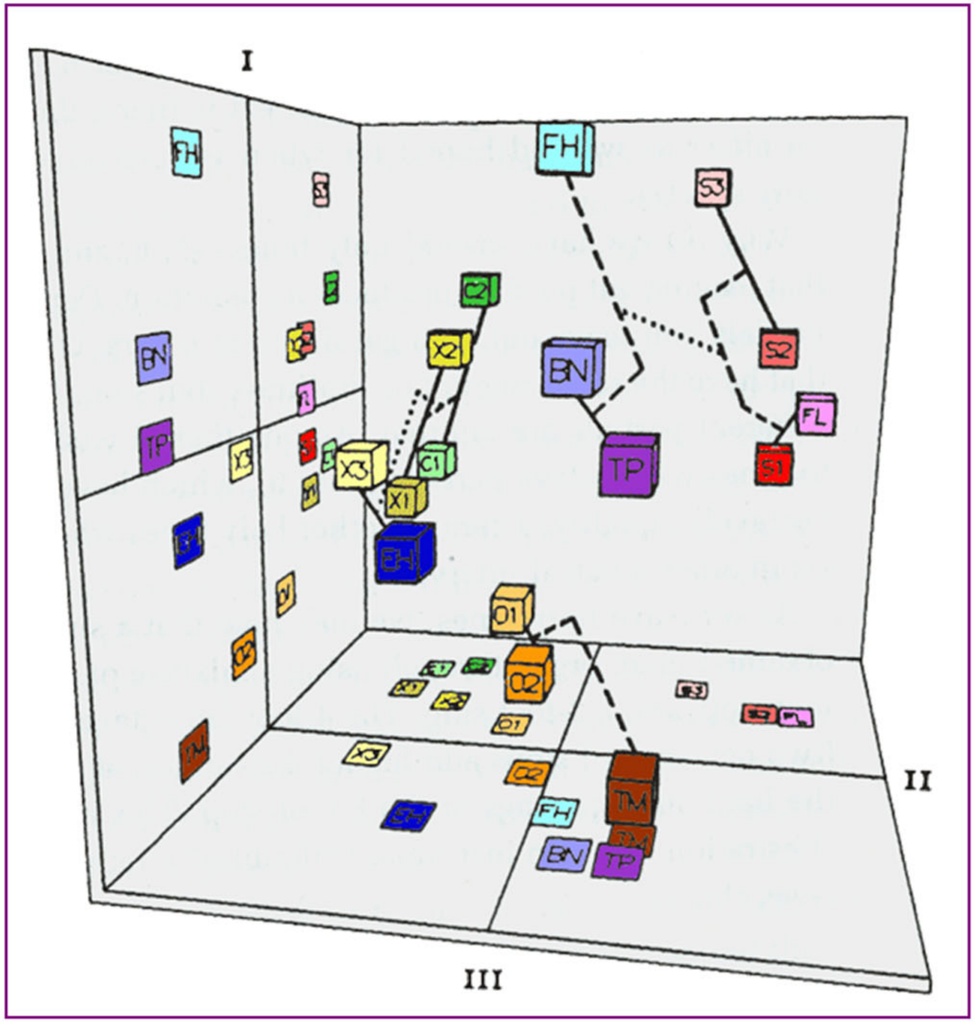
marimba, timpani, gong



Audio samples at this page:

http://www.biostat.jhsph.edu/~ririzarr/Demo/demo.html

Timbre dimensions: spectrum, attack, decay

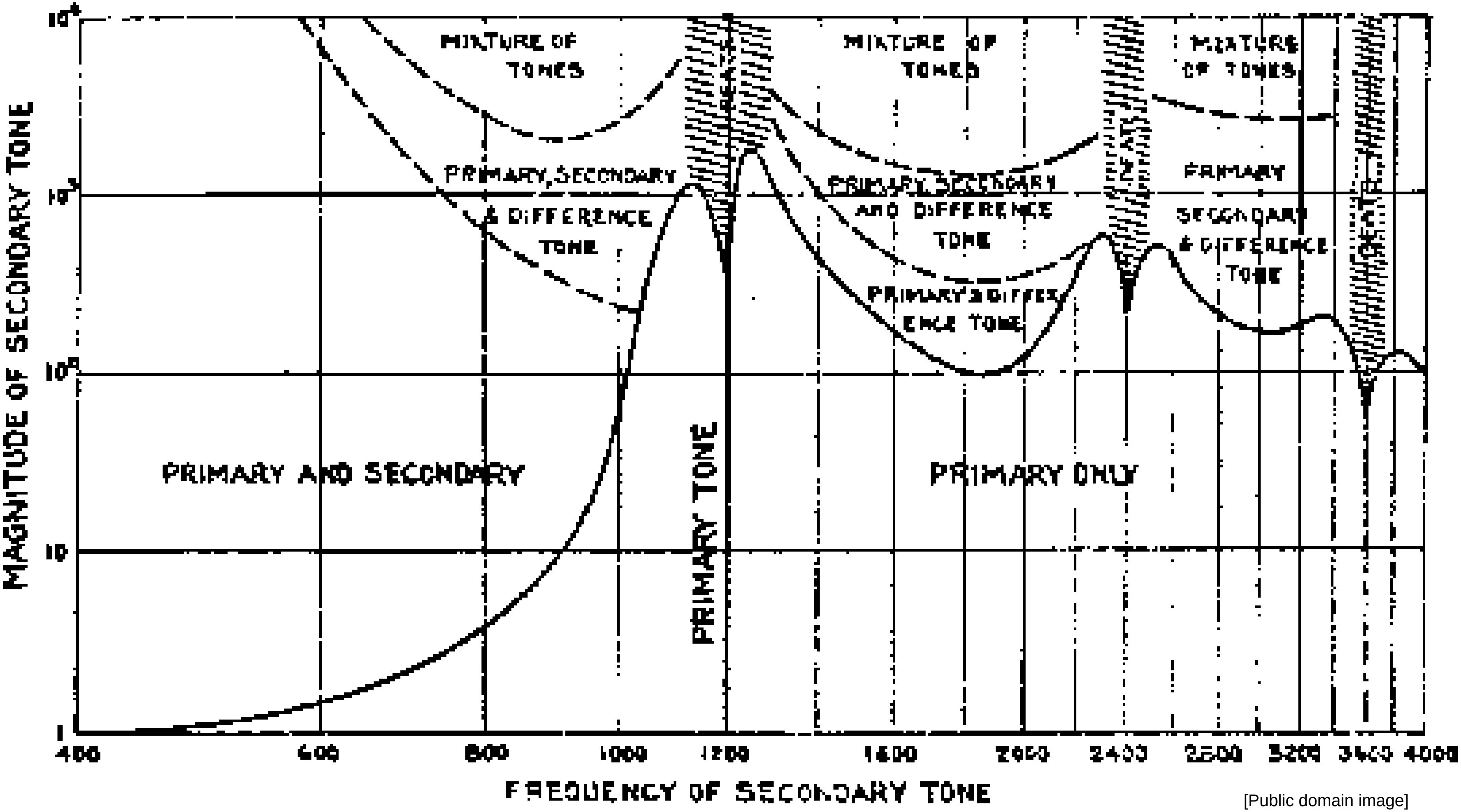


- BN - Bassoon
- C1 - E flat Clarinet
- C2 - B flat Bass Clarinet
- EH - English Horn
- FH - French Horn
- FL - Flute
- O1 - Oboe
- O2 - Oboe (different instrument and player)
- S1 - Cello, muted *sul ponticello*
- S2 - Cello
- S3 - Cello, muted *sul tasto*
- TM - Muted Trombone
- TP - B flat Trumpet
- X1 - Saxophone, played *mf*
- X2 - Saxophone, played *p*
- X3 - Soprano Saxophone

- Dimension I: spectral energy distribution, from broad to narrow
- Dimension II: timing of the attack and decay, synchronous to asynchronous
- Dimension III: amount of inharmonic sound in the attack, from high to none

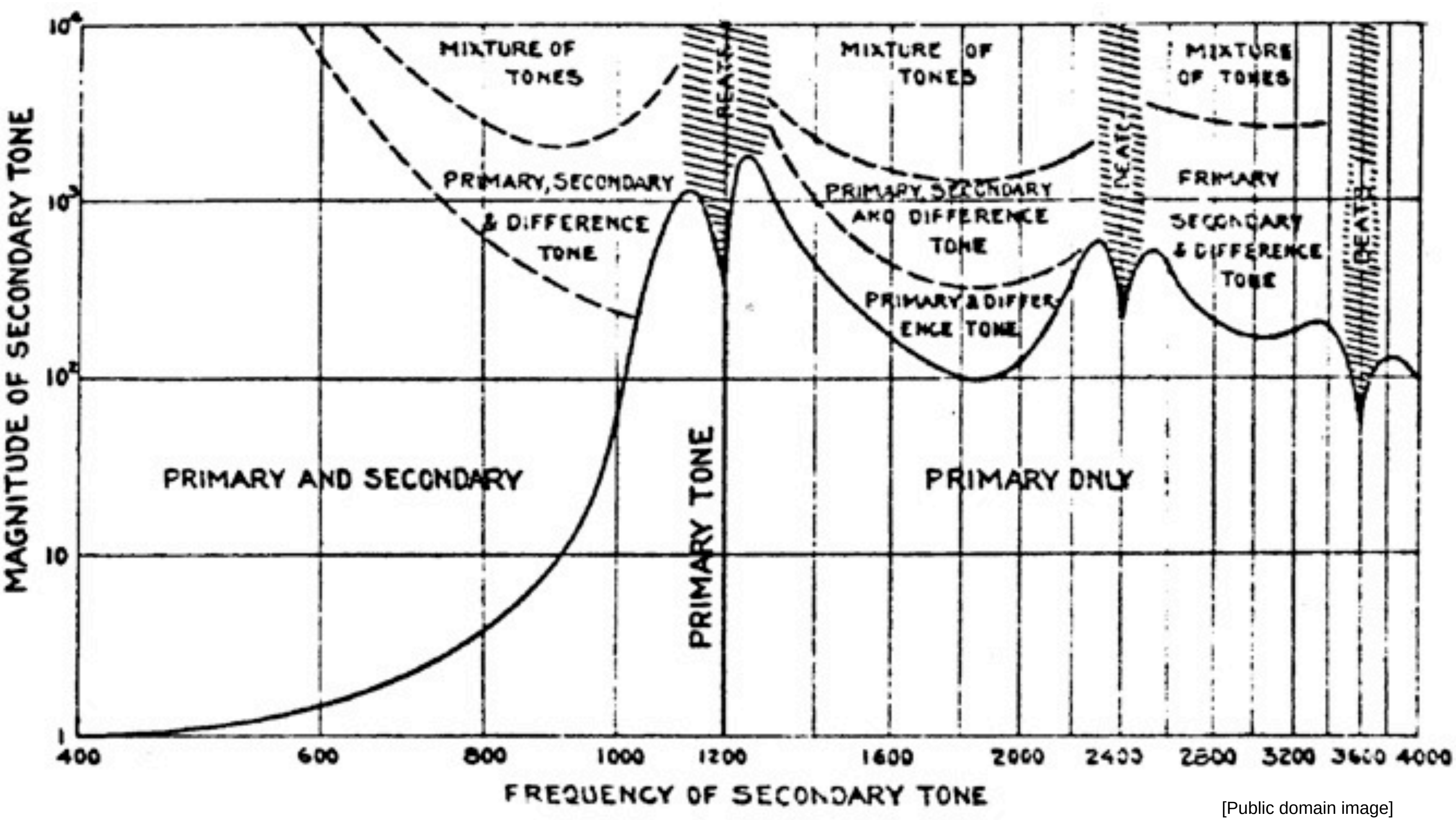
Courtesy of Hans-Christoph Steiner. Used with permission. After J. M. Grey, Stanford PhD Thesis (1975) and Grey and Gordon, JASA (1978).

Interference interactions between tones



[Public domain image]

Masking audiograms



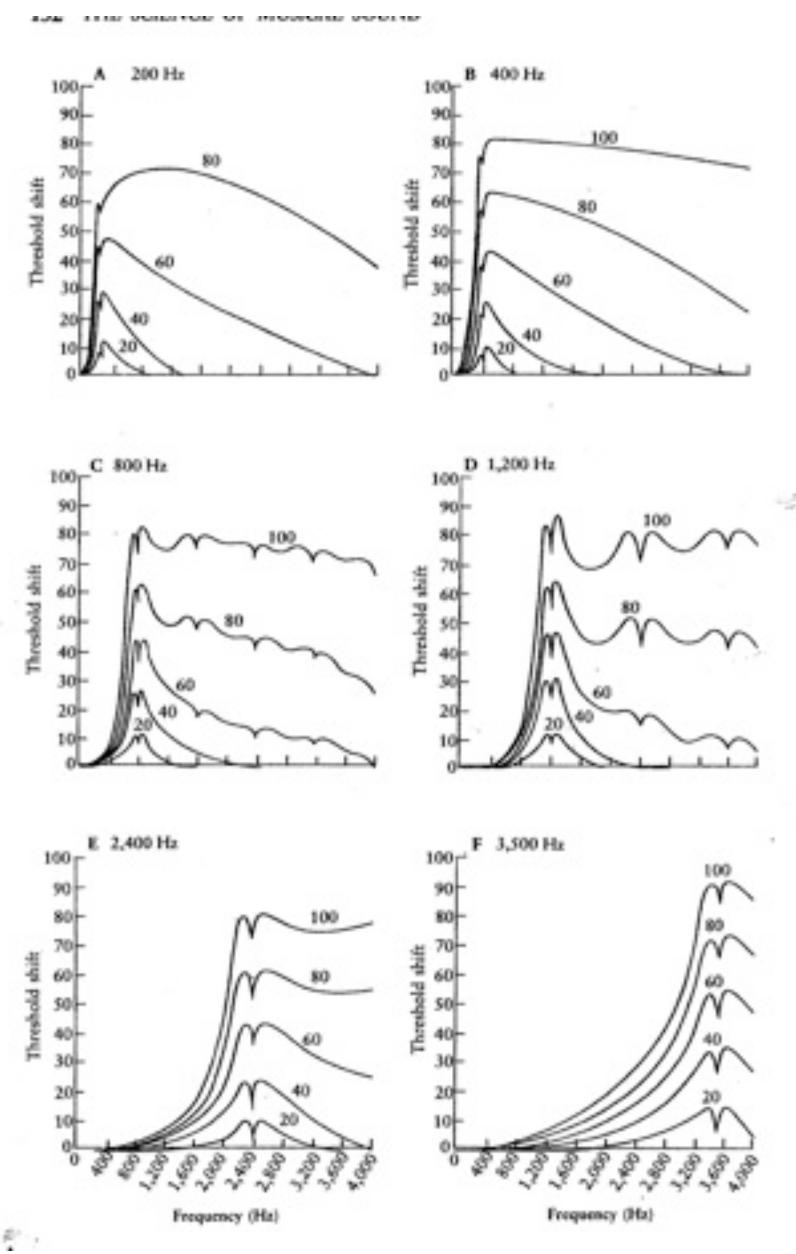
[Public domain image]

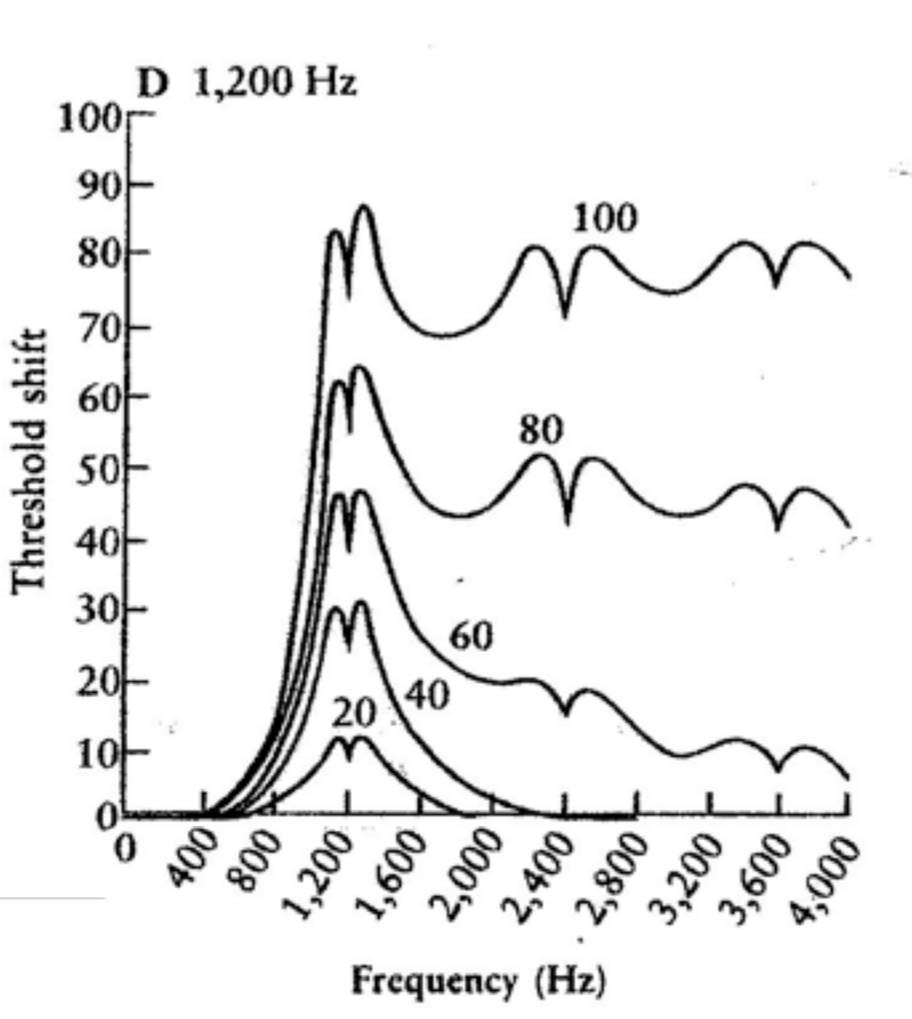
Wegel & Lane, 1924

1000 Hz pure tone masker

Graph removed due to copyright restrictions.
See Fig. 3, "Average masking patterns for 1000 cps based upon three listeners" in Ehmer, Richard H. "Masking Patterns of Tones." The Journal of the Acoustical Society of America, vol. 31, no. 8 (1959): 1115. http://www.zainea.com/masking2.htm

Tone on tone masking curves (Wegel & Lane, 1924)





[Public domain images]

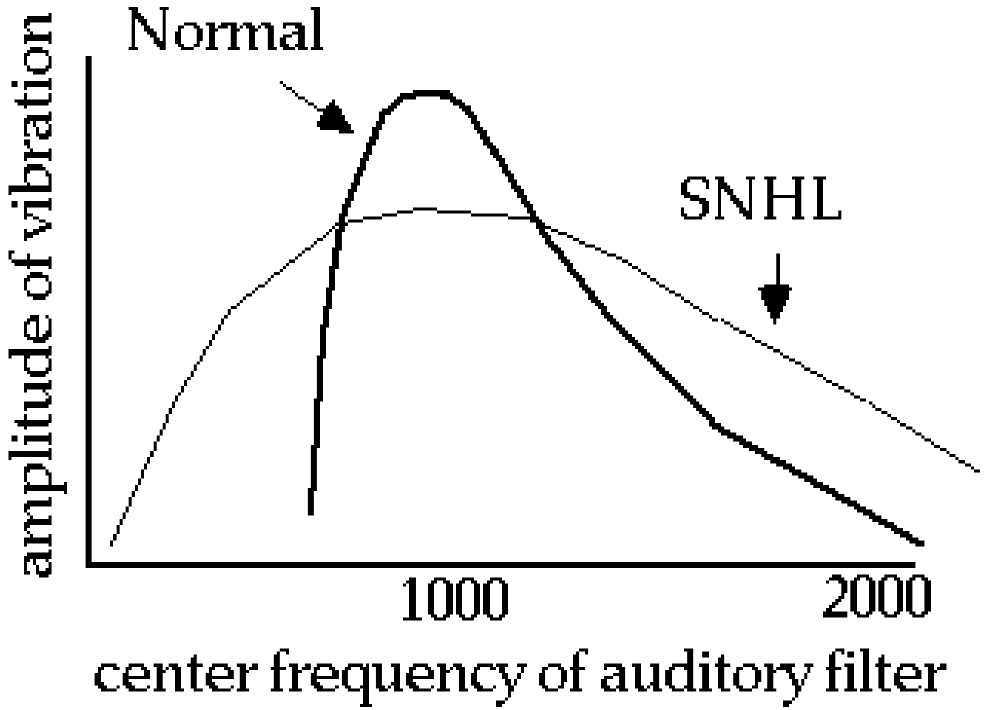
Resolvability of harmonics (Plomp, 1976)

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.

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

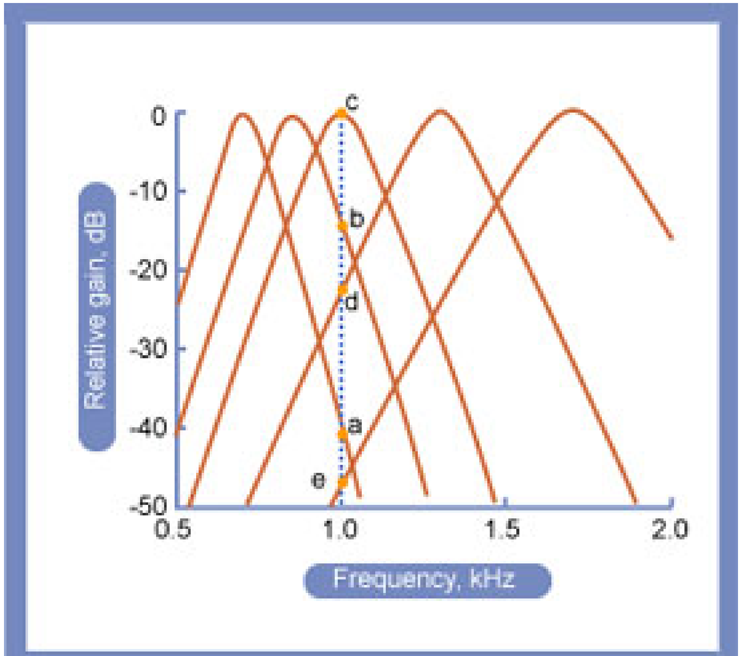
Power spectrum
Filter metaphor

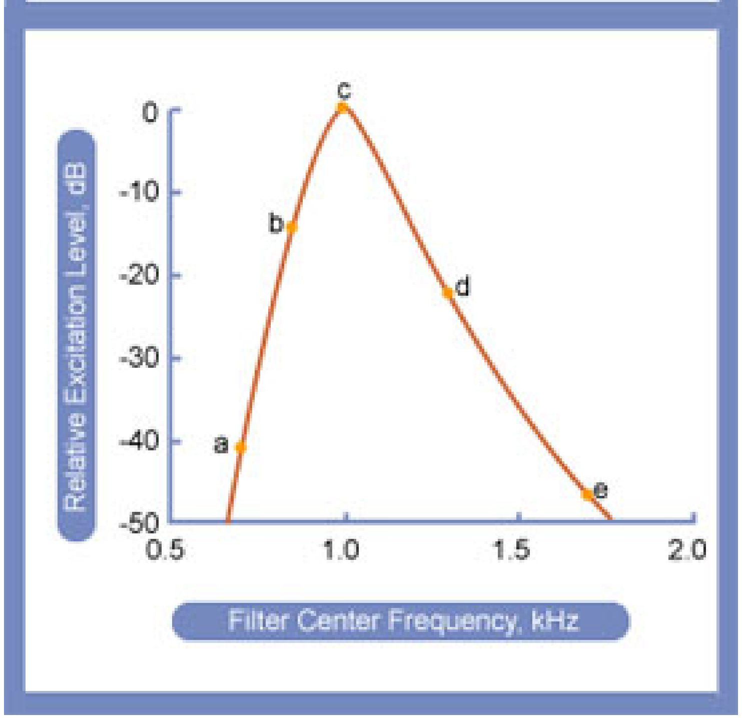
Notion of one central spectrum that subserves



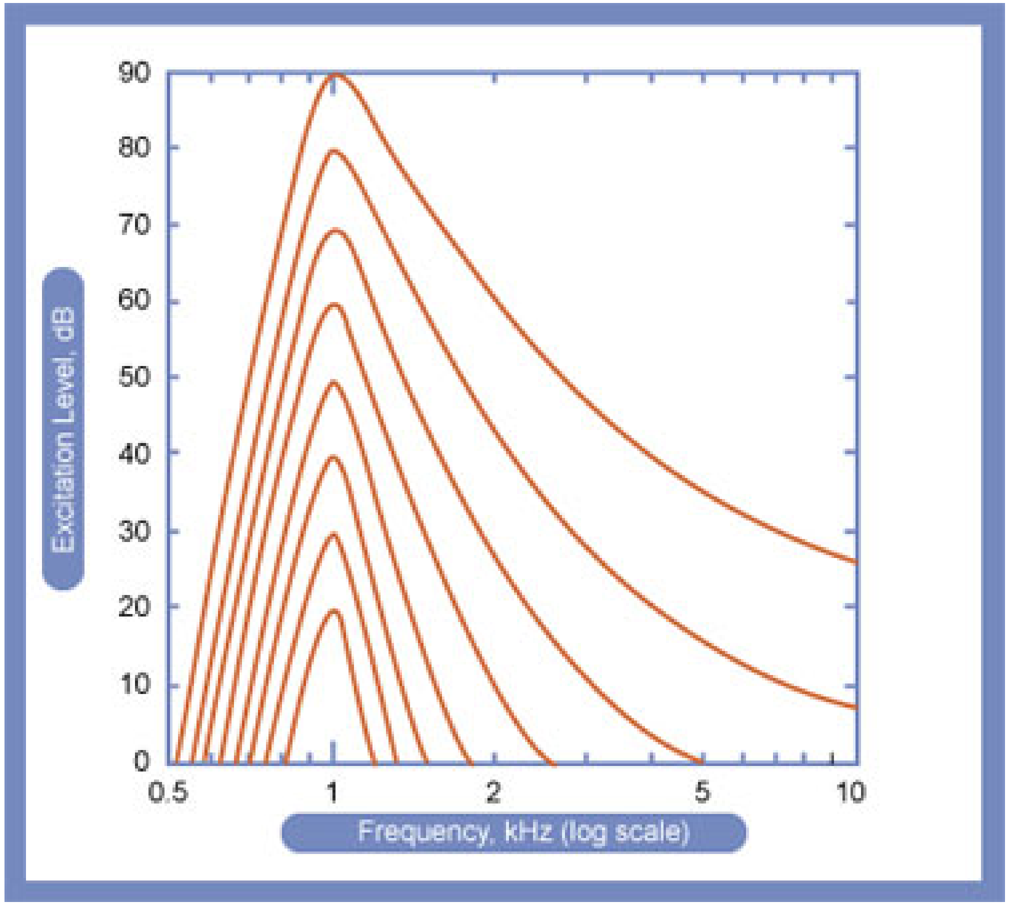
2.2. Excitation pattern. Using the filter shapes and bandwidths derived from masking experiments we can produce the excitation pattern produced by a sound. The excitation pattern shows how much energy comes through each filter in a bank of auditory filters. It is analogous to the pattern of vibration 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

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.

Resolution of harmonics

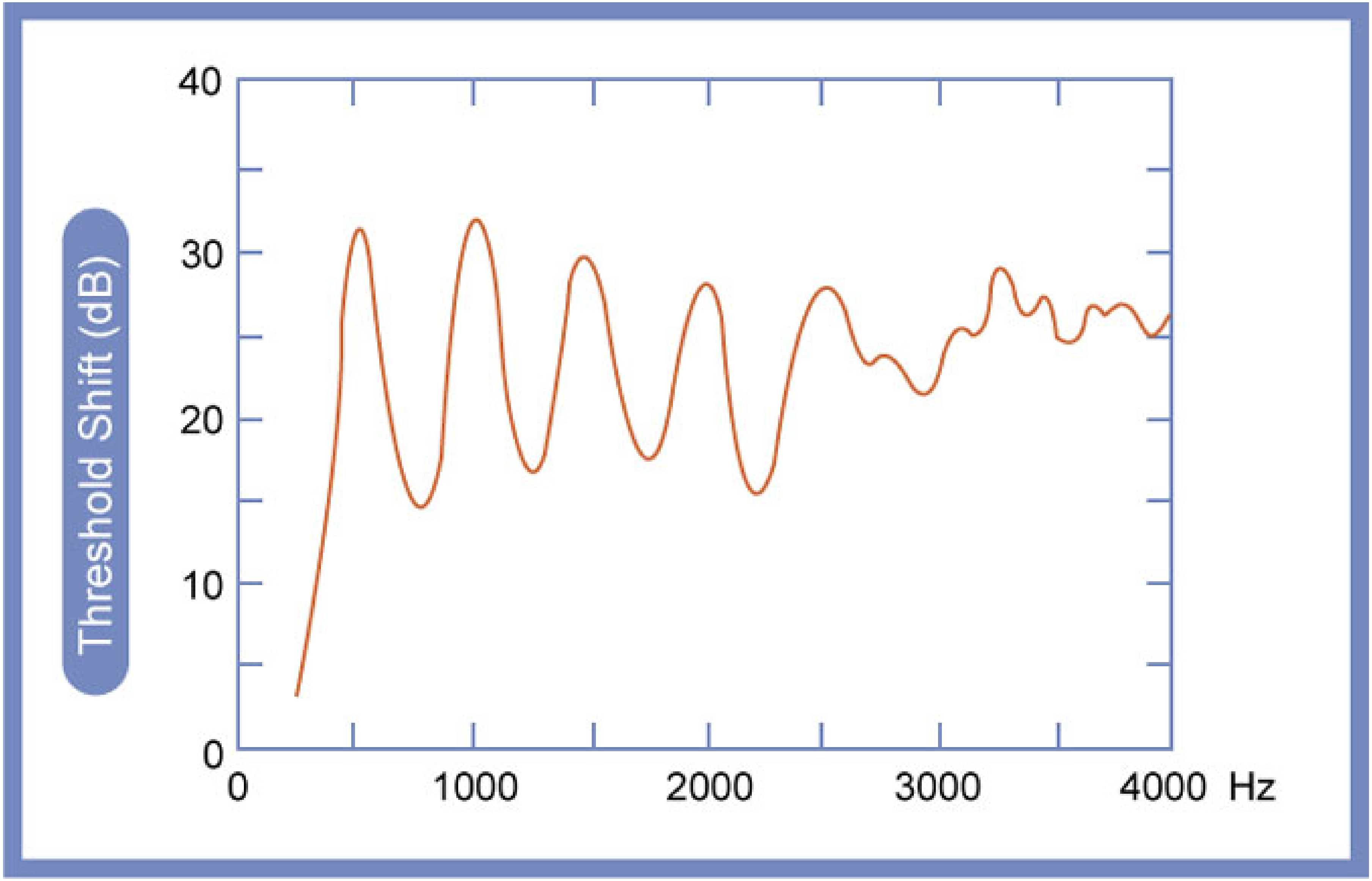
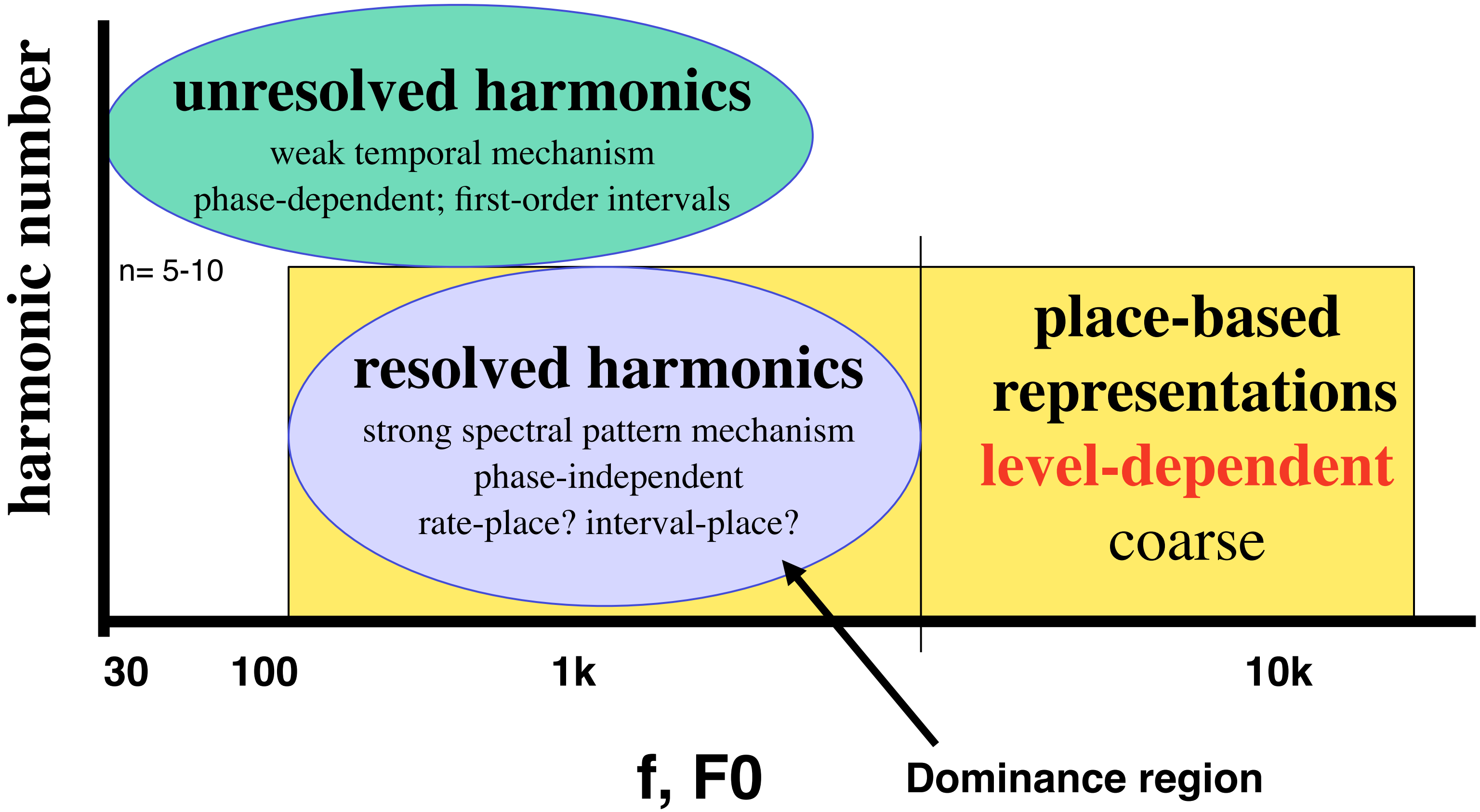


Figure by MIT OpenCourseWare.

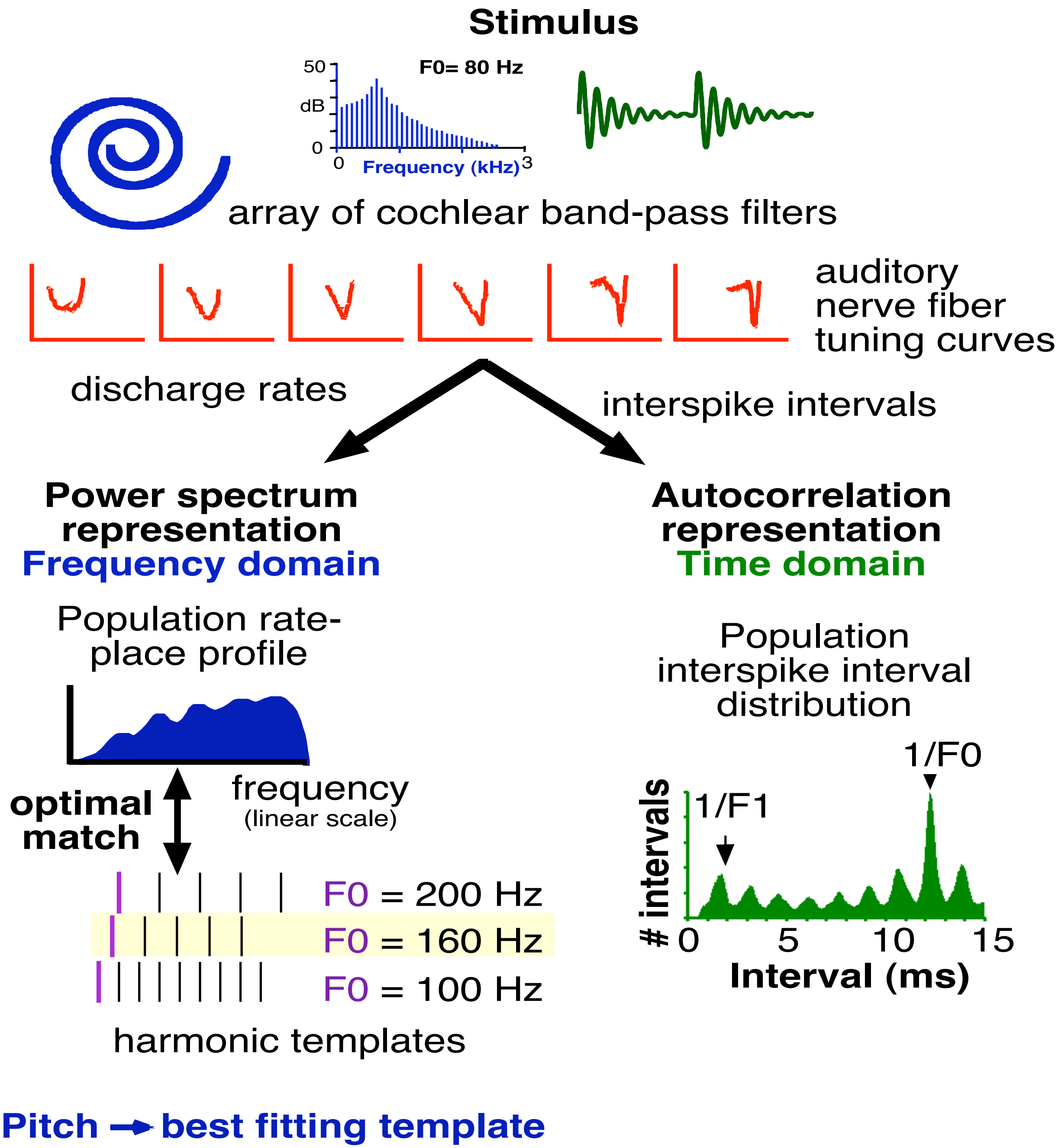
A "two-mechanism" perspective (popular with some psychophysicists)



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



JND's for pure tone frequency

(Roederer, 1995)

Graph removed due to copyright restrictions.
Fig. 2.9 in Roederer, J. G. *The Physics and Psychophysics of Music: An Introduction*. New York, NY: Springer, 1995.

Critical bands as tonal fusion

(Roederer, 1995)

Partial fusion,
partial loudness summation

Two sine waves,
one fixed at 400 Hz,
the other ascending
from 400 Hz to 510 Hz
at which point it is
separated from the
first by a
critical bandwidth.

Graph removed due to copyright restrictions.
Fig. 2.12 in Roederer, J. G. *The Physics and Psychophysics of Music: An Introduction*. New York, NY: Springer, 1995.
See 2nd graph on this page:
http://www.sfu.ca/sonic-studio/handbook/Critical_Band.html

Critical bandwidths

(Roederer, 1995)

Graph removed due to copyright restrictions.
Fig. 2.13 in Roederer, J. G. *The Physics and Psychophysics of Music: An Introduction*.
New York, NY: Springer, 1995.
See 1st graph on this page: http://www.sfu.ca/sonic-studio/handbook/Critical_Band.html

Critical bands (conventional, place-based interpretation)

CRITICAL BAND and CRITICAL BANDWIDTH

For a given FREQUENCY, the critical band is the smallest BAND of frequencies around it which activate the same part of the BASILAR MEMBRANE. Whereas the DIFFERENTIAL THRESHOLD is the just noticeable difference (jnd) of a single frequency, the critical bandwidth represents the ear's resolving power for simultaneous tones or partials.

In a COMPLEX TONE, the critical bandwidth corresponds to the smallest frequency difference between two PARTIALs such that each can still be heard separately. It may also be measured by taking a SINE TONE barely MASKed by a band of WHITE NOISE around it; when the noise band is narrowed until the point where the sine tone becomes audible, its width at that point is the critical bandwidth. See: RESIDUE.

In terms of length (see diagram under BASILAR MEMBRANE) the critical bandwidth is nearly constant at 1.2 mm, within which are located about 1300 receptor cells, and is generally independent of intensity (unlike COMBINATION TONES). Twenty-four critical bands of about one-third octave each comprise the audible spectrum.

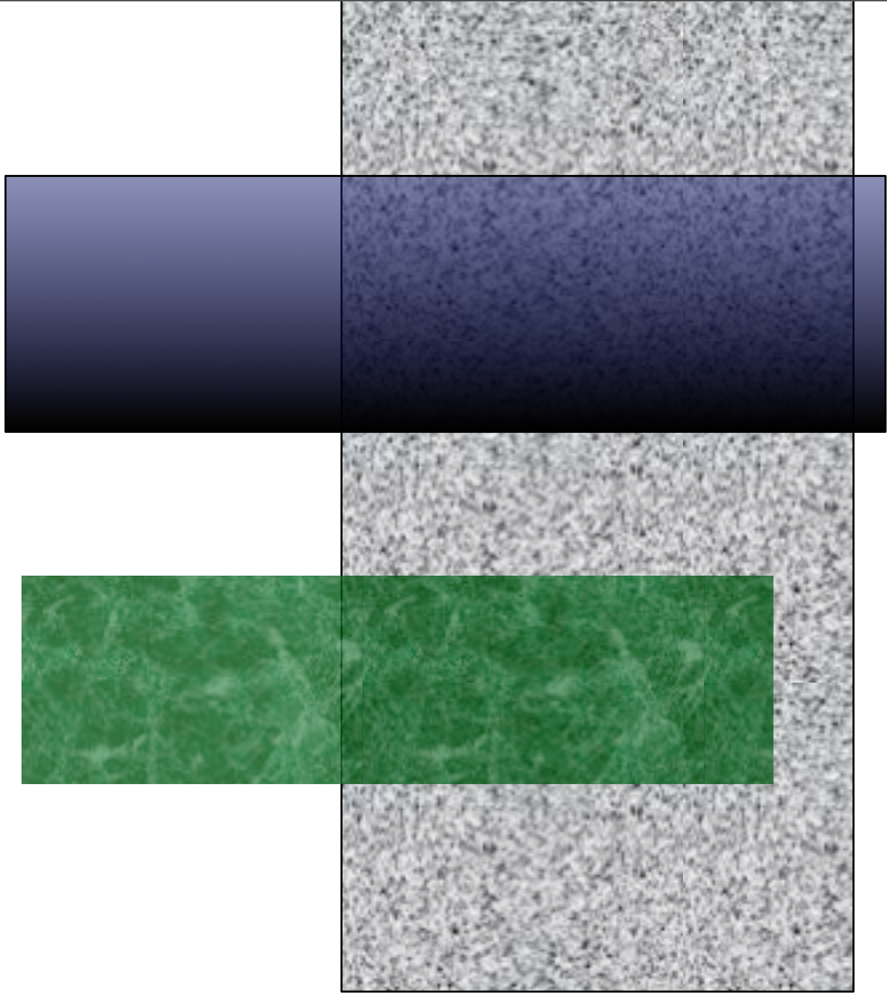
Truax, B., ed. From "CRITICAL BAND and CRITICAL BANDWIDTH." http://www.sfu.ca/sonic-studio/handbook/Critical_Band.html
Handbook for Acoustic Ecology. 2nd edition, 1999. Courtesy of Barry Truax. Used with permission.

Critical bands (usually interpreted in terms of frequency analysis)

Simultaneous tones lying within a critical bandwidth do not give any increase in perceived loudness over that of the single tone, provided the sound pressure level remains constant. For tones lying more than a critical bandwidth apart, their combination results in increased loudness.

When two tones are close together in frequency, BEATS occur, and the resulting tone is a fusion of the two frequencies. As the frequency difference increases, roughness in the tones appears, indicating that both frequencies are activating the same part of the basilar membrane. Further apart, the two frequencies can be discriminated separately, as shown below by DfD, whereas roughness only disappears at a frequency separation equal to the critical bandwidth DfCB. At this point, the two frequencies activate different sections of the basilar membrane. This phenomenon only applies to monaural listening with pure tones. With DICHOTIC listening, the basilar membrane of each ear is activated separately, and therefore no roughness results. With complex tones, frequency discrimination is improved but the critical bandwidth remains the same for each of the component partials.

Alternative interpretation is that critical bandwidths are the result of fusion of (e.g. interspike interval) representations rather than cochlear proximity per se.



Masking by signal swamping
Reduce signal/noise to disrupt signal detection

Camouflage: pattern fusion
Disruption of pattern detection

Photo courtesy of kirklandj on Flickr.

Re: varieties of masking in the auditory system, see Delgutte (1988) Physiological mechanisms of masking. In. Duifhuis, Horst & Wit, eds. Basic Issues in Hearing. London. Academic Press, 204-14.

Spatial hearing

Azimuth:

interaural time differences (20-600 usec)

interaural level differences

Elevation:

received spectrum of broadband sounds (pinna effects)

Distance

Spatial form (size, shape)

Enclosure size, shape

Reverberation pattern

Patterns of long delays

Assignment of spatial attributes to auditory objects

Diagram showing effect of interaural path-length differences.

Figure 2.1 in Warren, R. M. *Auditory Perception: A New Synthesis*. New York, NY: Pergamon Press, 1982. ISBN: 9780080259574.

Interaural time difference and localization of sounds

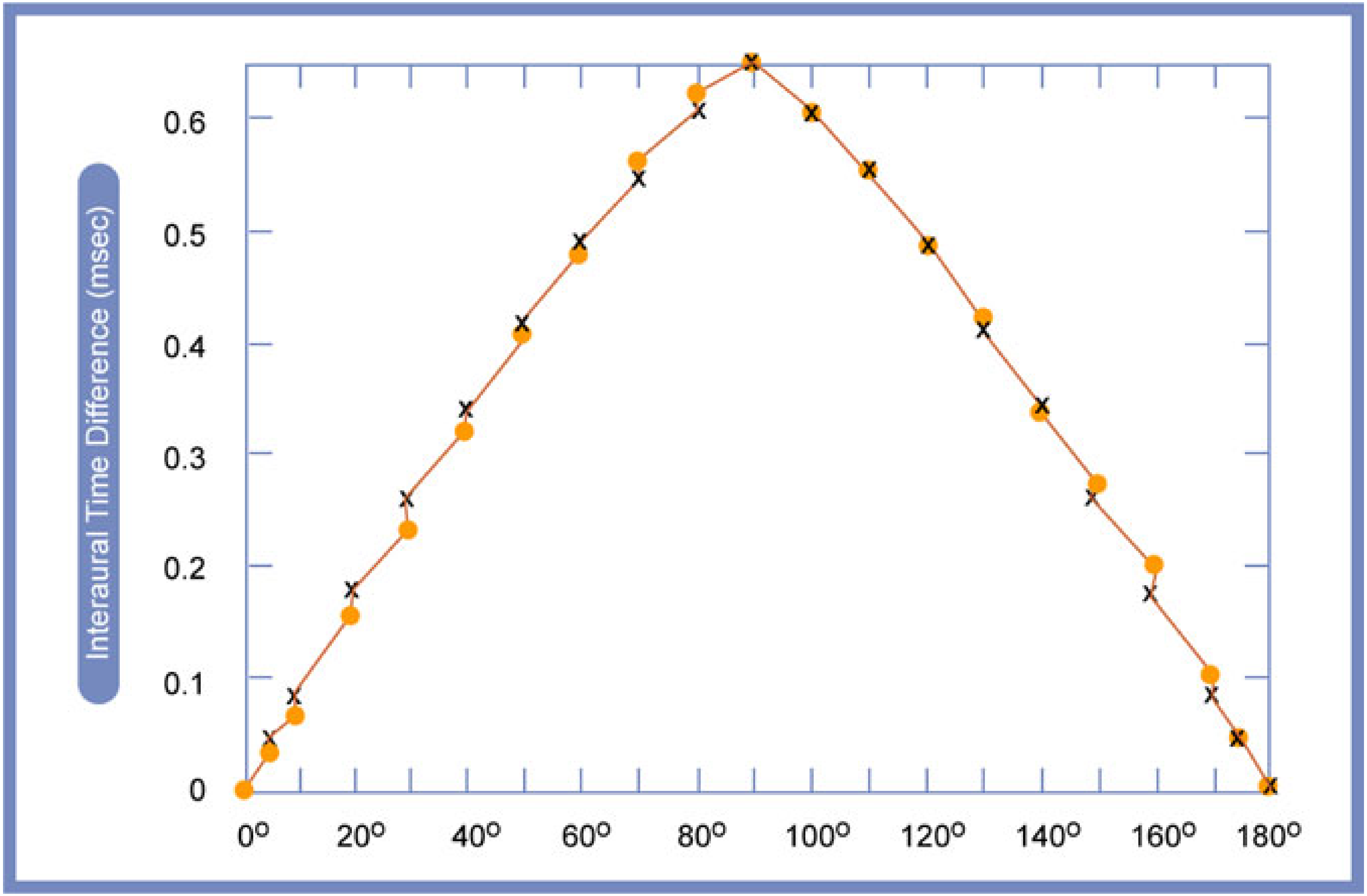


Figure by MIT OpenCourseWare.

Binaural cross-correlation and cancellation

Binaurally-created pitches

Tones (F0 from one harmonic in each ear)
Phase-disparity pitches (auditory analog
of Julez random-dot stereodiagrams)
Repetition pitches (weak)

Binaural masking release (BMLD)

A tone in noise that is just masked is
presented to one ear. The tone cannot be heard initially.
Now also present the identical noise alone in the other ear
and the tone pops out. The noise appears to be cancelled
out, providing up to 15 dB of unmasking.

Generalist vs. specialist sensory systems (conjectures)

- General-purpose vs. special-purpose systems
- Adaptability vs. adaptedness

Adaptable: optimized for many different env's

Adaptedness: high degree of optimization

Tradeoff between the two

Panda gut (highly adapted) vs. human gut (omnivore, high adaptability)

In sensory systems, high adaptability is favored when appearances are highly variable; adaptedness when appearances are highly constrained

Intra-species communications: adaptedness is favored

Signal production and reception under same genetic coordination

e.g. pheromone systems

Inter-species interactions (predator or prey): adaptability is favored under varying relations, adaptedness under stable relations

(e.g. navigation systems, early warning systems, predator or prey recognition)

Reading for Tuesday, Feb 13

Next meeting we will introduce neural coding and give an overview of the auditory system.

Weinberger chapter in Deutsch (3)

Look over auditory physiology chapter in Handel (12)

HST.725 Music Perception and Cognition
Spring 2009