Psychoacoustics

Fundamental Definitions

- **Acoustics**
  - branch of physics concerned with the study of sound/ science of sound
  - studies audibility and characteristics of sound propagation (typically in a closed space (room); reverberation, etc)
- **Psychoacoustics**
  - branch of science studying the psychological and physiological responses associated with sound (including noise, speech and music). It can be further categorized as a branch of psychophysics which aims at **linking perception of acoustical stimuli with auditory sensations**

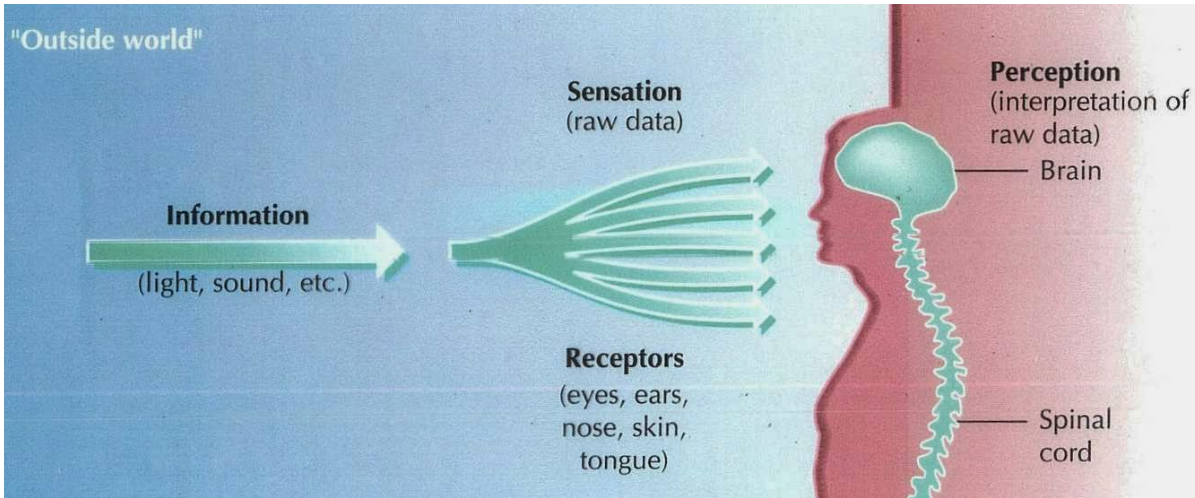
Fundamental Question of Psychoacoustics

- how do certain perceptual qualities of sound depend on the physical properties of sound?

SENSATION ? PERCEPTION

SENSATION

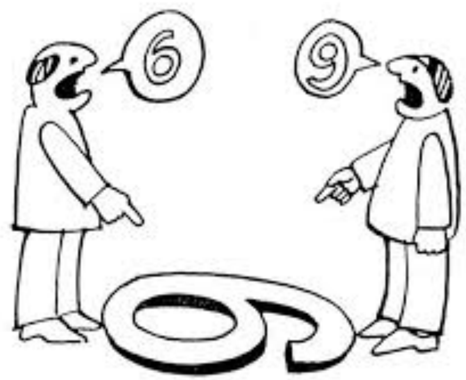
- gathering info from the environment via your senses (ex: visual, auditory, etc) and transmit them to the brain
- sensory information that has registered in the brain but has not been interpreted



PERCEPTION

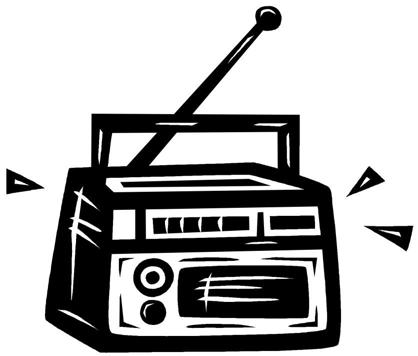
the ILLUSION of REALITY

- process by which sensory information is actively organized and interpreted by the brain
- understanding what is being sensed
- integration of multiple sensory streams (eg: audio + visual)



SENSATION

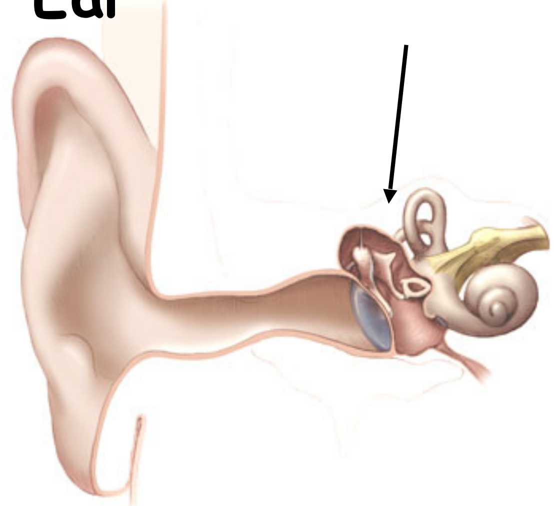
How do we sense sound?



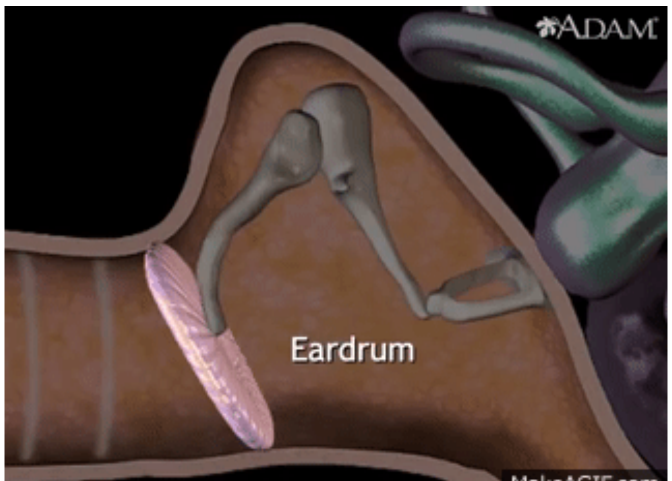
Sound Waves

Outer Ear

Middle Ear

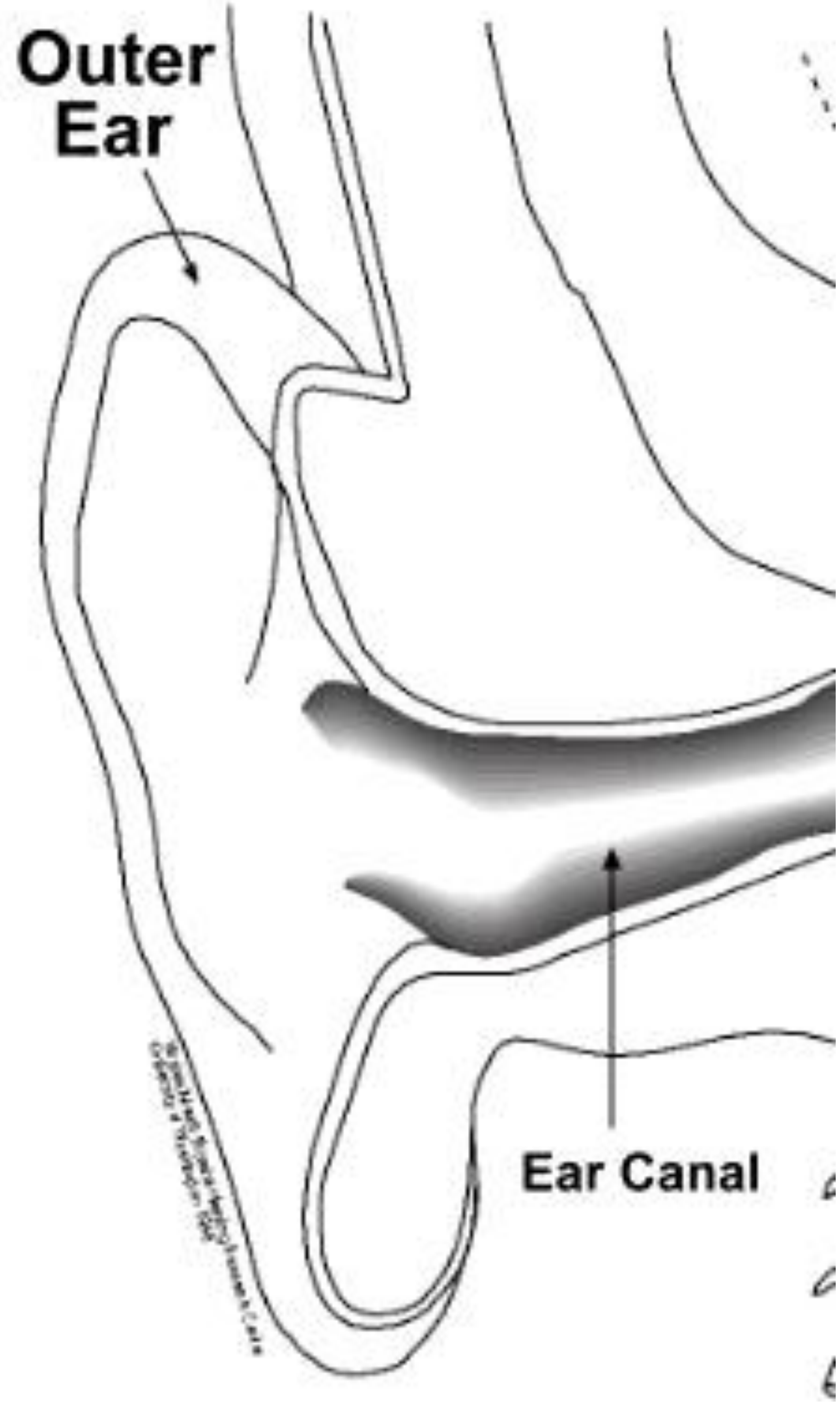


Inner Ear



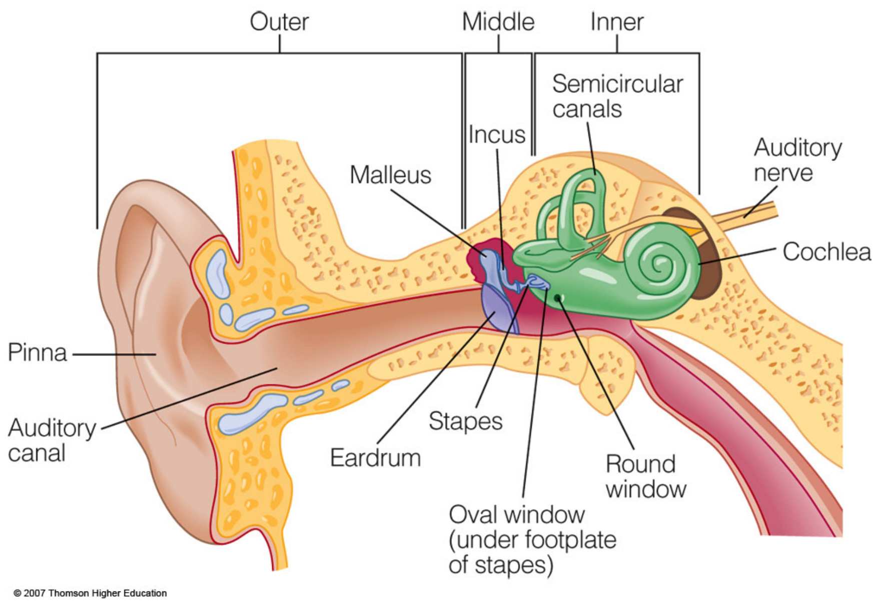
The Outer Ear

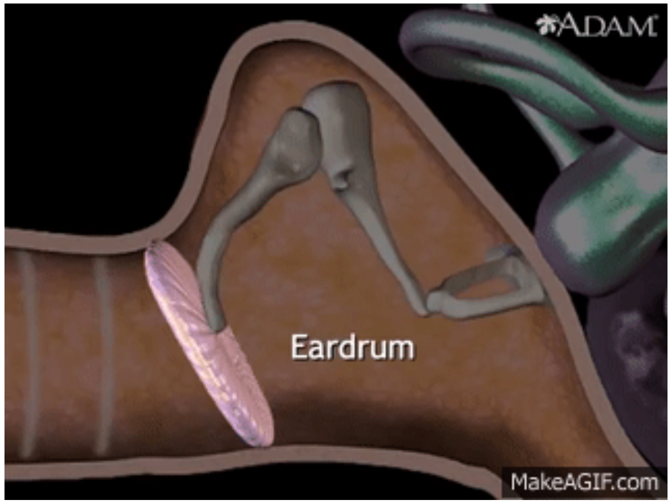
- The curved formation on the outside (the pinna) helps funnel sound down the ear canal to the eardrum



The Middle Ear

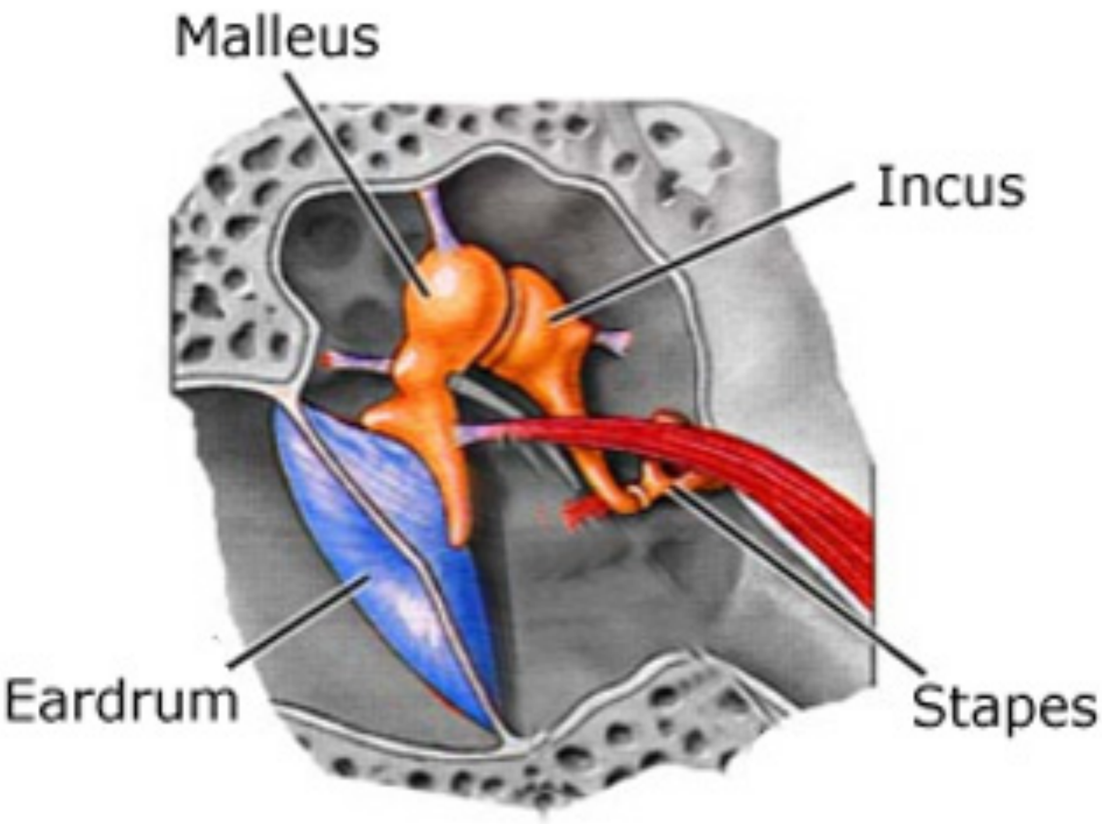
- transfers the energy of a sound wave by vibrating the three bones found there.
- ossicles are arranged and interact with each other as a lever system
- **amplifier** - without them, only about 0.1 percent of sound energy would make it into the inner ear.





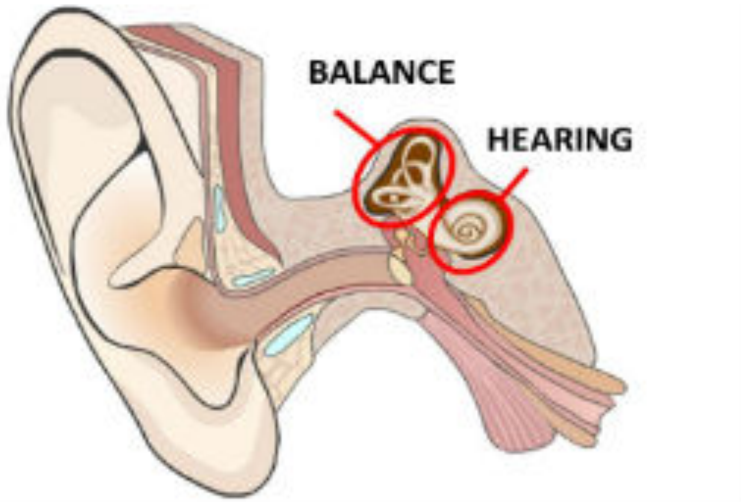
Bones of the Middle Ear

- These are the smallest bones in your body!



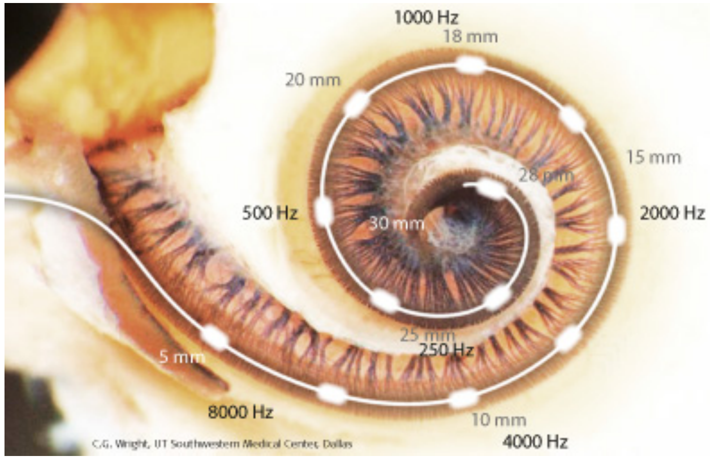
The Inner Ear

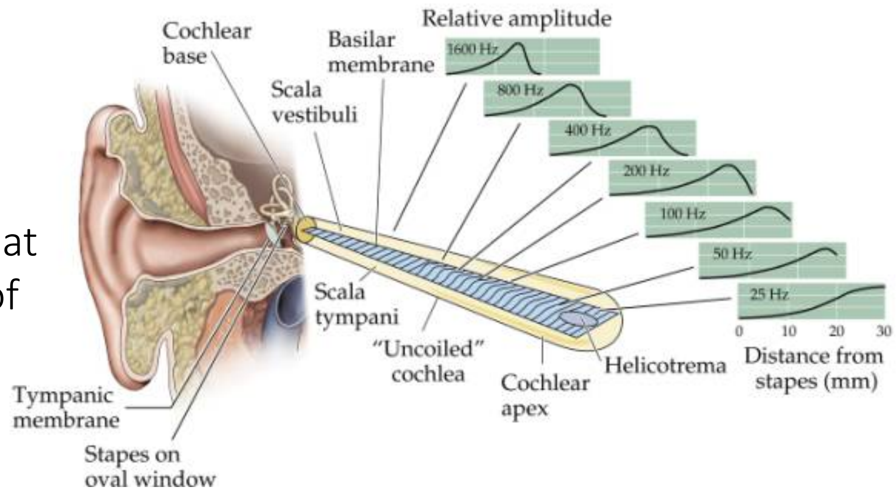
- Two main parts:
  - **Vestibular system** (**balance**) - Semicircular Canals- Fluid filled tubes attached to the cochlea that help us maintain our sense of balance
  - **Cochlea** (**hearing**) converting sound pressure patterns from the outer ear into electrochemical impulses which are passed on to the brain via the auditory nerve.



The Inner Ear

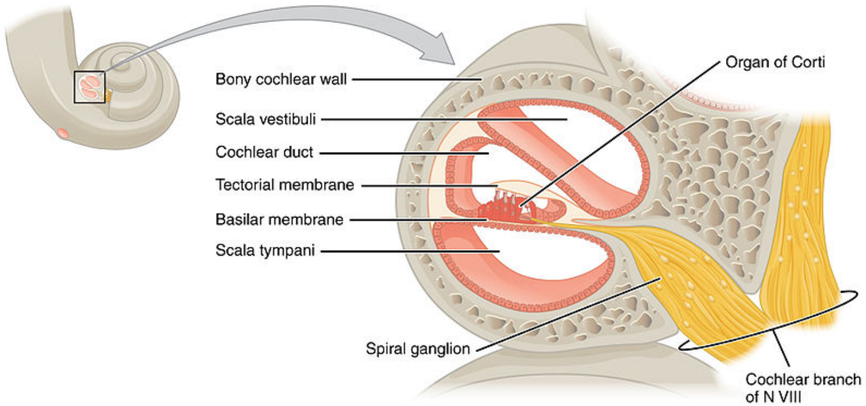
- **Cochlea**
  - Coiled like a snail shell, fluid-filled; it is lined with cilia (tiny hair) that move when vibrated and cause a nerve impulse to form.
  - tonotopic map - a location code formed on the cochlea
- **Basilar membrane**
  - stiff long structural element varying in width and stiffness and bends rows of hair cells beneath it based on incoming signal
  - base for the sensory cells of hearing & acts as a frequency analyser
  - the front end (base) of the membrane to resonate with higher frequencies, and its rear end (apex) with lower frequencies
  - bending of the hair cells gives rise to electric impulses that encode information about the periodicity and intensity of the sound.

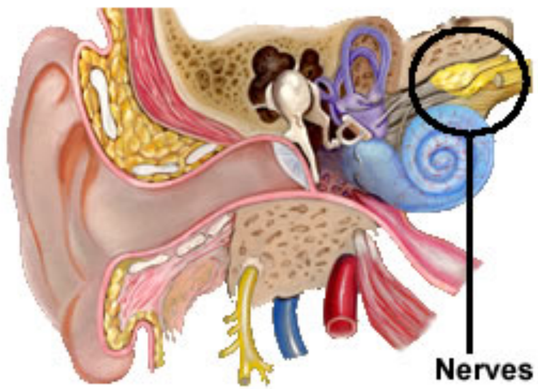




The Inner Ear

- **Organ of Corti**
  - contains auditory receptors that transduce auditory signals into nerve impulses' action potential
- **Auditory/Cochlear nerve**
  - these carry electro-chemical signals from the inner ear (the cochlea) to the brain.

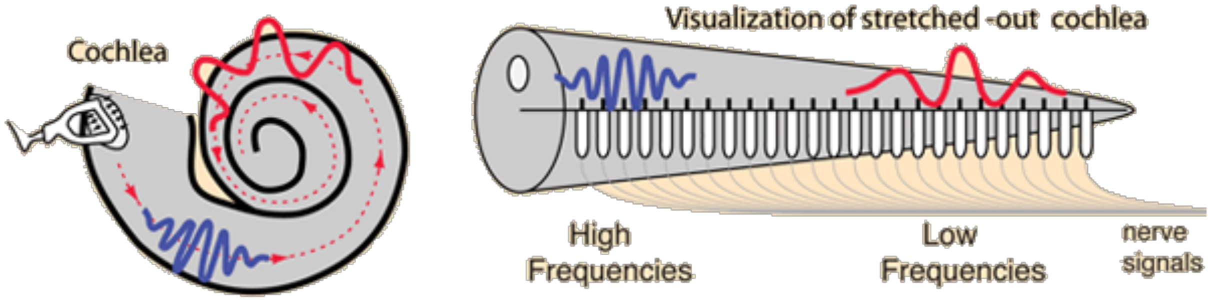


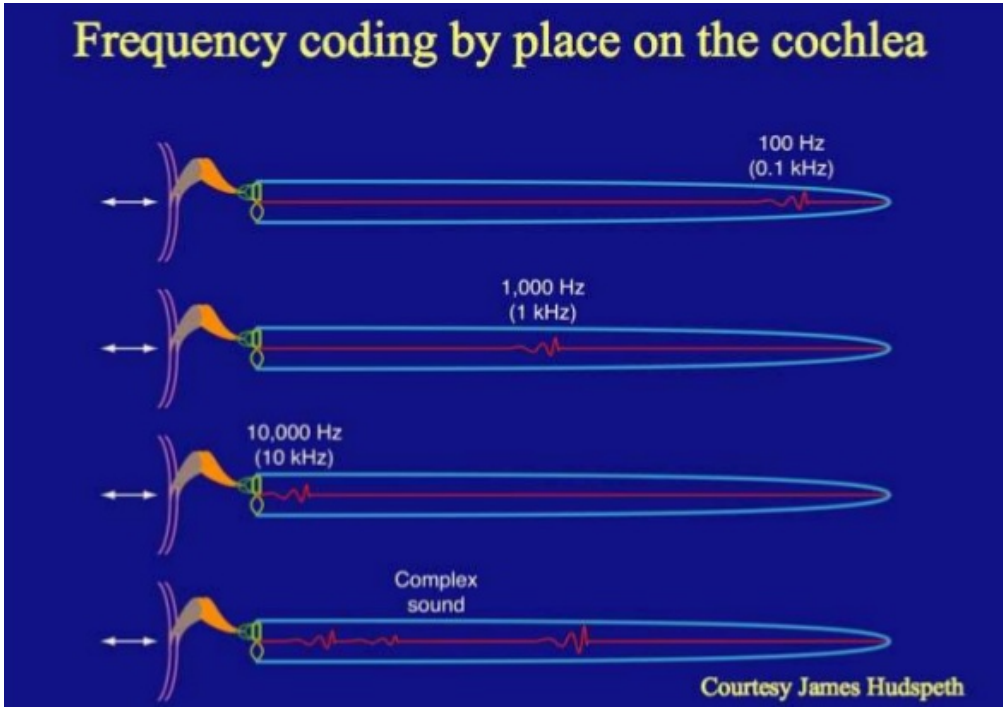


Theories of Hearing

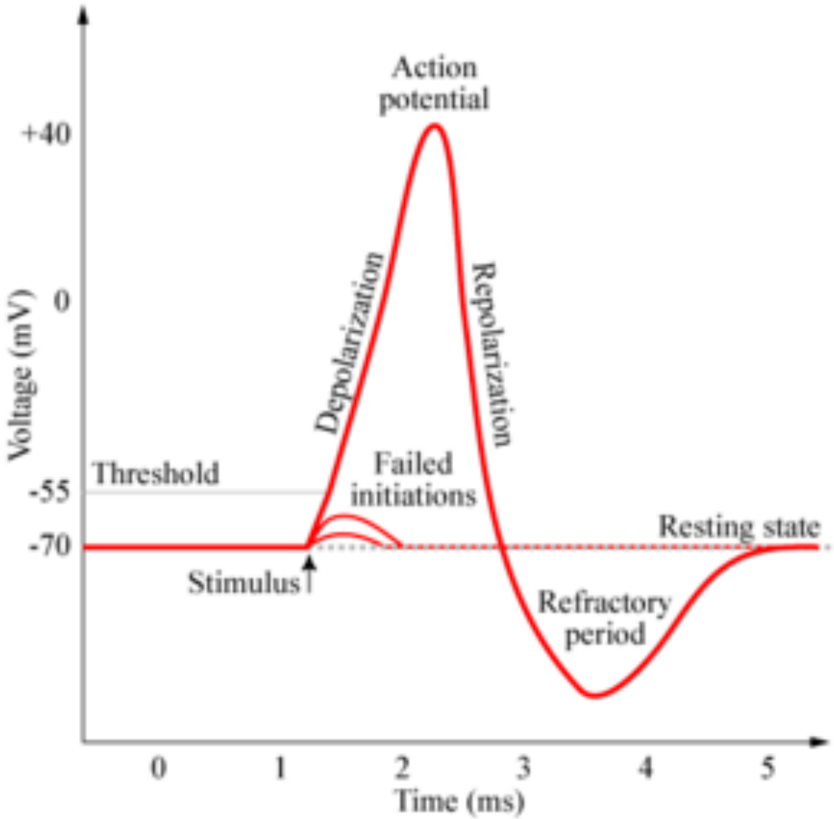
- **Place Theory**
  - encoded at different places on the basilar membrane
- **Temporal Theory**
- **Volley Theory**

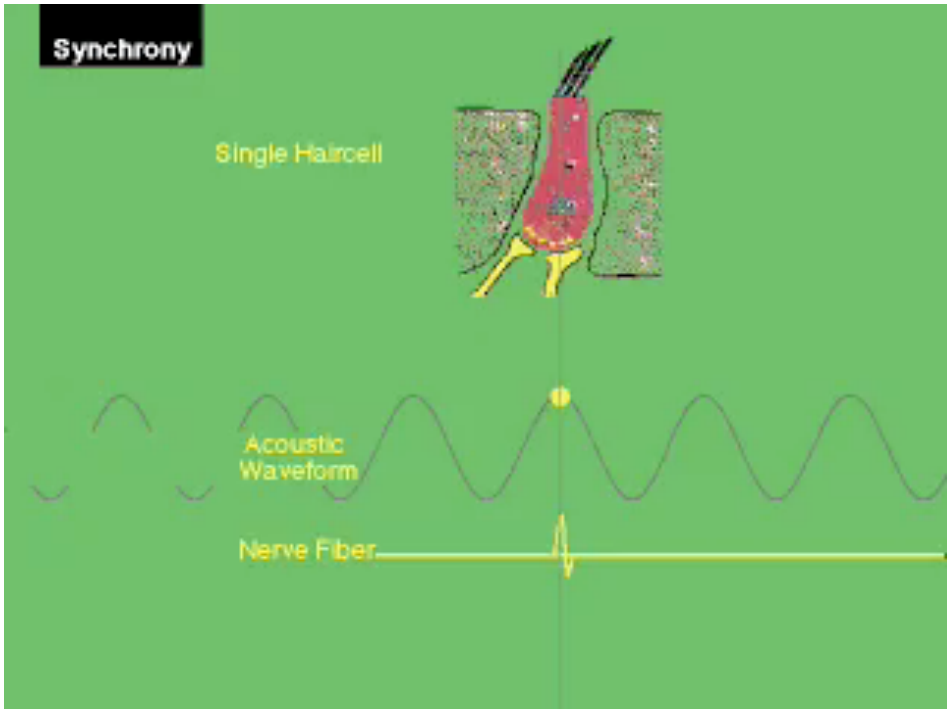
Frequency Coding by place on Cochlea
Place Theory





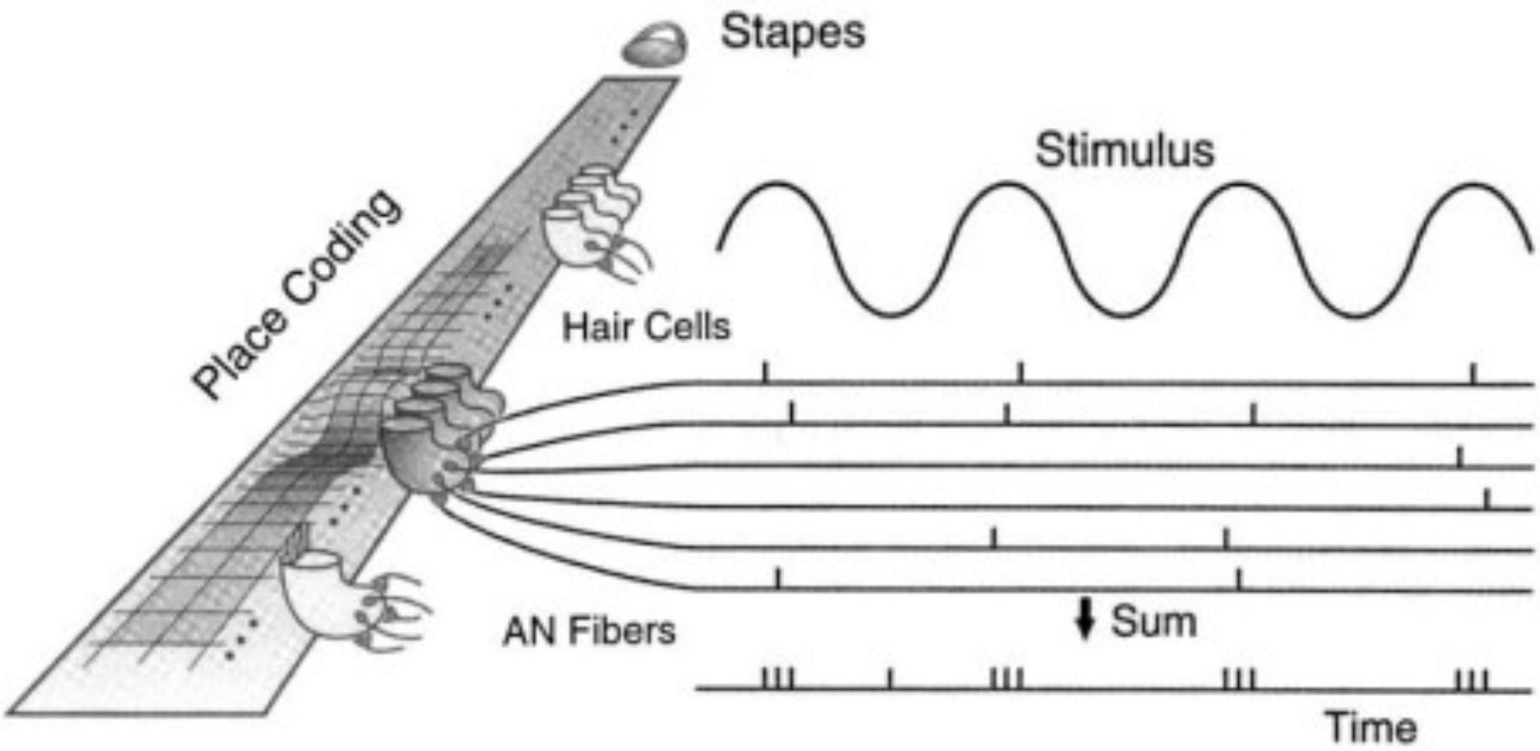
But: neurons are slow





Each neuron can send at most ~1000 impulses/sec
How are frequencies > 1000 Hz encoded?

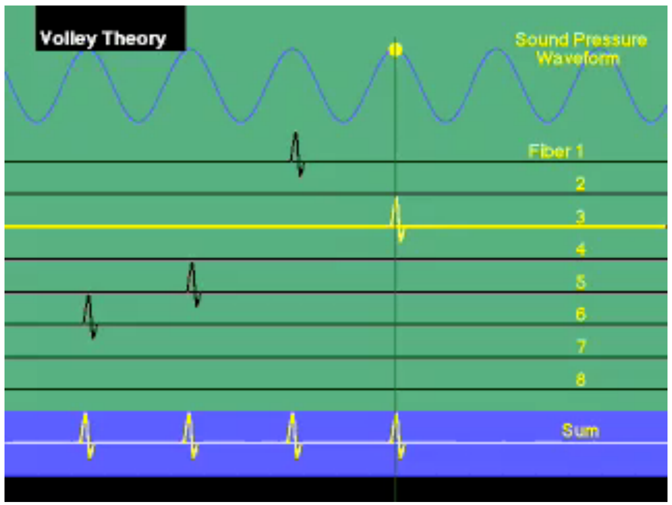
Temporal Coding



Temporal Coding

Volley Theory

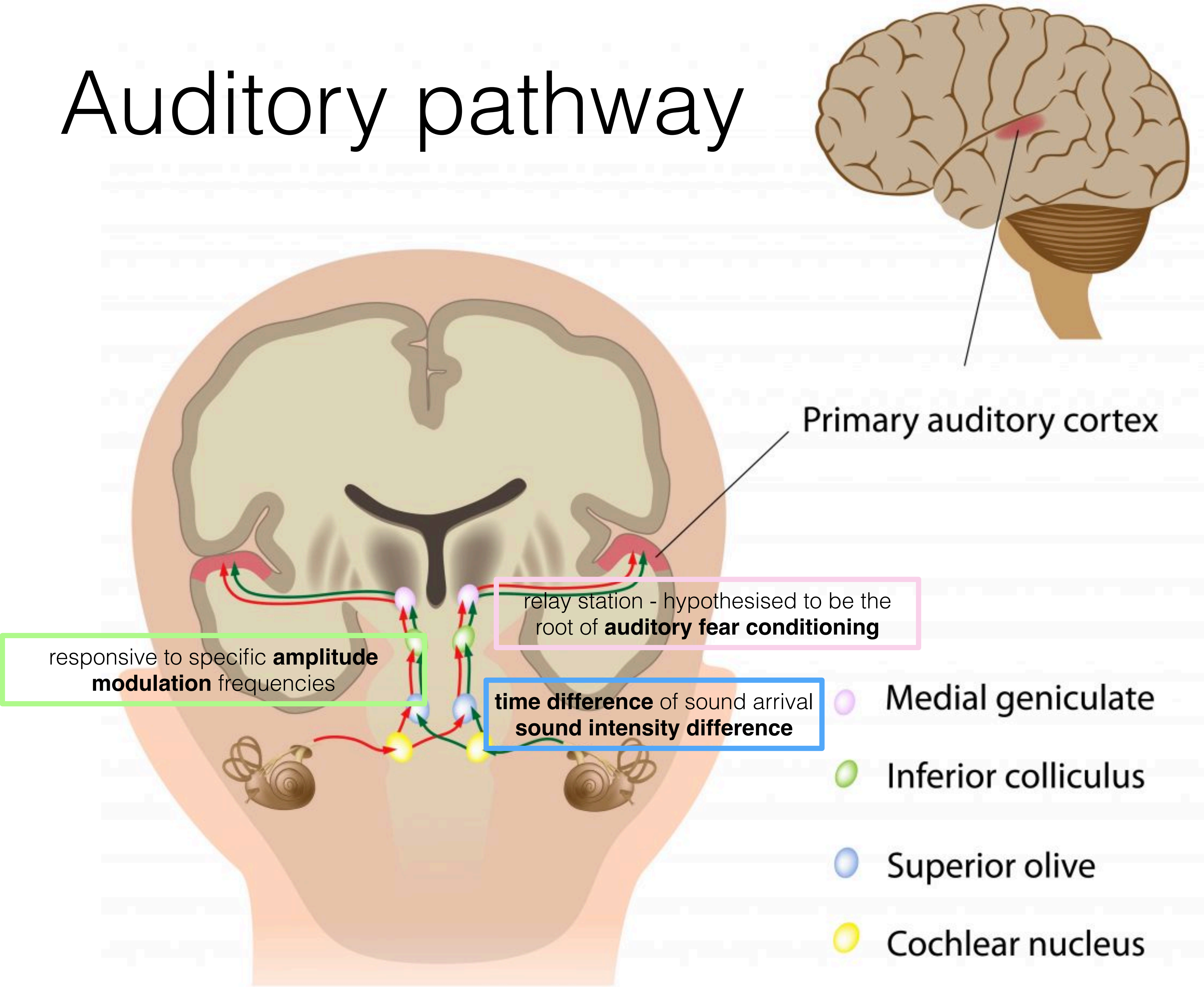
- synchronization of several adjacent neurons encodes periodicity information
- groups of auditory neurons use phase-locking to represent subharmonic frequencies of one harmonic sound



Theories of Hearing

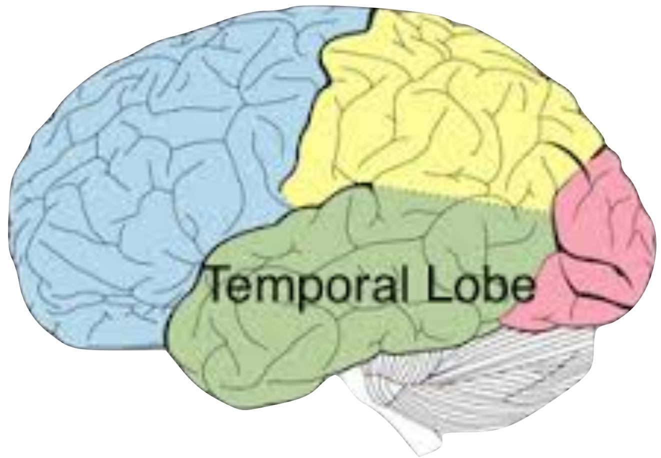
- **Place Theory**
  - encoded at different places on the basilar membrane
- **Temporal Theory**
  - pitch encoded in timing of neural firings along the basilar membrane
  - nerve firings occur at particular phases of the waveform
- **Volley Theory**
  - synchronization of several adjacent neurons encodes periodicity information
  - groups of auditory neurons use phase-locking to represent subharmonic frequencies of one harmonic sound

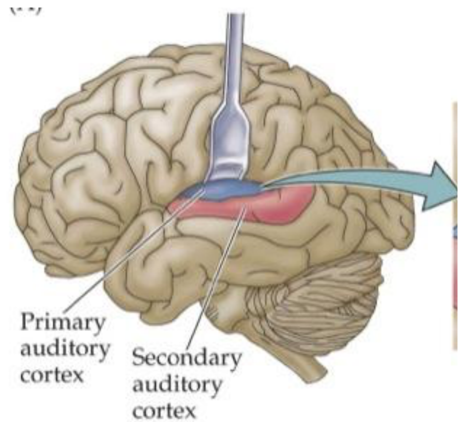
Auditory pathway



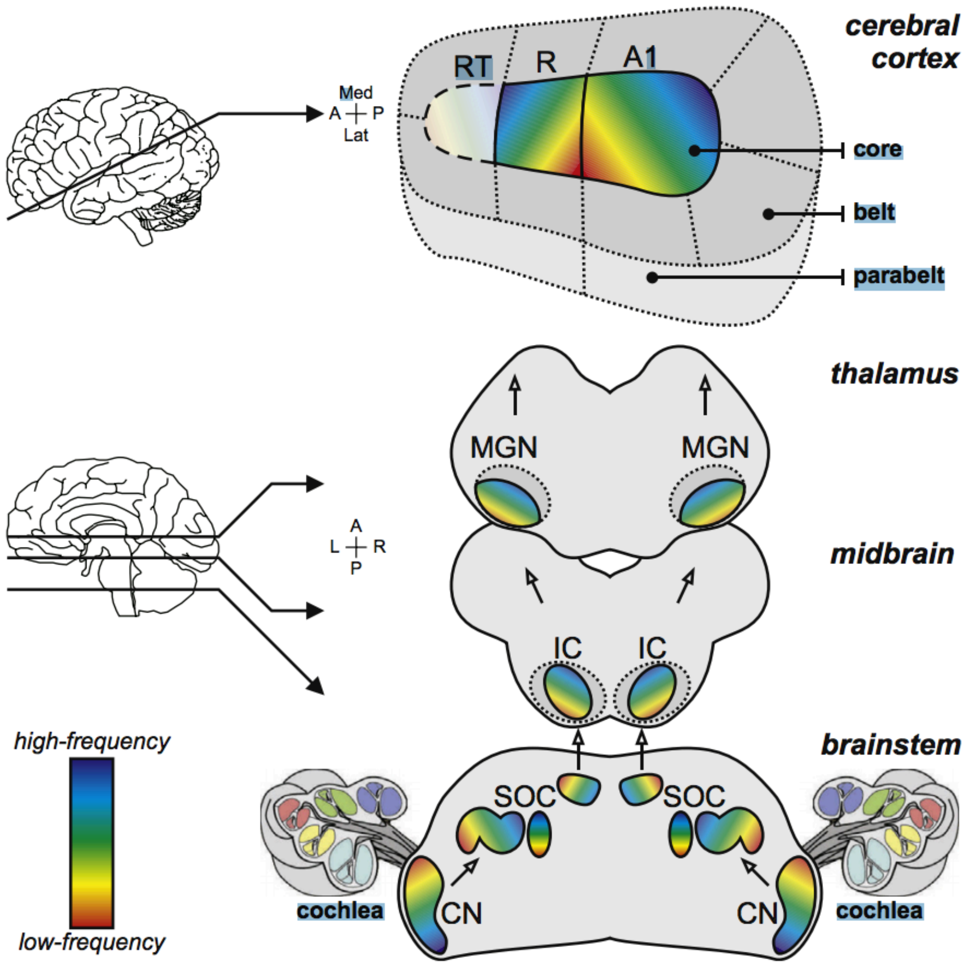
Auditory cortex

- **primary auditory cortex (PAC)**
  - BA 41 (42) - Heschl's gyrus
  - in the temporal lobe - involved with sense of hearing
  - tonotopic organisation
  - projects to numerous secondary cortical areas including multisensory areas (allow us to recognise animals or humans by both sound and sight) and to regions specifically involved in communication



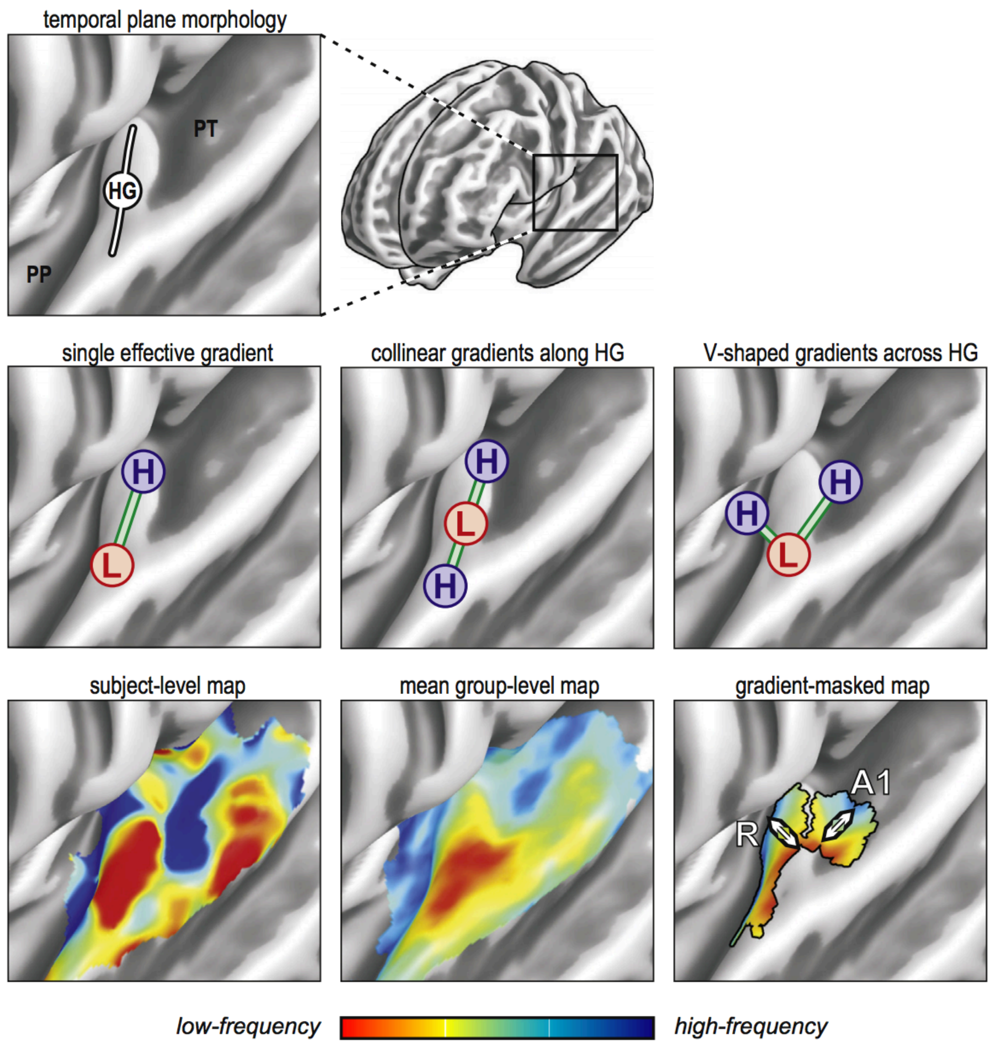


Frequency Encoding



Saenz & Langers (2013) Tonotopic mapping of human auditory cortex. Hearing Research.

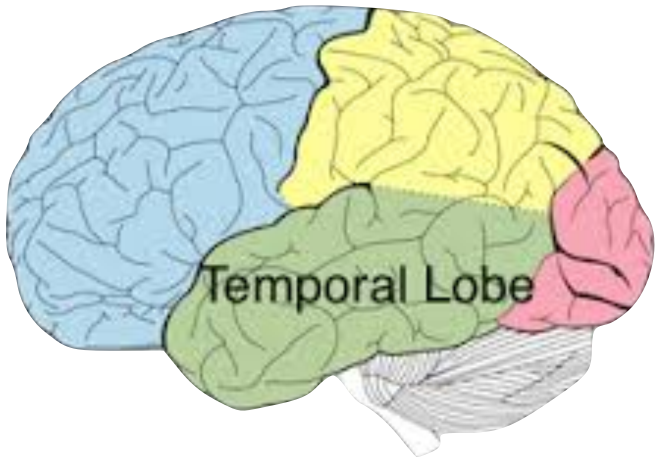
Evolution of Frequency Encoding Evidence in PAC

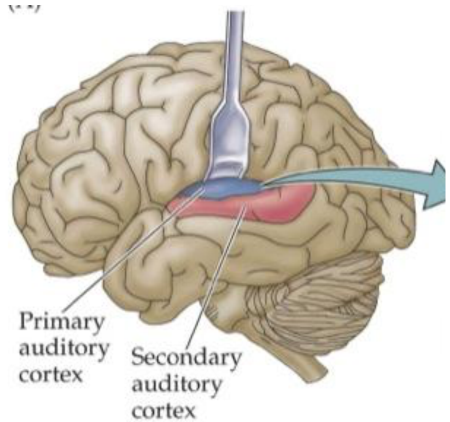


Saenz & Langers (2013) Tonotopic mapping of human auditory cortex. Hearing Research.

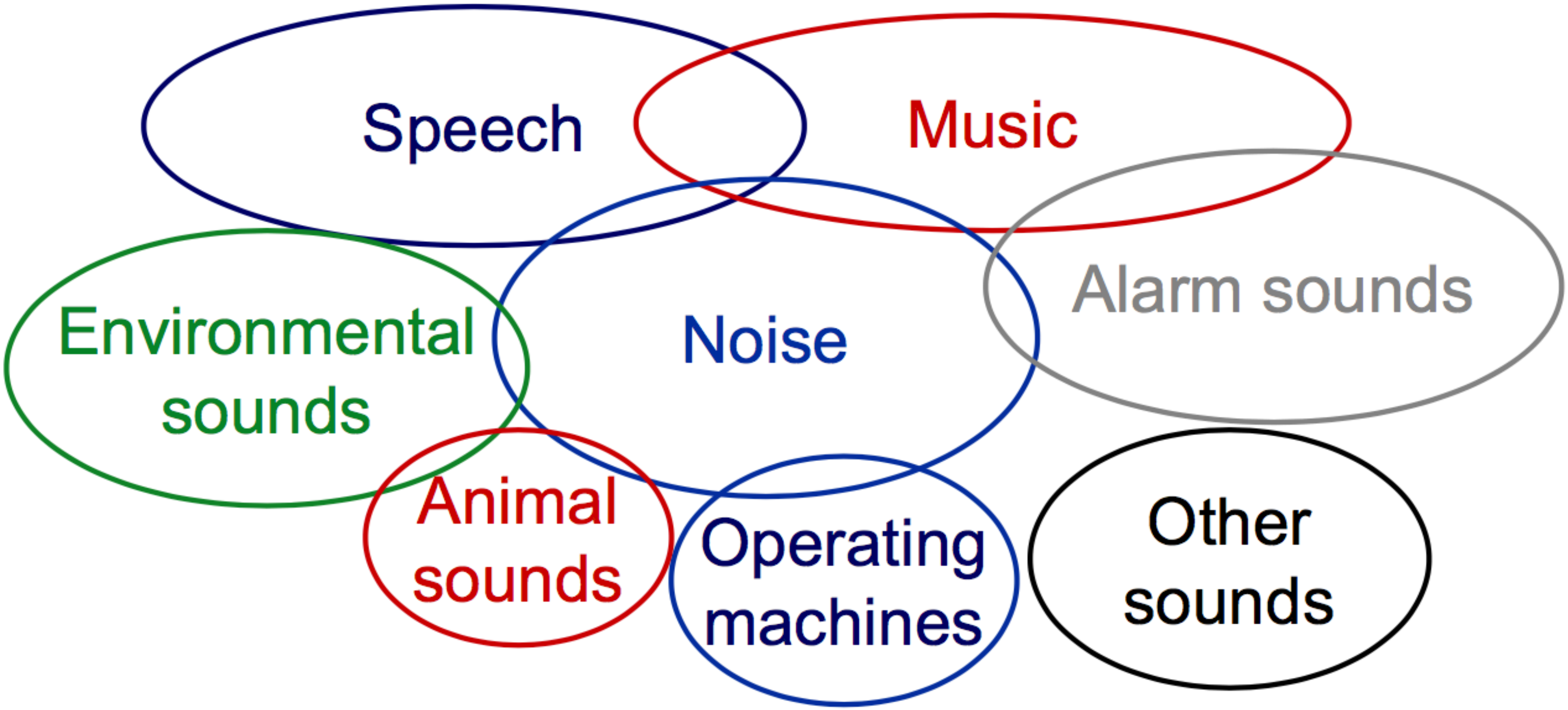
Auditory cortex

- **secondary auditory cortex**
  - BA 22 - Superior temporal gyrus
  - home to Wernicke's area (association)
  - left possesses greater temporal resolution
  - right associated with greater spatial resolution



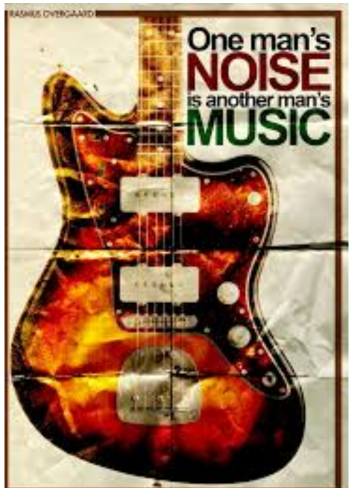


Types of sounds



Noise

- disturbing sound
- typically an environmental problem
- can cause unwanted physiological and psychological effects
- sometimes it is a matter of taste
  - E.g. an outdoor concert,



Noise Sensitivity

- Weinstein Noise Sensitivity Scale (1978)

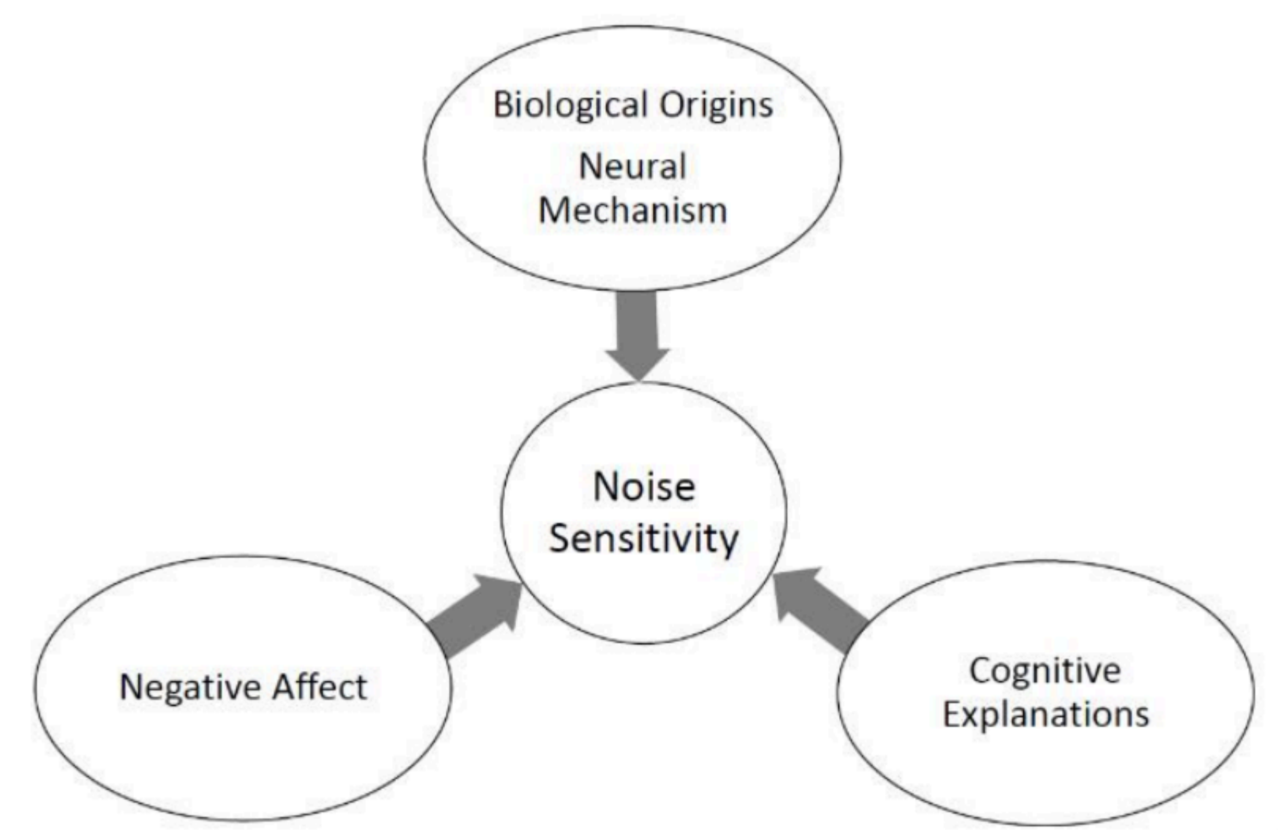
Table 2. Weinstein Noise Sensitivity Scale.

Instructions – Circle the number corresponding to how well you agree or disagree. Don't be disturbed by the reversals of order from one line to another. At the end, add up the numbers for your score.

1. I wouldn't mind living on a noisy street if the apartment I had was nice.	AGREE	1	2	3	4	5	6	DISAGREE
2. I am more aware of noise than I used to be.	AGREE	6	5	4	3	2	1	DISAGREE
3. No one should mind much if someone turns up his stereo full blast once in a while.	AGREE	1	2	3	4	5	6	DISAGREE.
4. At movies, whispering and crinkling candy wrappers disturb me.	AGREE	6	5	4	3	2	1	DISAGREE
5. I am easily awakened by noise.	AGREE	6	5	4	3	2	1	DISAGREE
6. If it's noisy where I'm studying, I try to close the door or window or move someplace else.	AGREE	6	5	4	3	2	1	DISAGREE
7. I get annoyed when my neighbors are noisy.	AGREE	6	5	4	3	2	1	DISAGREE
8. I get used to most noises without much difficulty.	AGREE	1	2	3	4	5	6	DISAGREE
9. How much would it matter to you if an apartment you were interested in renting was located across from a fire station.	A LOT	6	5	4	3	2	1	NOT MUCH
10. Sometimes noises get on my nerves and get me irritated.	AGREE	6	5	4	3	2	1	DISAGREE
11. Even music I normally like will bother me if I'm trying to concentrate.	AGREE	6	5	4	3	2	1	DISAGREE
12. It wouldn't bother me to hear the sounds of everyday living from my neighbors (footsteps, running water, etc).	AGREE	1	2	3	4	5	6	DISAGREE
13. When I want to be alone, it disturbs me to hear outside noises.	AGREE	6	5	4	3	2	1	DISAGREE
14. I'm good at concentrating no matter what is going on around me.	AGREE	1	2	3	4	5	6	DISAGREE
15. In a library, I don't mind if people carry on a conversation if they do it quietly.	AGREE	1	2	3	4	5	6	DISAGREE
16. There are often times when I want complete silence.	AGREE	6	5	4	3	2	1	DISAGREE
17. Motorcycles ought to be required to have bigger mufflers.	AGREE	6	5	4	3	2	1	DISAGREE
18. I find it hard to relax in a place that's noisy.	AGREE	6	5	4	3	2	1	DISAGREE
19. I get mad at people who make noise that keeps me from falling asleep or getting work done.	AGREE	6	5	4	3	2	1	DISAGREE
20. I wouldn't mind living in an apartment with thin walls.	AGREE	1	2	3	4	5	6	DISAGREE
21. I am sensitive to noise.	AGREE	6	5	4	3	2	1	DISAGREE

TOTAL SCORE ___________

Noise Sensitivity



Heinonen-Guzejev et al. (2018) Studying the origins of noise sensitivity–negative affect or biological factors

Noise Sensitivity

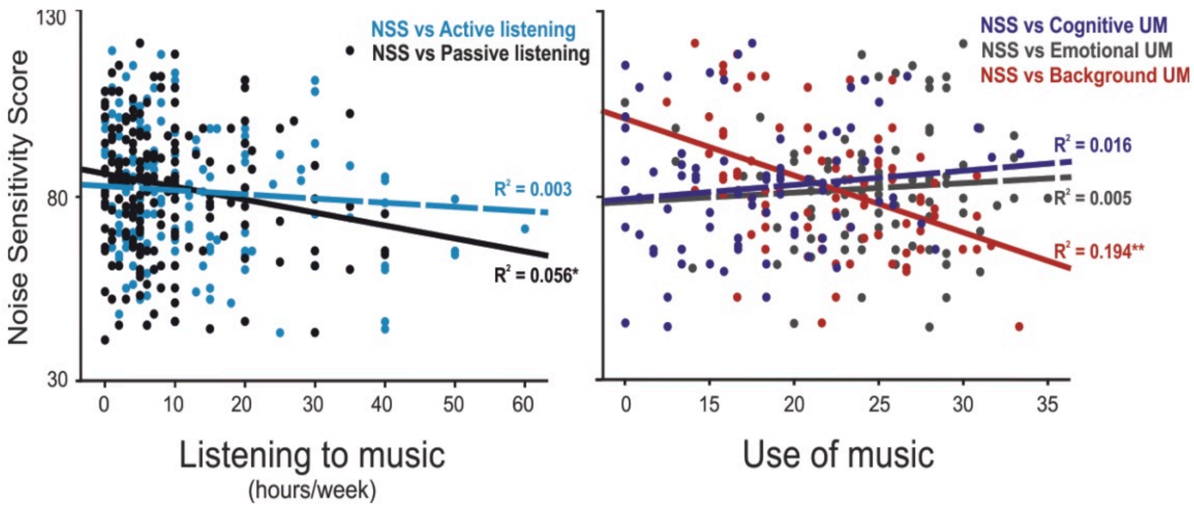
- **Negative Affect:** NS a dispositional tendency to negatively evaluate situations and the self
  - people with NS might also have heightened sensitivity to other stimuli (e.g., smells, bright lights)
  - however, Shepherd et al. (2015) failed to support this notion
  - instead - “noise vulnerability hypothesis”, suggests that noise impacts sensitive individuals the most, rather than negative affect being the cause
- **Biological origins:**
  - NS has a potential heritability factor (36%)
  - NS only to noise and not to other sound features such as pitch, location or intensity
  - MRI studies (Kliuchko et al., 2017) found NS correlates with larger auditory cortex, hippocampus, and insula (regions are involved in sound perception, emotional processing, and sensory awareness)
  - appears to be a more acceptable theory for NS

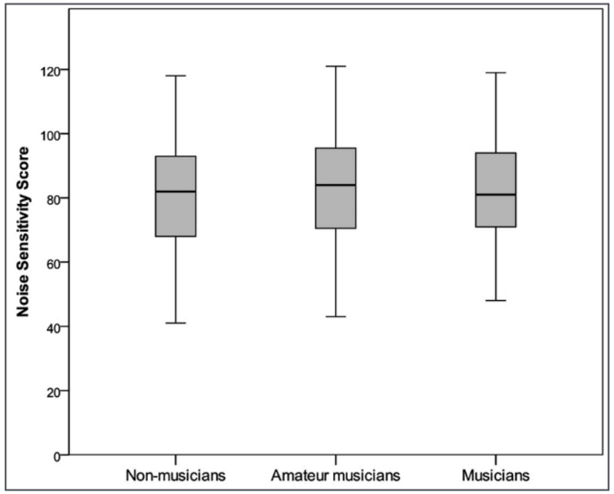
Noise Sensitivity

- **Cognitive Explanations:**
  - NS exerts a negative effect on cognitive functions, such as attention, working memory and episodic recall
  - noise-induced memory and attentional deficits lead to annoyance or distress
  - NS individuals may have less cognitive control over filtering out irrelevant noise, leading to higher distraction and frustration
  - Not all NS individuals react negatively due to emotional traits—many experience real cognitive disruptions

Noise

- NS moderates how and why individuals listen to music





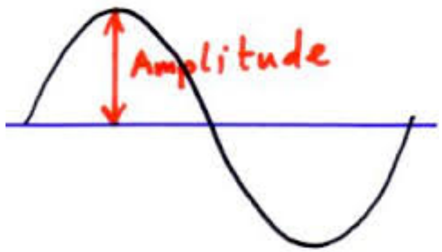
while NS individuals seem to be able to enjoy the sound of music, use it for mood regulation, and they attentively (actively) listen to music like non-sensitive individuals, they prefer not having it in the background.

Noise

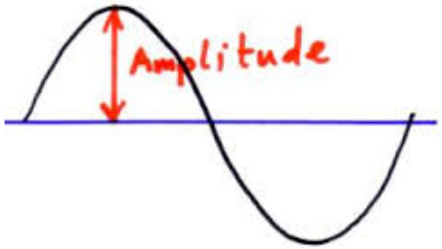
- higher NS
  - rated music as less important in their life than did individuals with lower sensitivity to noise
  - not associated with musical training
  - spend less time in passive (background) listening to music than those with lower sensitivity to noise

Thesis: https://helda.helsinki.fi/bitstream/handle/10138/217878/NOISESEN.pdf?sequence=1

Physical Perceptual



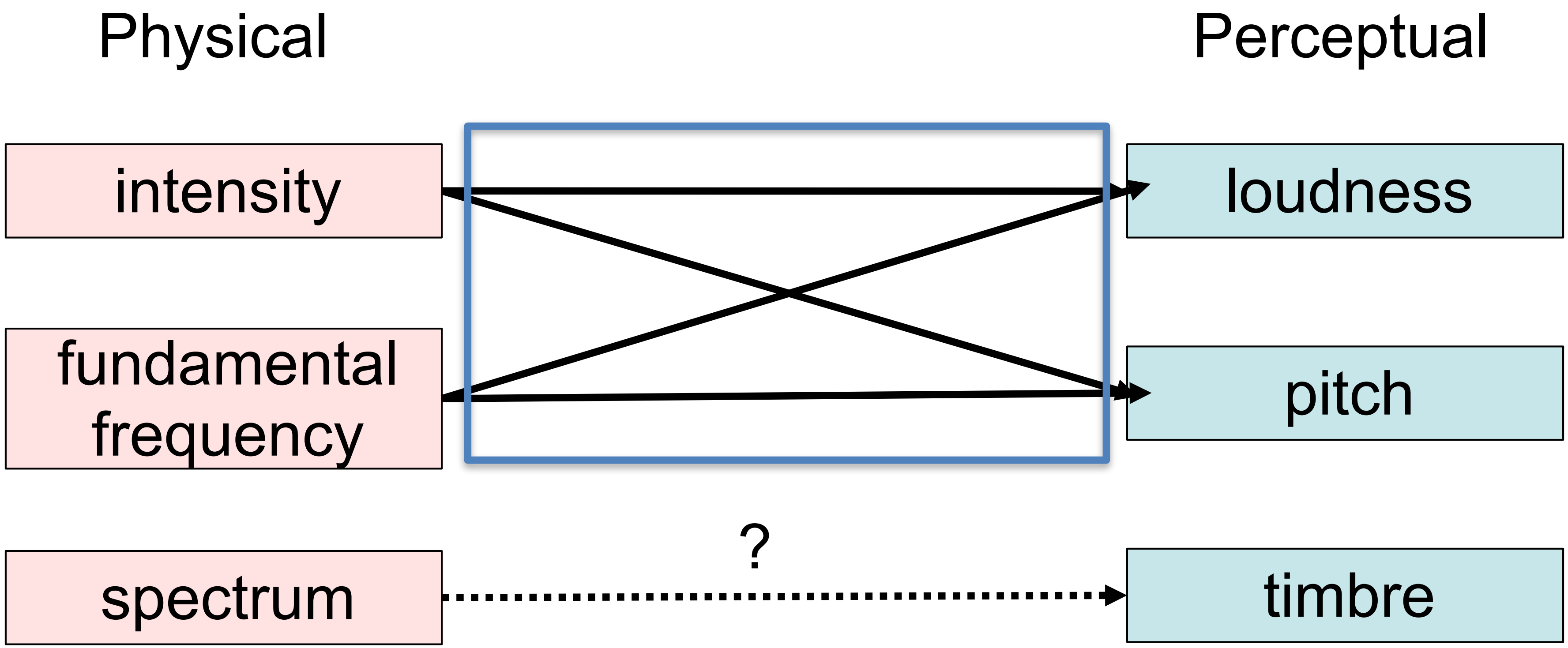
Physical

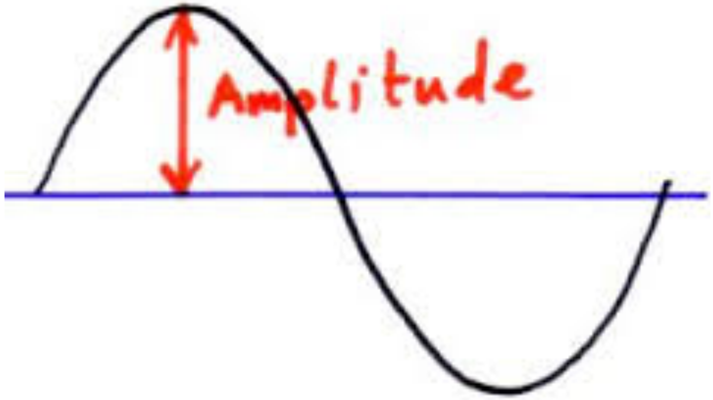


?

Perceptual

Fundamental Question of Psychoacoustics



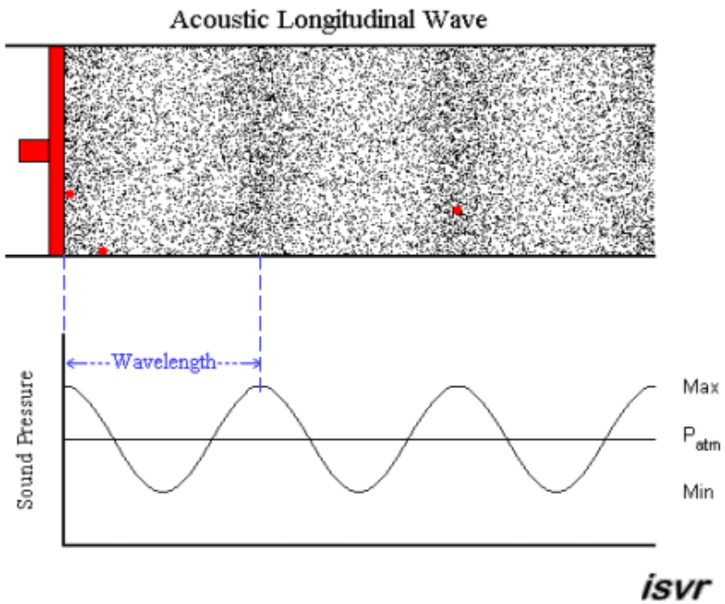


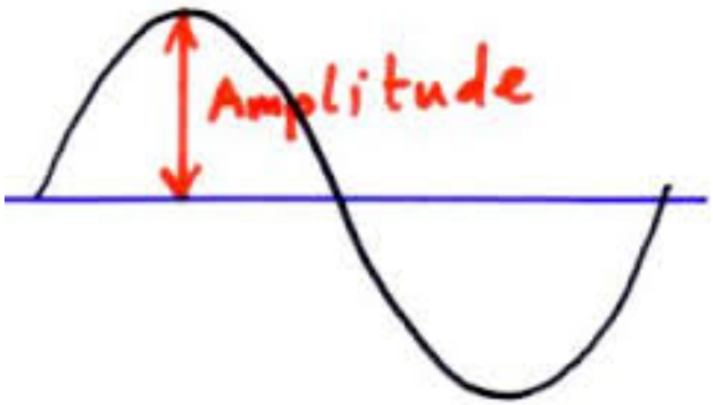
amplitude - the objective measurement of the degree of change in atmospheric pressure

- directly related to the acoustic energy or intensity of a sound

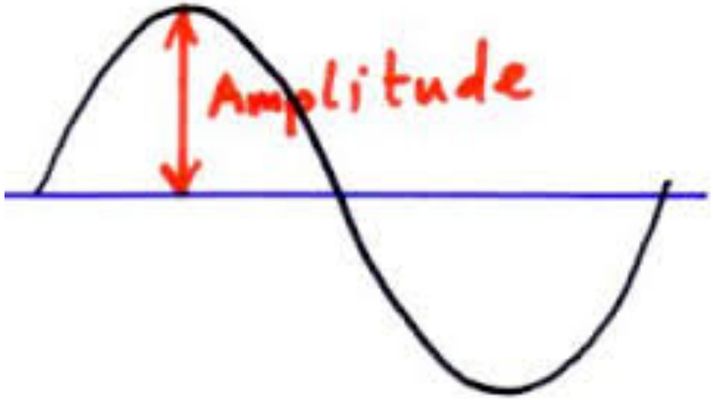
loudness - psycho-physiological correlate of amplitude

- physical measure of sound strength typically sound pressure, sound pressure level (in decibels), sound intensity or sound power



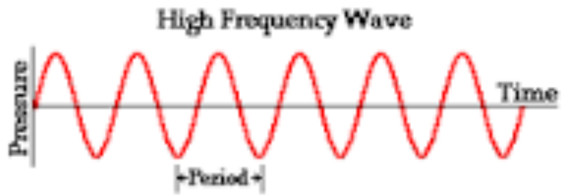


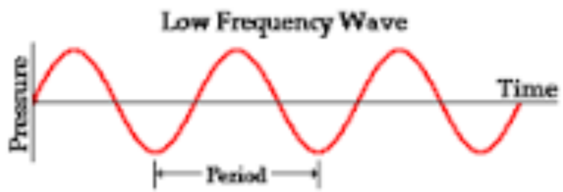
- ear can respond to various levels of sound intensities
  - Ratio of the smallest and largest air pressure = 10,00,000 : 1 = 10^6 : 1
- **decibel** - unit used to measure the intensity of a sound (logarithmic) (130 dB and 0dB)



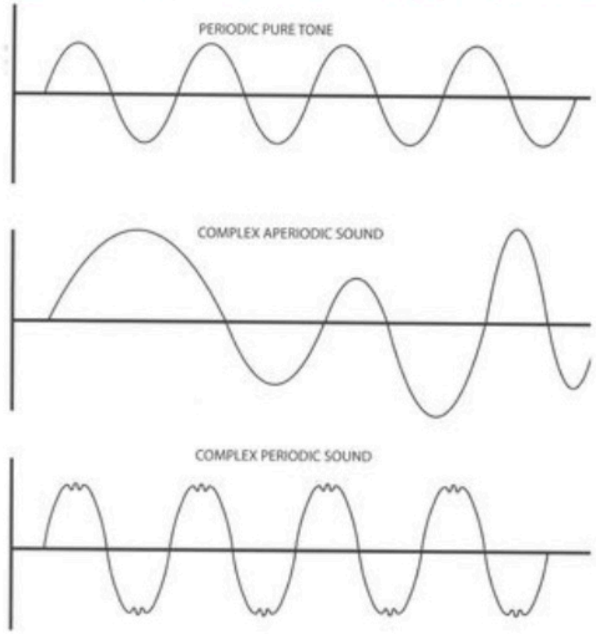
Sound sources (noise) Examples with distance	Sound pressure Level L_p dB SPL
Jet aircraft, 50 m away	**140**
Threshold of pain	**130**
Threshold of discomfort	**120**
Chainsaw, 1 m distance	**110**
Disco, 1 m from speaker	**100**
Diesel truck, 10 m away	**90**
Kerbside of busy road, 5 m	**80**
Vacuum cleaner, distance 1 m	**70**
Conversational speech, 1 m	**60**
Average home	**50**
Quiet library	**40**
Quiet bedroom at night	**30**
Background in TV studio	**20**
Rustling leaves in the distance	**10**
Hearing threshold	**0**

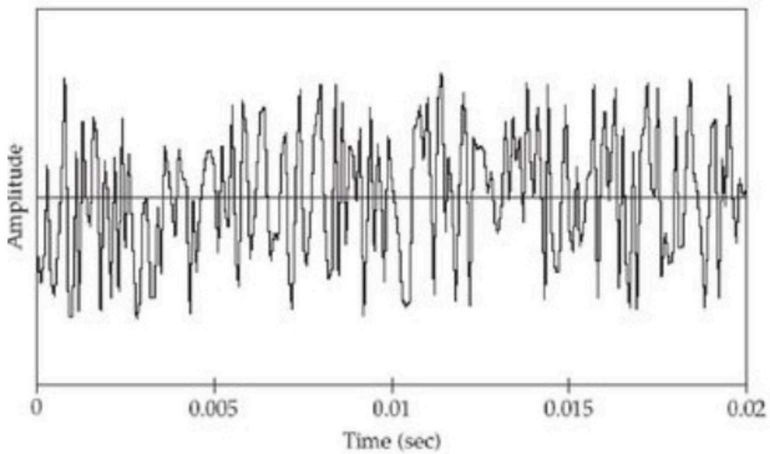
- **frequency** - number of sound wave cycles per second
  - measured in hertz (Hz)
  - wide audible frequency range: 20 — 20 000 Hz
  - varies according to age and exposure to noise
- **pitch** - psycho-physiological correlate of frequency

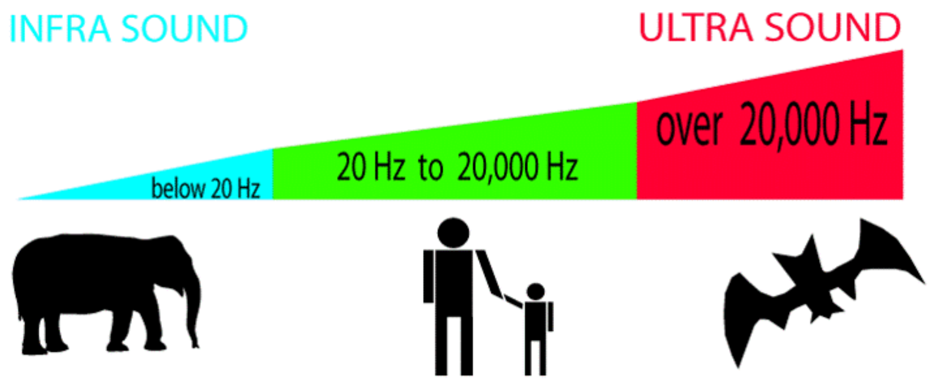




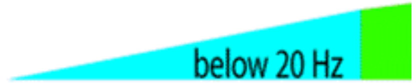

Frequency


PERIODIC PURE TONE
COMPLEX APERIODIC SOUND
COMPLEX PERIODIC SOUND


Amplitude
0
0.005
0.01
0.015
0.02
Time (sec)


INFRA SOUND
ULTRA SOUND
below 20 Hz
20 Hz to 20,000 Hz
over 20,000 Hz

INFRA SOUND



- can cause various unwanted physiological and psychological reactions as well as hearing damage at very high levels
- caused for example by machines or structural vibrations

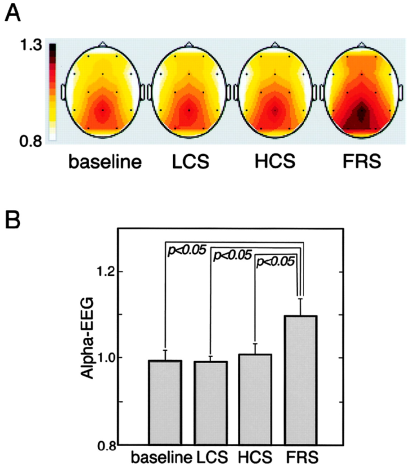
Frequency

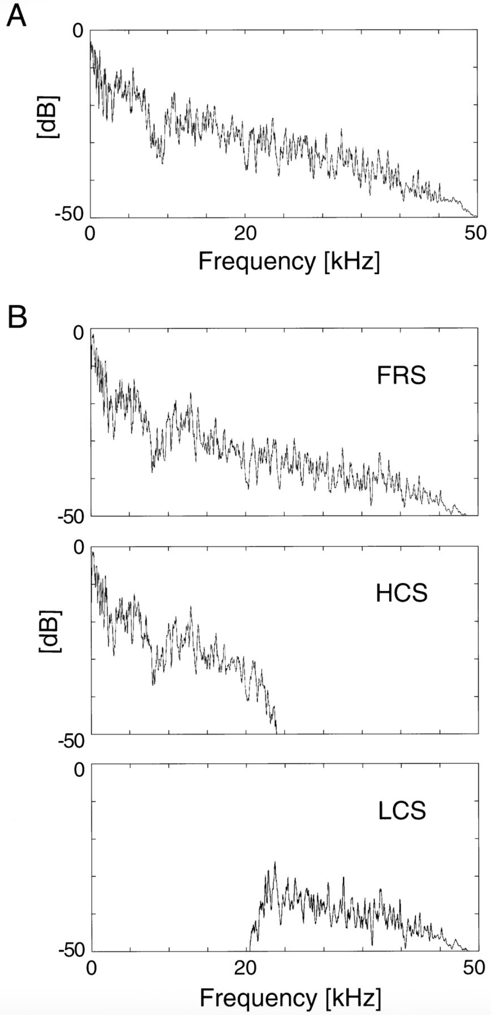
Inaudible High-Frequency Sounds Affect Brain Activity: Hypersonic Effect

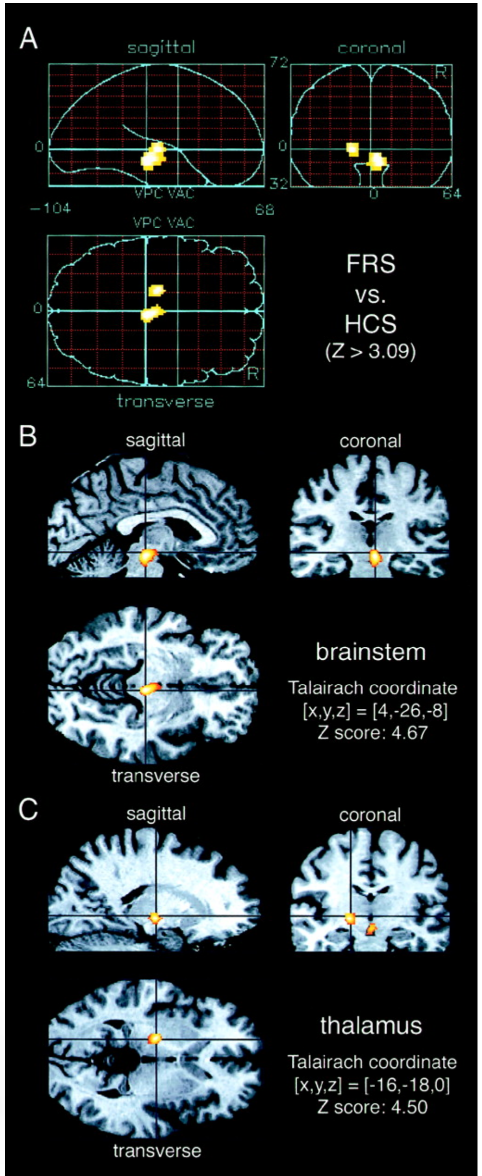
Oohashi et al. (2000). Journal of Neurophysiology, 83 (6) 3548-3558;

ULTRA SOUND

over 20,000 Hz







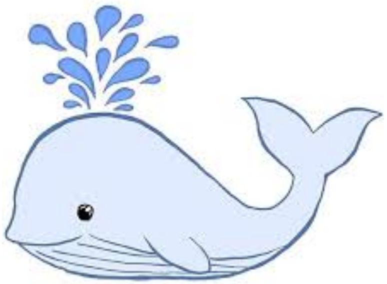
sound containing HFC to be more pleasant than the same sound lacking an HFC

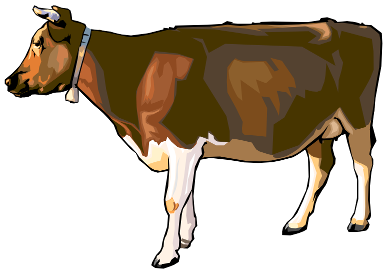
pitch

- Pitch is that auditory attribute of sound according to which sounds can be ordered on a scale from low to high

- The property of a sound and especially a musical tone that is determined by the frequency of the waves producing it: highness or lowness of sound

pitch

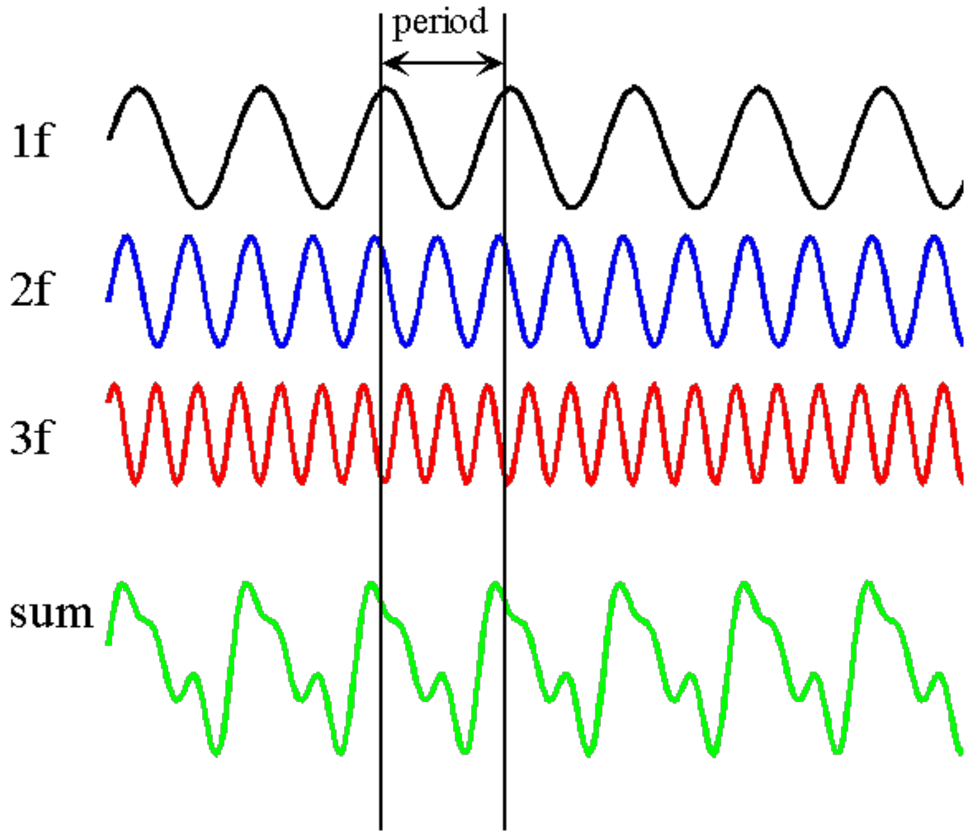




low

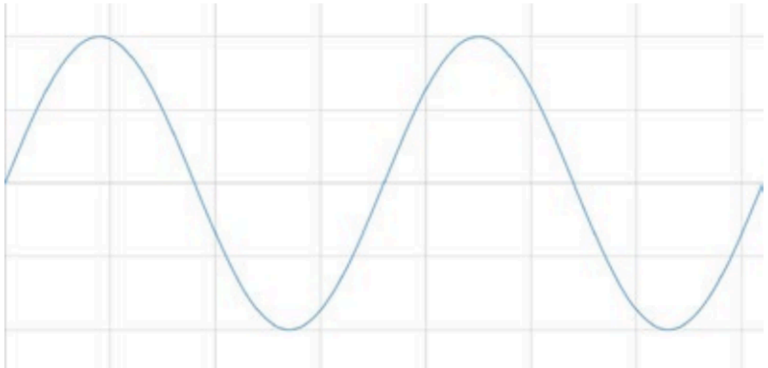
high

Klofstad, C.A., Anderson, R.C., Peters, S., 2012.**Sounds like a winner**: voice pitch influences perception of leadership capacity in both men and women. Proc. R. Soc. B: Biol.Sci. 279, 2698–2704.



MATLAB: ADDITIVE SYNTHESIS

How does a sine wave sound?

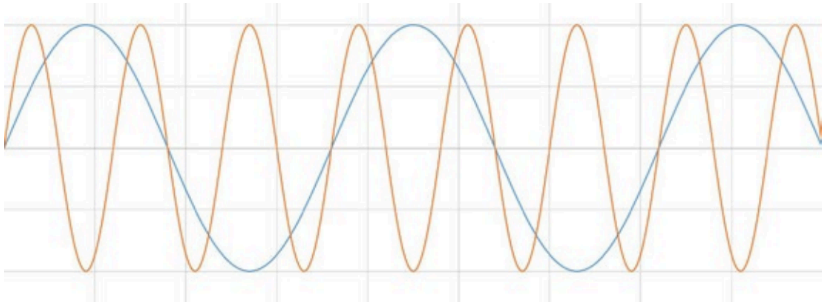


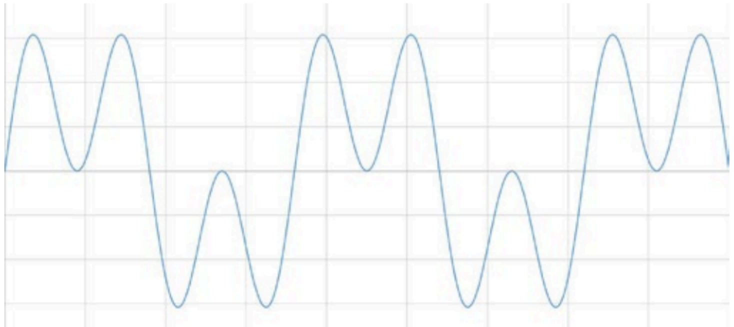
150 Hz

150 Hz

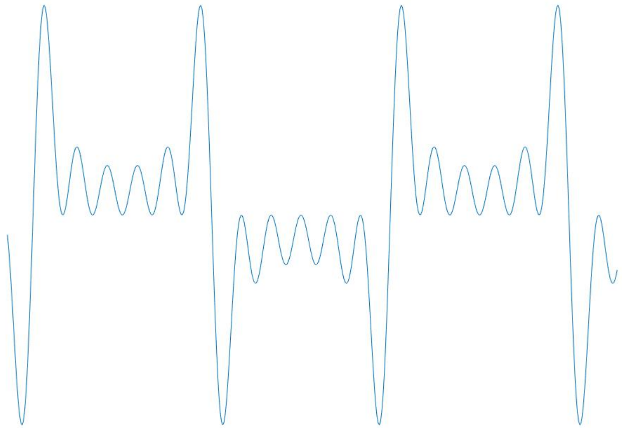
Sine wave + **Odd** Harmonics

150 Hz + 450 Hz

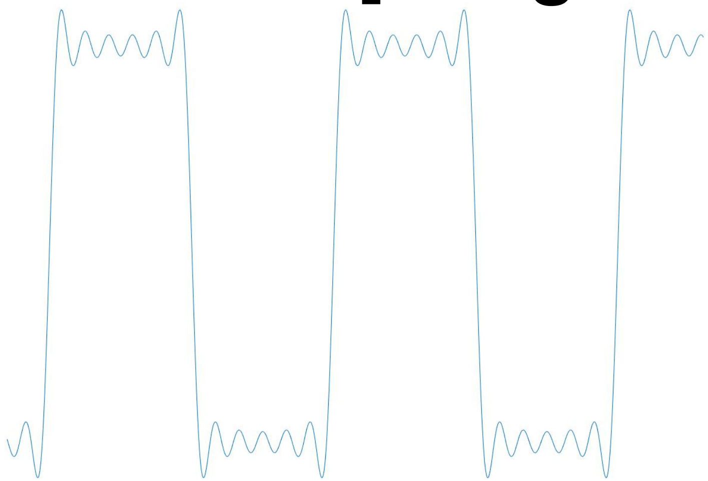




[1 3 5 7 9 11]

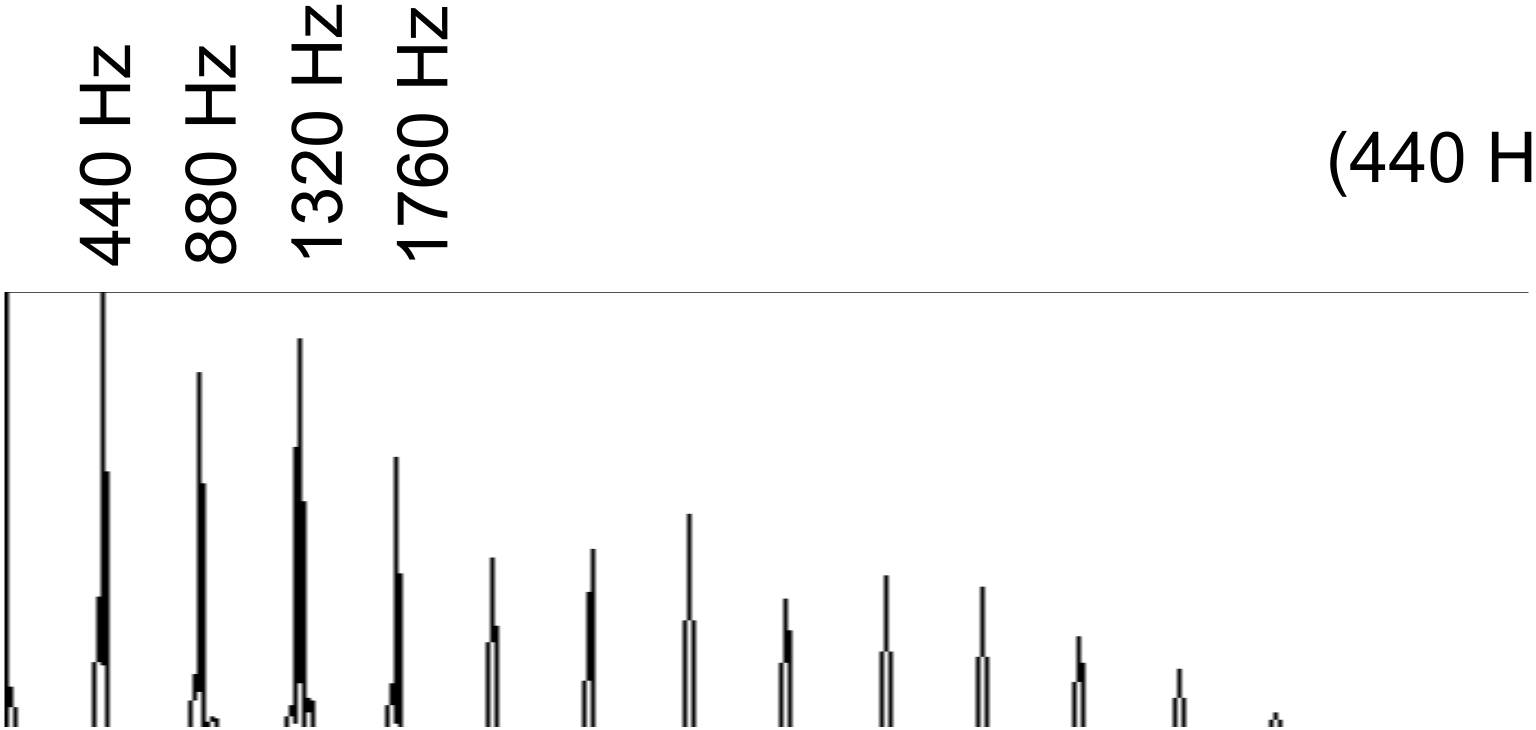


[1 3 5 7 9 11]-weighted



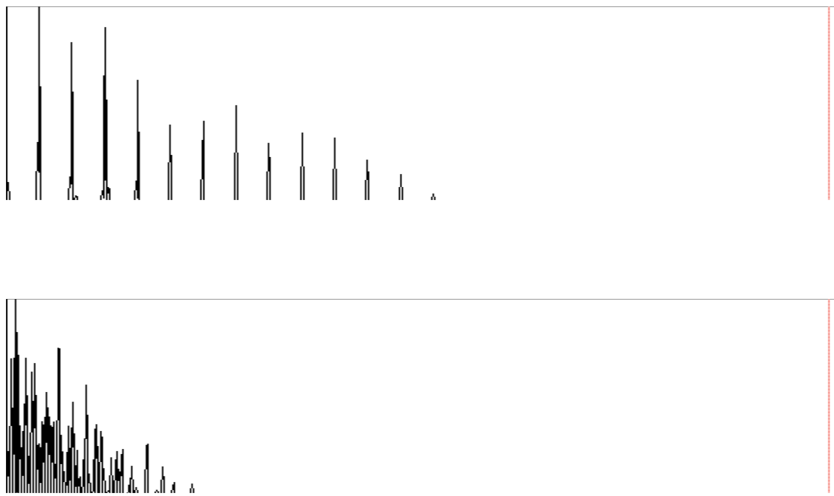
Pitch of periodic tones

≈ pitch of the sine tone with frequency equal to the frequency of the first partial

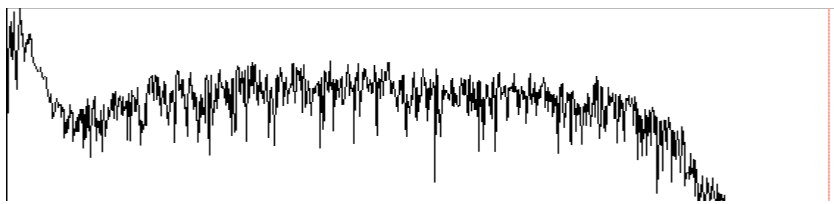


(440 Hz sine tone)

Which one sounds higher?



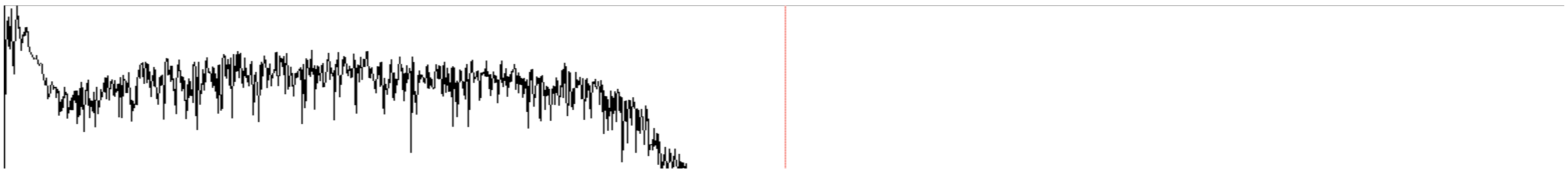
periodic tones
(harmonic spectrum)



aperiodic tones
(non-harmonic spectrum)

Pitch of aperiodic tones

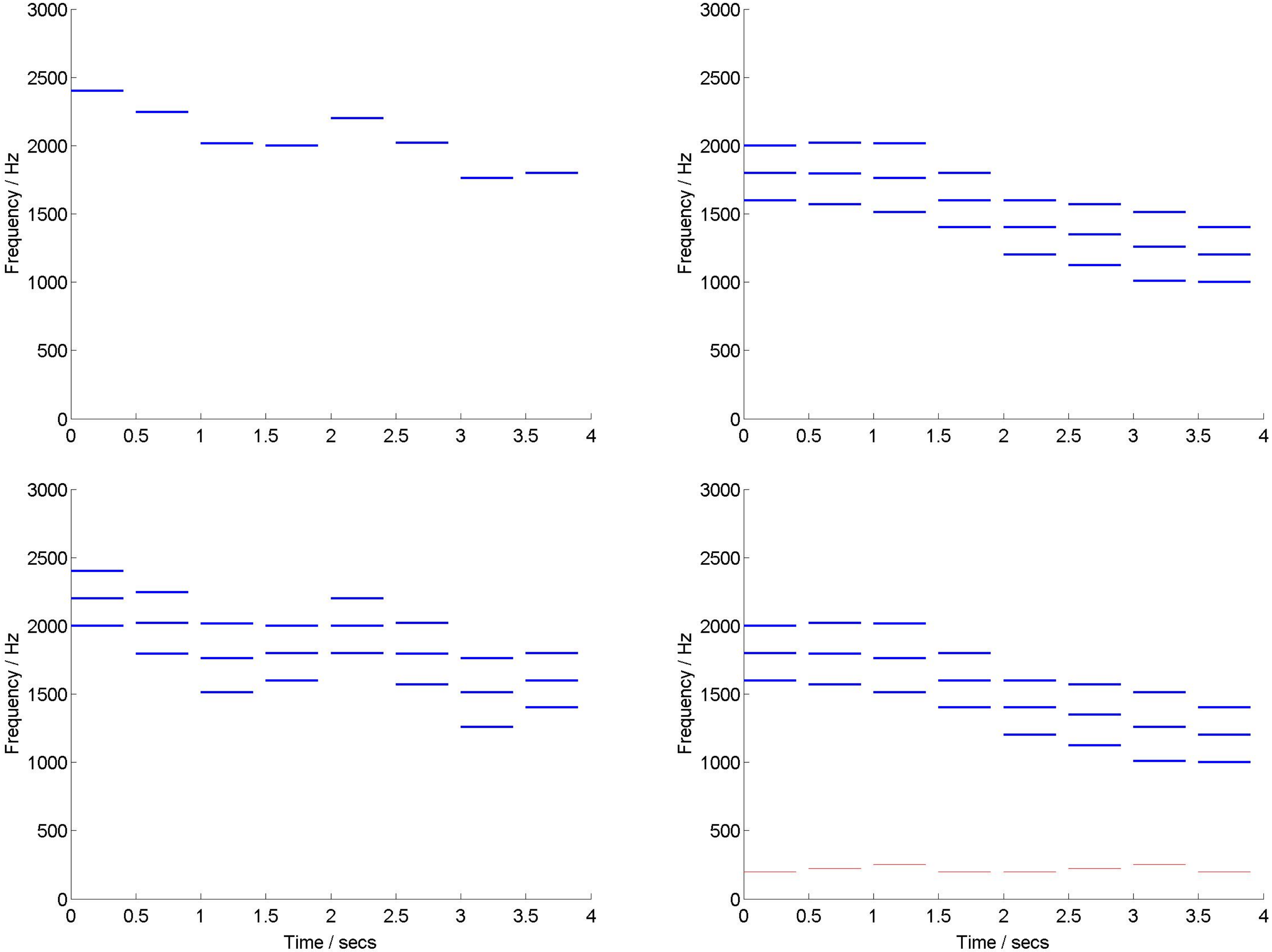
- difficult to define generally
- pitch salience
  - depends on the degree of harmonicity of the tone (degree of coincidence of subharmonics of partials) (Terhardt, E., Stoll, G., Seewann, M.1982)



What would be the fundamental pitch?

MATLAB DEMO

Do you hear a familiar melody?

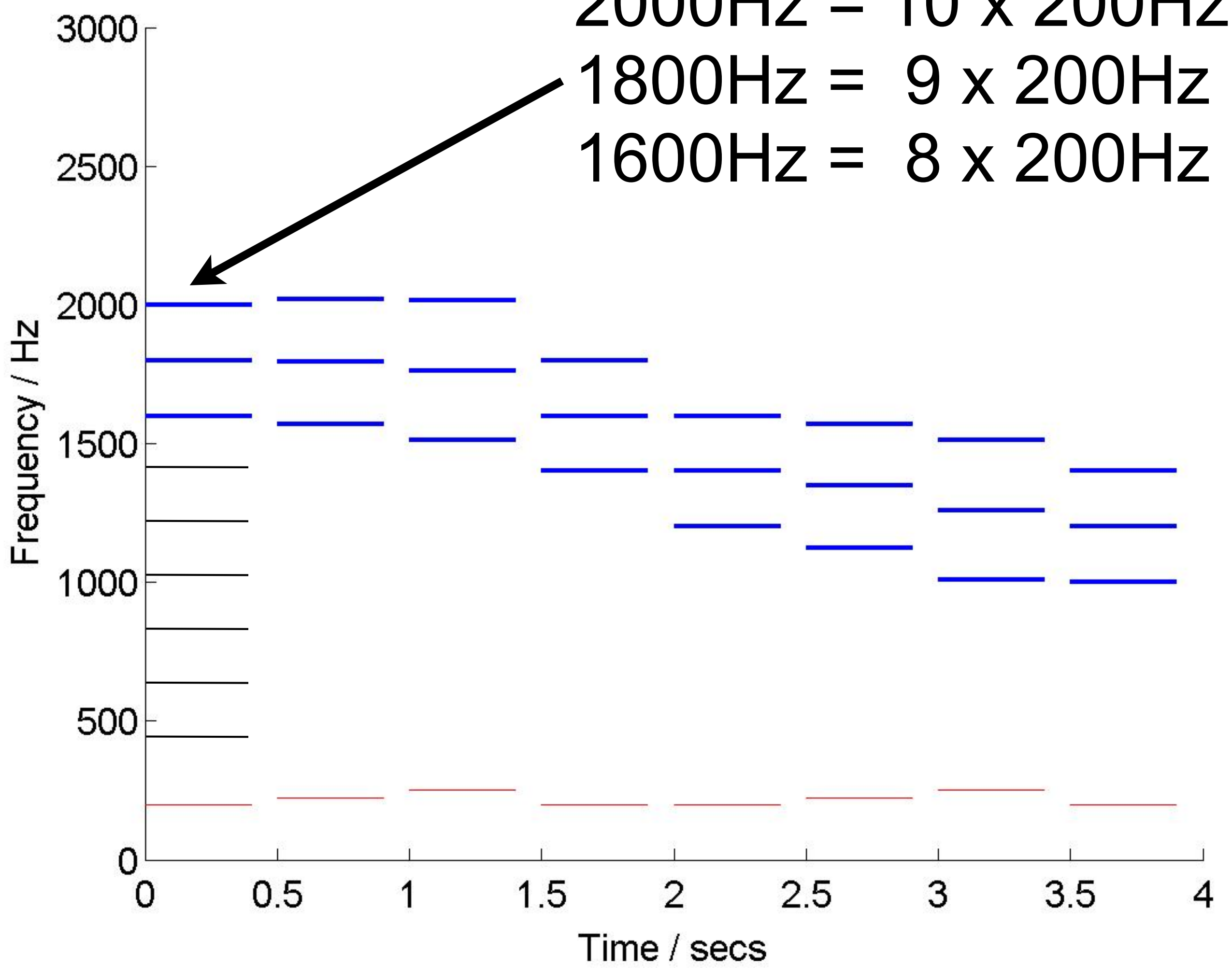


Two kinds of pitch

2000Hz = 10 x 200Hz
1800Hz = 9 x 200Hz
1600Hz = 8 x 200Hz

spectral pitch

virtual pitch



200 225 252 200 Hz

Spectral vs. virtual pitch

- spectral pitch
  - corresponds to frequencies present in the tone
- virtual pitch
  - corresponds to frequencies not necessarily present in the tone

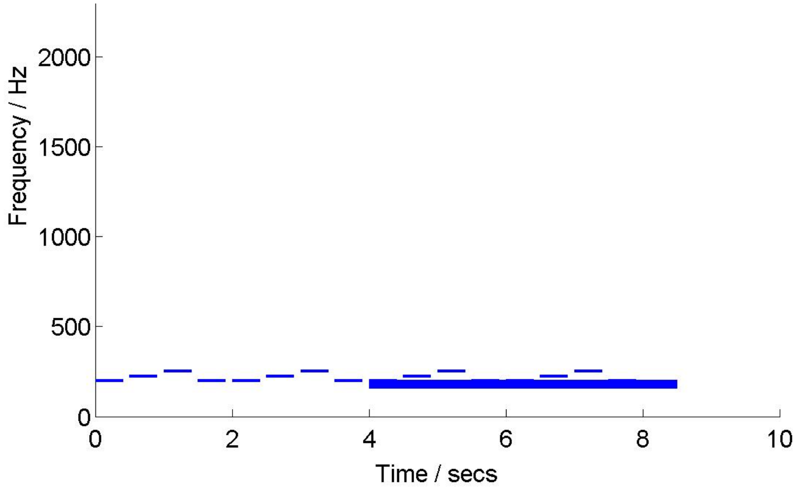
Virtual pitch in everyday life

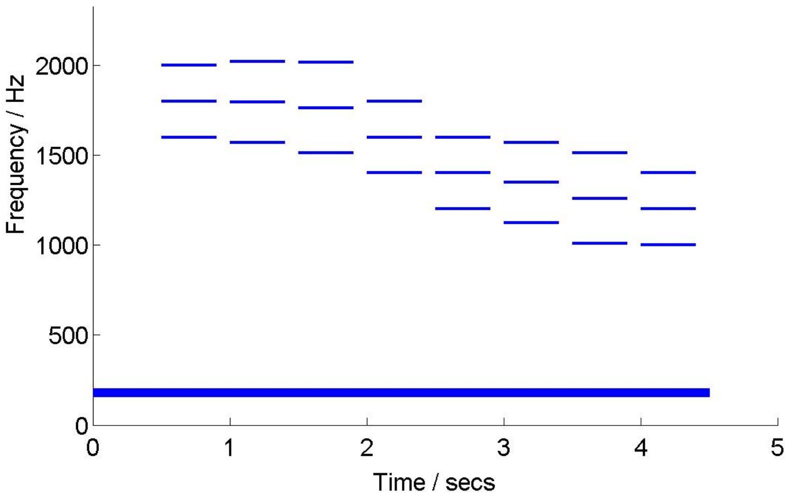
- mobile phone
  - bandwidth 500-3000 Hz
  - male voice 100-200 Hz
- small multimedia speakers
  - music 50- Hz

Explanation of virtual pitch

- Rate theory of pitch perception (volley theory)
    - pitch is encoded in the periodicity pattern of neural firings

Evidence for rate/temporal coding





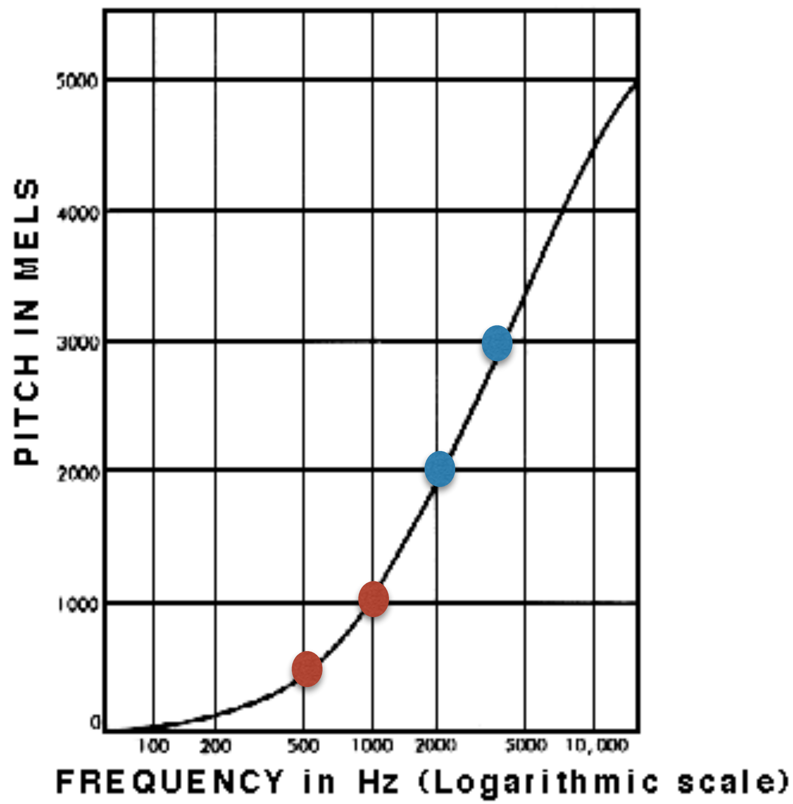
- virtual pitch is not masked
- thus: it is perceived in cochlea at a location different from that of corresponding spectral pitch

What is twice as high?

MATLAB DEMO

What is twice as high?

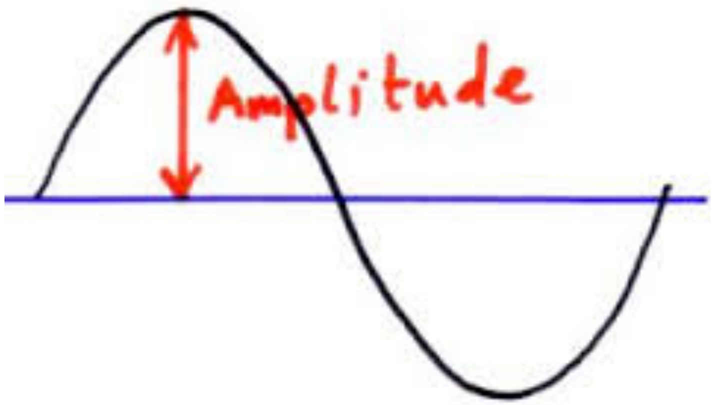
- **mel scale** is a scale of pitches judged by listeners to be equal in distance one from another
  - 1000 Hz, 40 dB sine tone = 1000 mels
- at high frequencies, larger and larger intervals are needed to produce equal pitch increments



$$m = 2595 \log_{10}\left(\frac{f}{700} + 1\right) = 1127 \log_e\left(\frac{f}{700} + 1\right)$$

Summarizing Pitch

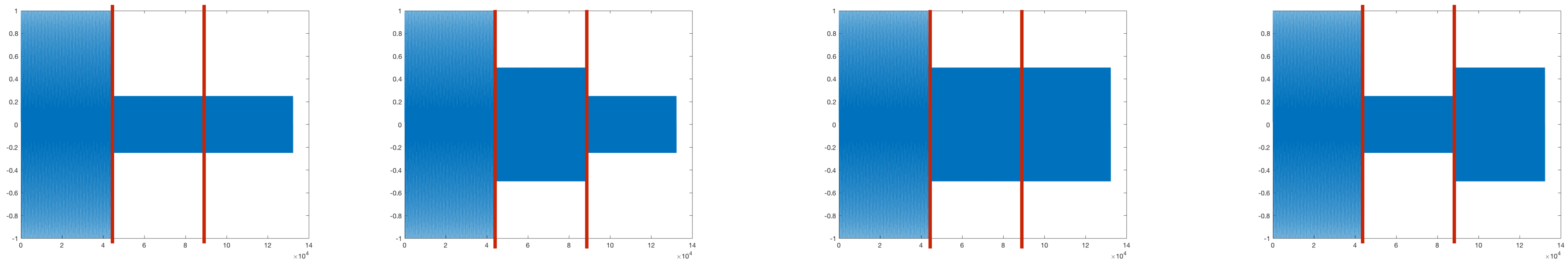
- pitch of harmonic tones ≈ fundamental frequency
- pitch depends slightly on intensity
- two kinds of pitch
  - spectral
  - virtual
- pitch perception occurs as a combination of
  - place coding: location of maximal oscillation along basilar membrane
  - frequency coding: periodicity information encoded in neural firing patterns of groups of neurons
  - volley theory: periodicity information encoded in neural firing clustered at certain places



interaction

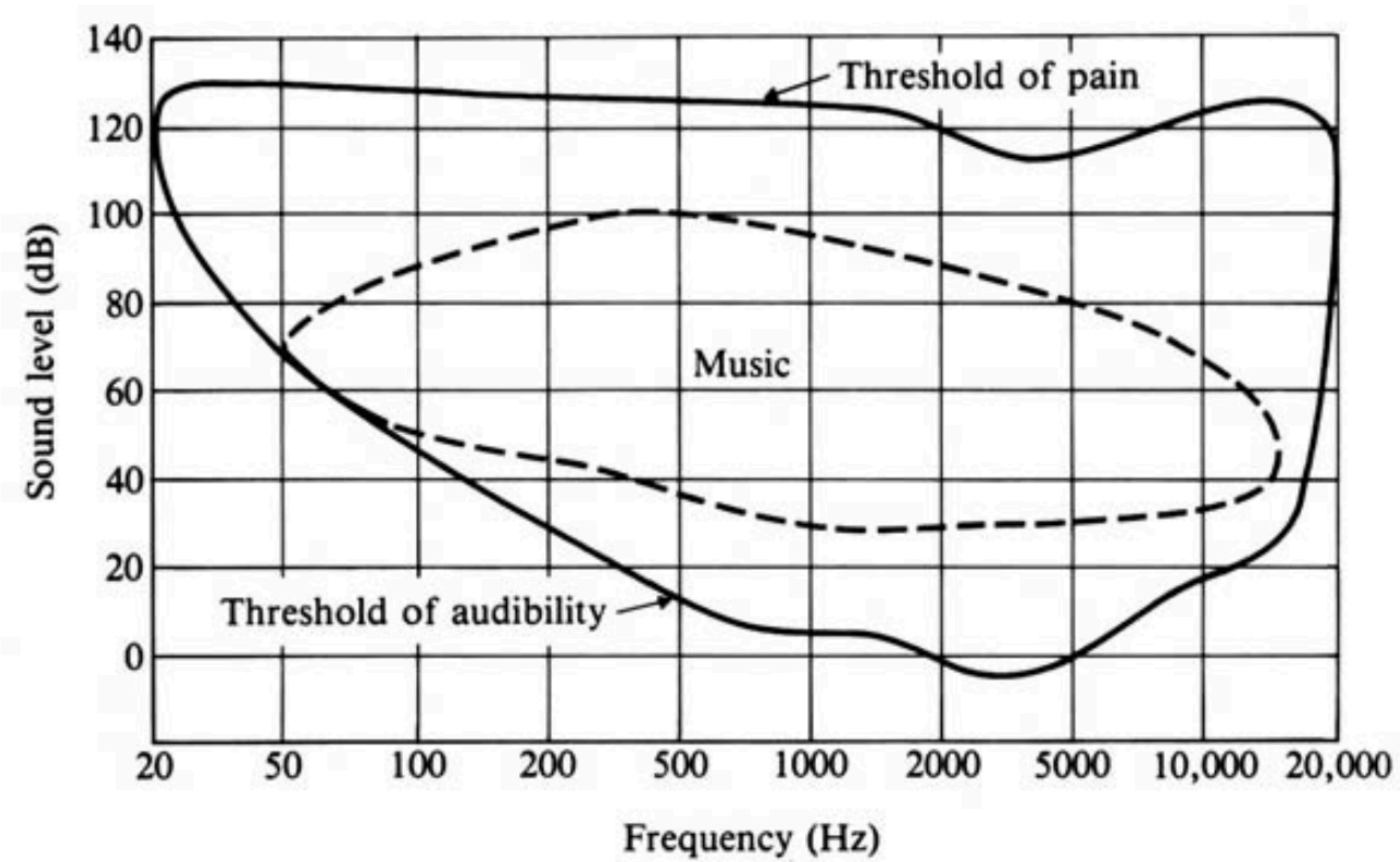
150 Hz **1000 Hz** **5000 Hz**

Which sine wave sounds louder?



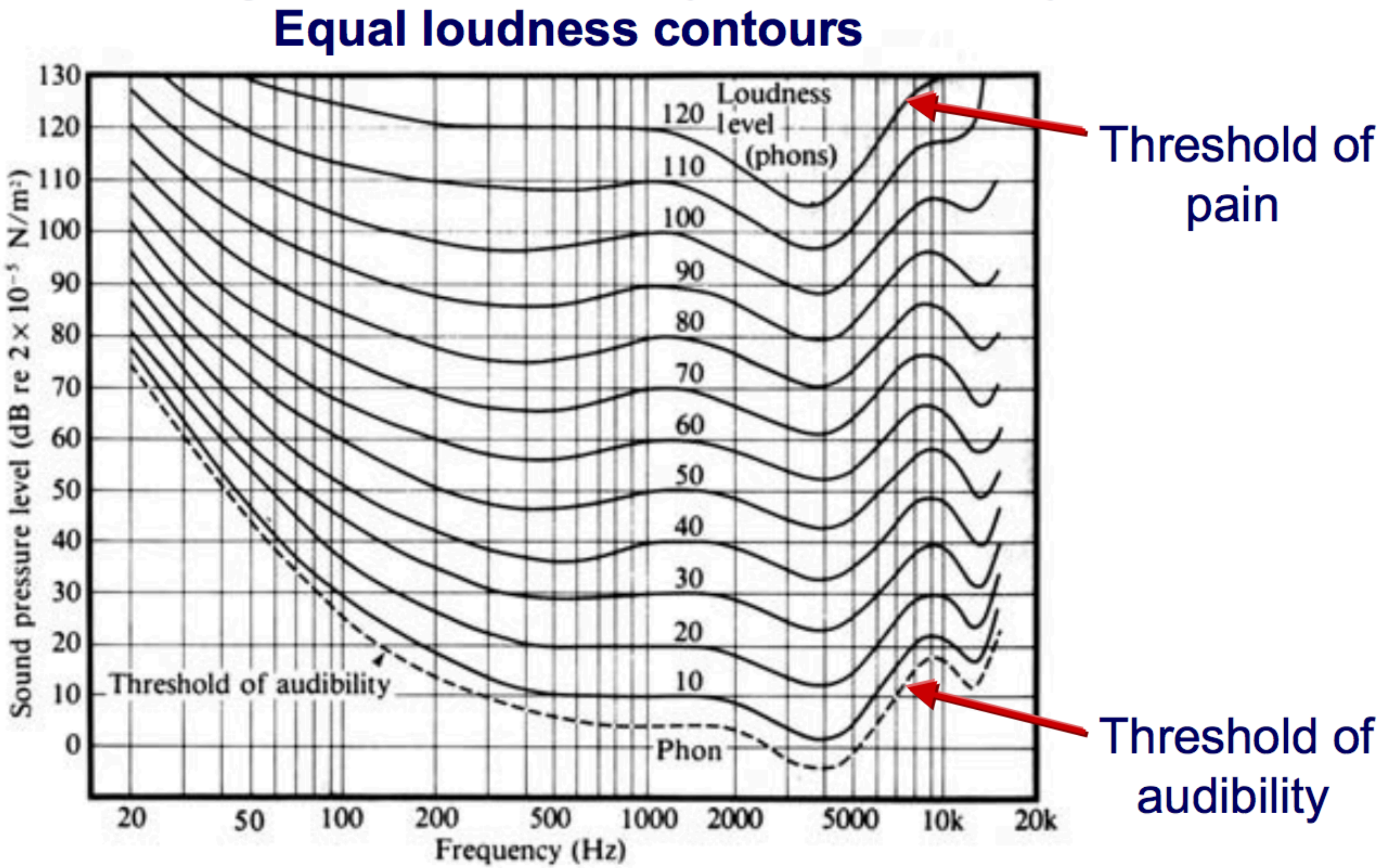
Hearing Sensitivity

- sounds targeted to humans go well with the properties of hearing



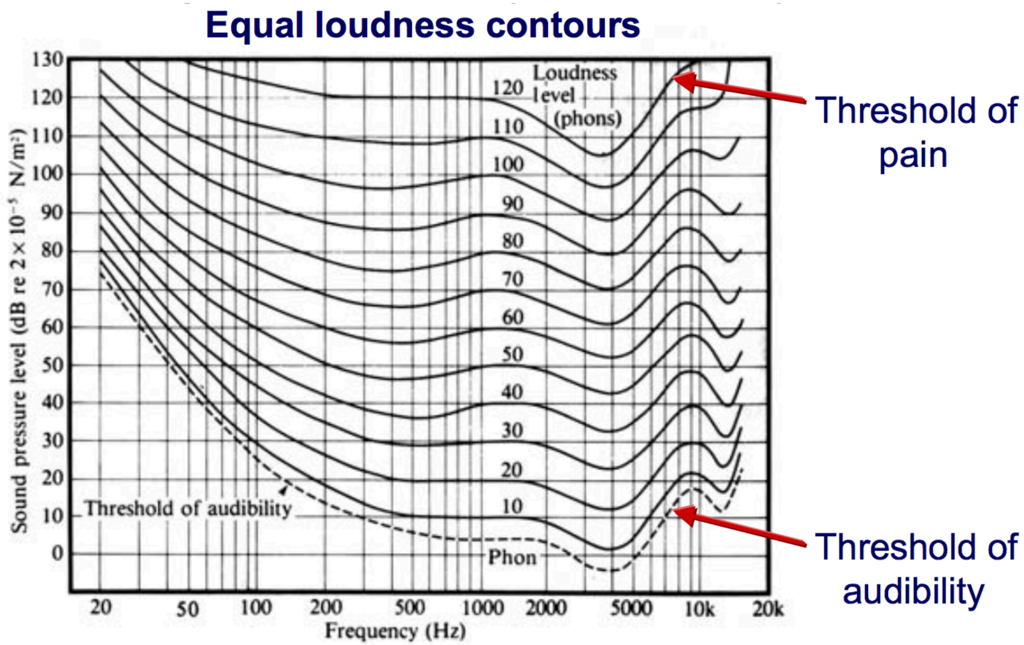
Hearing Sensitivity

- sensitivity of hearing depends heavily in frequency
- contours are based on psychoacoustic studies using sine tones



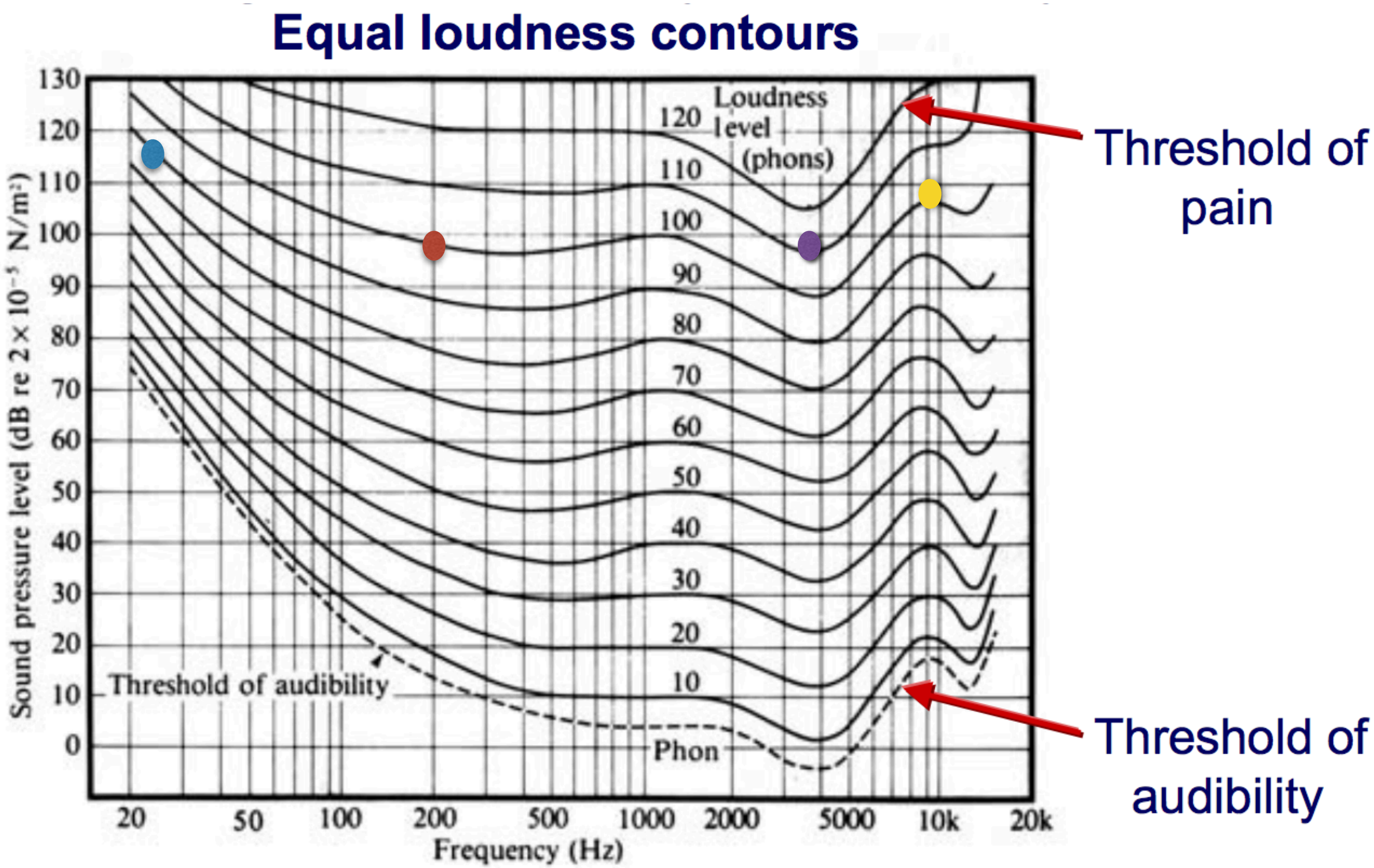
Hearing Sensitivity

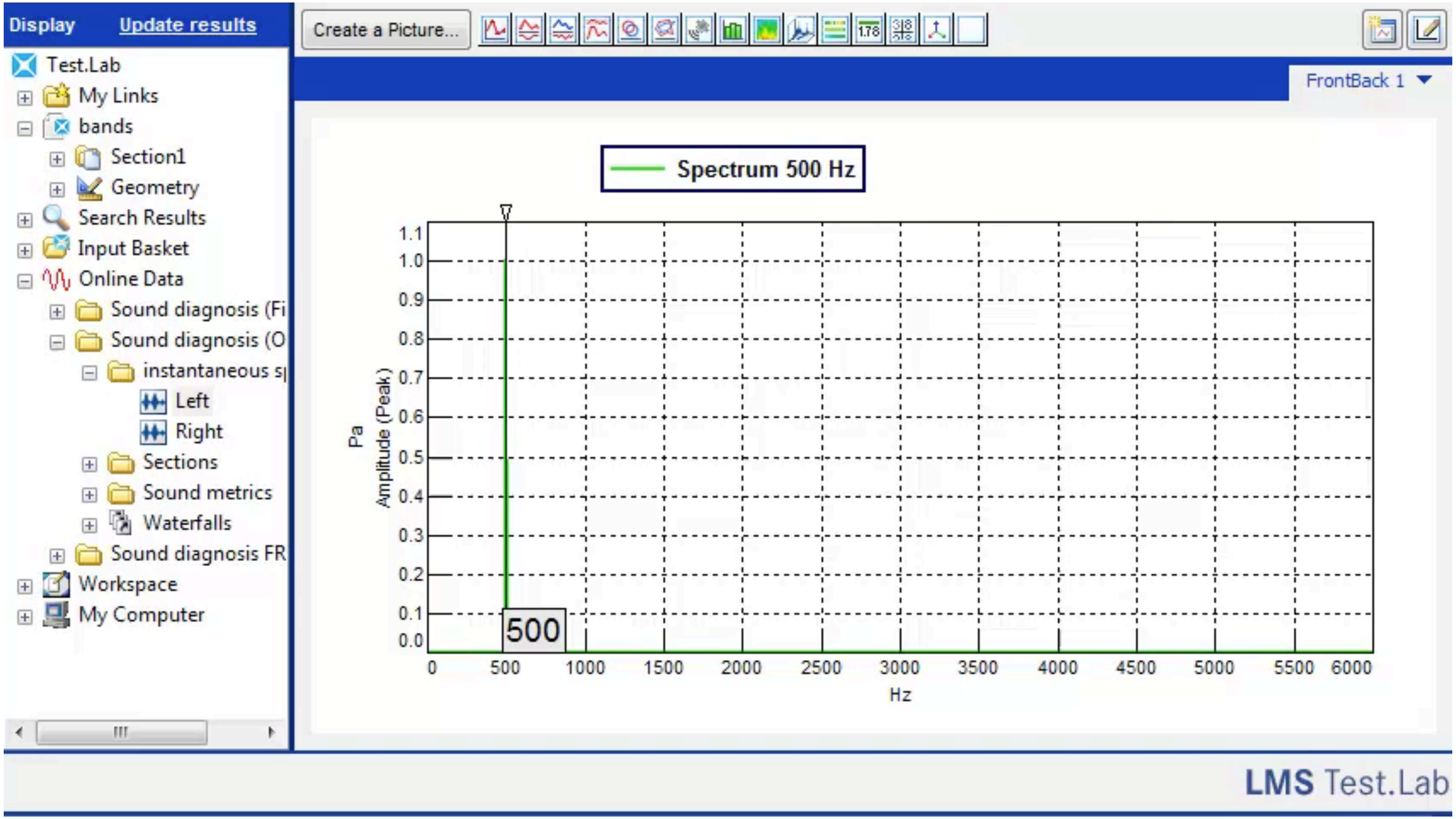
- ear most sensitive in the 1kHz-5kHz range
- absolute threshold of human hearing determined by human testing (describes energy in a pure/sine tone needed for audibility in a noiseless environment)
- phons = units of perceptual amplitude



Hearing Sensitivity

- which tone would be the loudest?

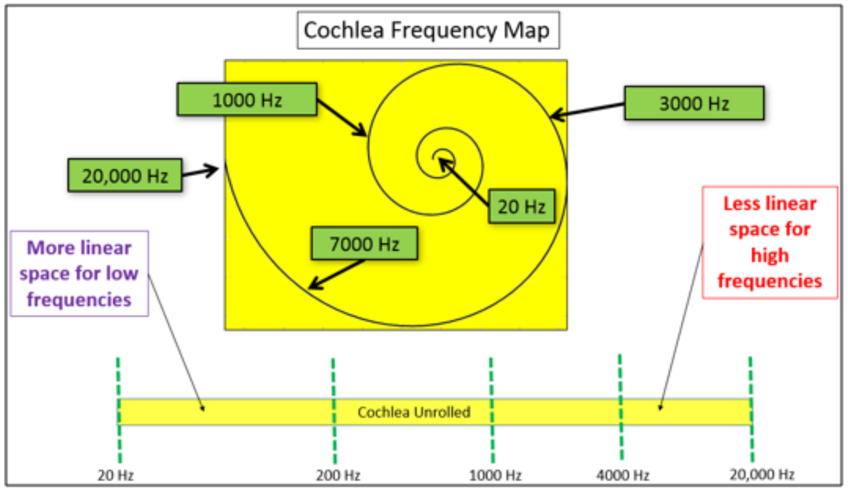


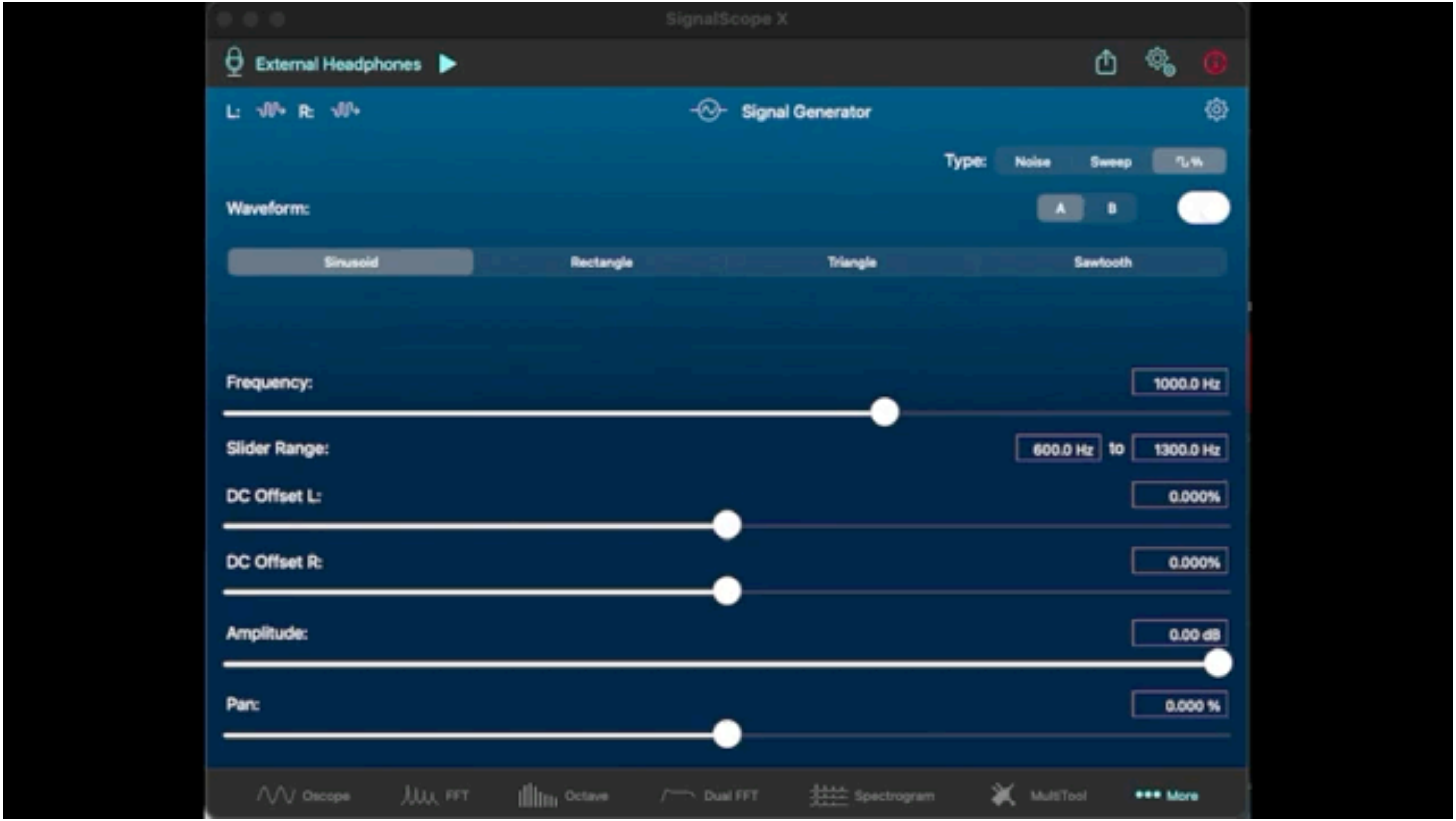


- ex: 20 Hertz difference between 500 and 520 Hertz tones more readily than a 5000 and 5020 Hertz tones.

Critical Bands

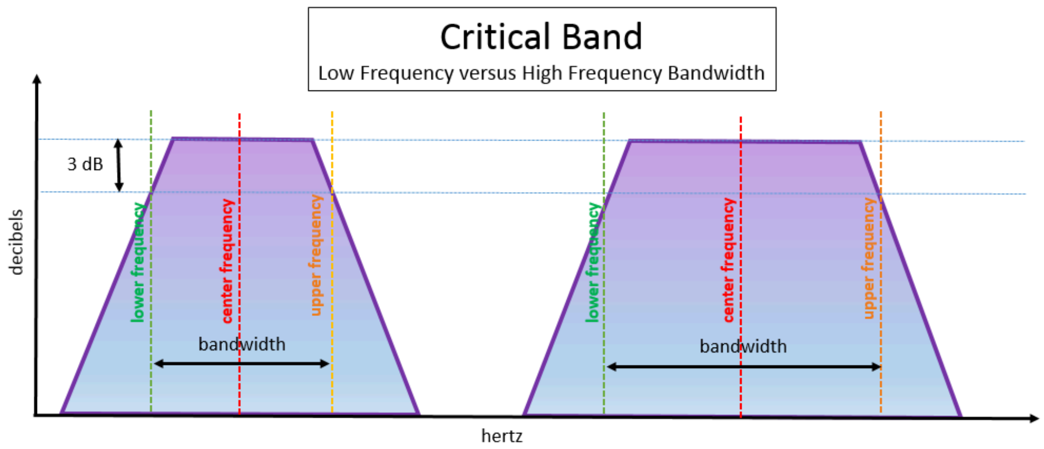
- human ear can hear from 20 to 20,000 Hertz - finer low frequency discrimination vs high frequencies
- ability to distinguish individual tones varies as a function of frequency
- hearing "bands" - used to quantify the ability of the human ear to distinguish between individual frequency tones
- loud response at one place on the basal membrane will mask softer response in the critical band around it




SignalScope X
External Headphones
L:
R:
Signal Generator
Type:
Noise
Sweep
Waveform:
A
B
Sinusoid
Rectangle
Triangle
Sawtooth
Frequency:
1000.0 Hz
Slider Range:
600.0 Hz
to
1300.0 Hz
DC Offset L:
0.000%
DC Offset R:
0.000%
Amplitude:
0.00 dB
Pan:
0.000 %
Oscope
FFT
Octave
Dual FFT
Spectrogram
MultiTool
More

Critical Bands

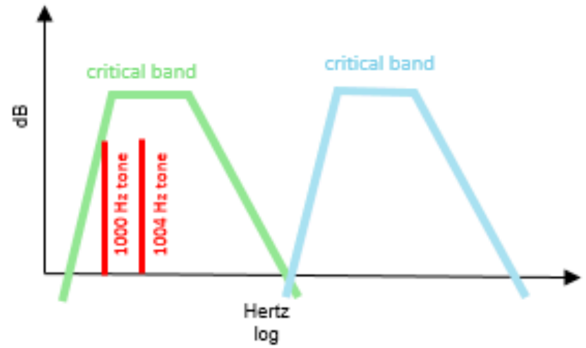
- any audible tone will create a critical band centred around it
- a pure tone (single line in the spectrum) can be represented by a psychological masking curve
- critical bands important (in perceptual coding) - they show that the ear discriminates between energy in and outside the band resulting in masking
- a critical band is the BW at which subjective responses change



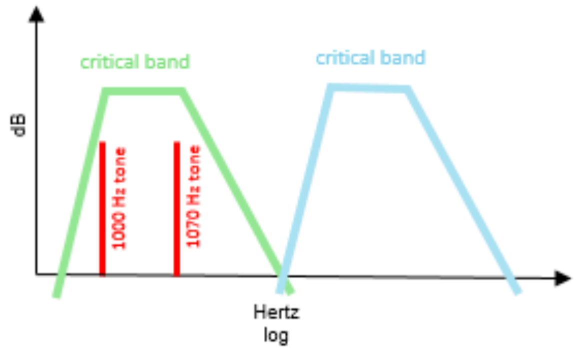
Critical Bands

- codecs rely on amplitude masking within critical bands to reduce information size
- critical bands also used to explain consonance and dissonance
  - tone intervals with a frequency difference greater than a critical band — more **consonant**
  - tone intervals with a frequency difference less than a critical band — more **dissonant**

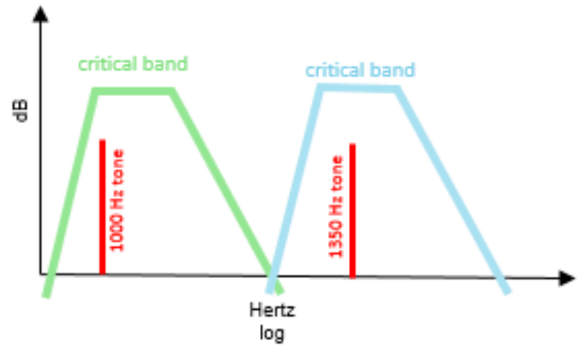
Critical Bands



- ex: tones are 4 Hertz apart - hear a single tone with a low frequency modulation or **beating**.



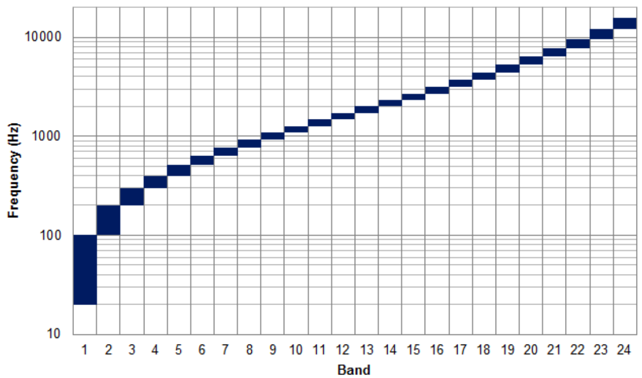
- ex: tones are about 70 Hertz apart, the ear hears a rapid modulation or beating giving rise to "roughness"



- ex: separation of 350 Hertz - tones in different critical bands - can distinguish them from each other

Critical Bands

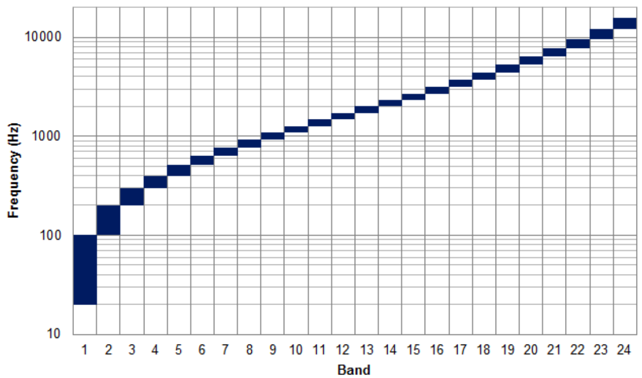
- Bark = unit of perceptual frequency
- Using a Bark scale - physical spectrum can be converted to a psychological spectrum
- 24 critical bands of hearing

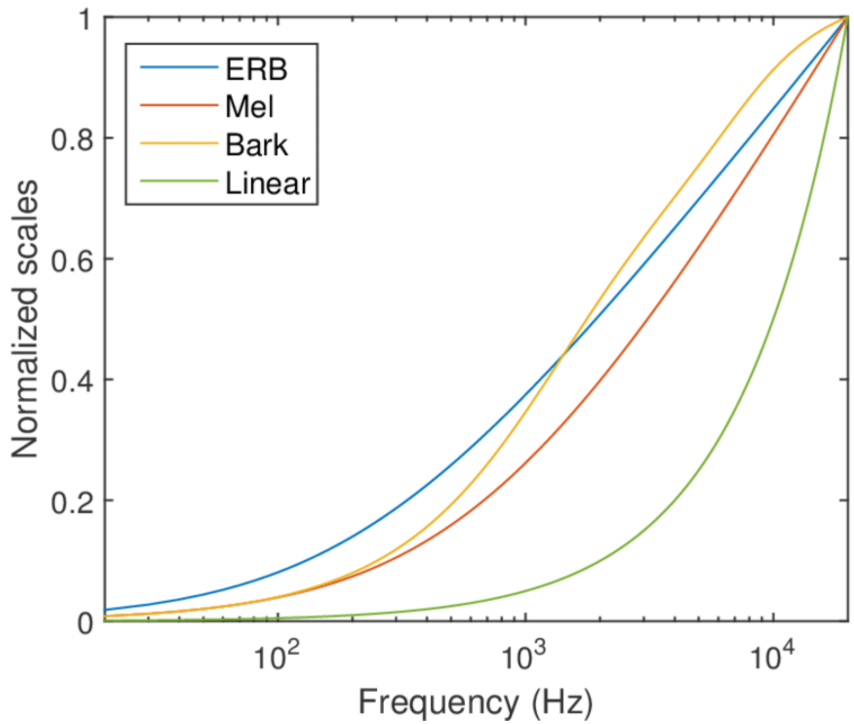


Critical Band	Lower cut-off (Hz)	Central Frequency (Hz)	Higher cut-off (Hz)	Bandwidth (Hz)
1	0	50	100	100
2	100	150	200	100
3	200	250	300	100
4	300	350	400	100
5	400	450	510	110
6	510	570	630	120
7	630	700	770	140
8	770	840	920	150
9	920	1000	1080	160
10	1080	1170	1270	190
11	1270	1370	1480	210
12	1480	1600	1720	240
13	1720	1850	2000	280
14	2000	2150	2320	320
15	2320	2500	2700	380
16	2700	2900	3150	450
17	3150	3400	3700	550
18	3700	4000	4400	700
19	4400	4800	5300	900
20	5300	5800	6400	1100
21	6400	7000	7700	1300
22	7700	8500	9500	1800
23	9500	10500	12000	2500
24	12000	13500	15500	3500

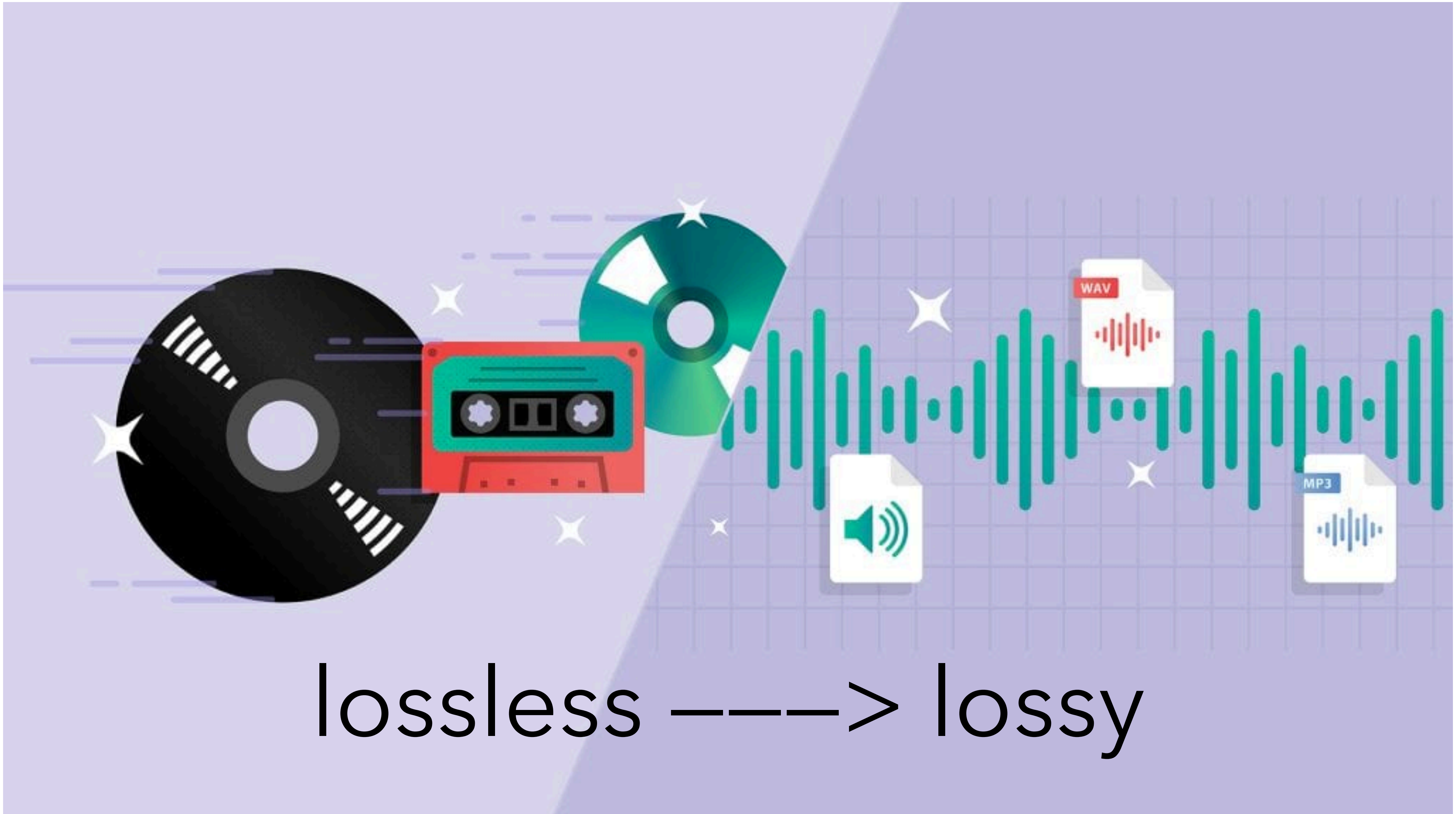
Critical Bands

- Bark = unit of perceptual frequency
- Using a Bark scale - physical spectrum can be converted to a psychological spectrum
- 24 critical bands of hearing





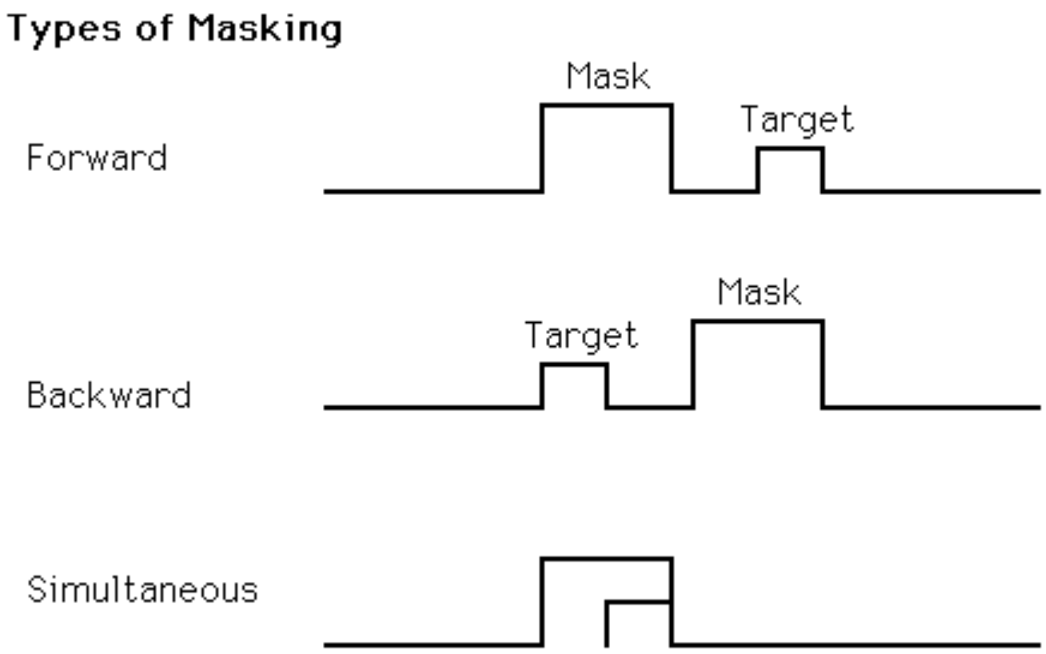
Audio Masking



lossless ———> lossy

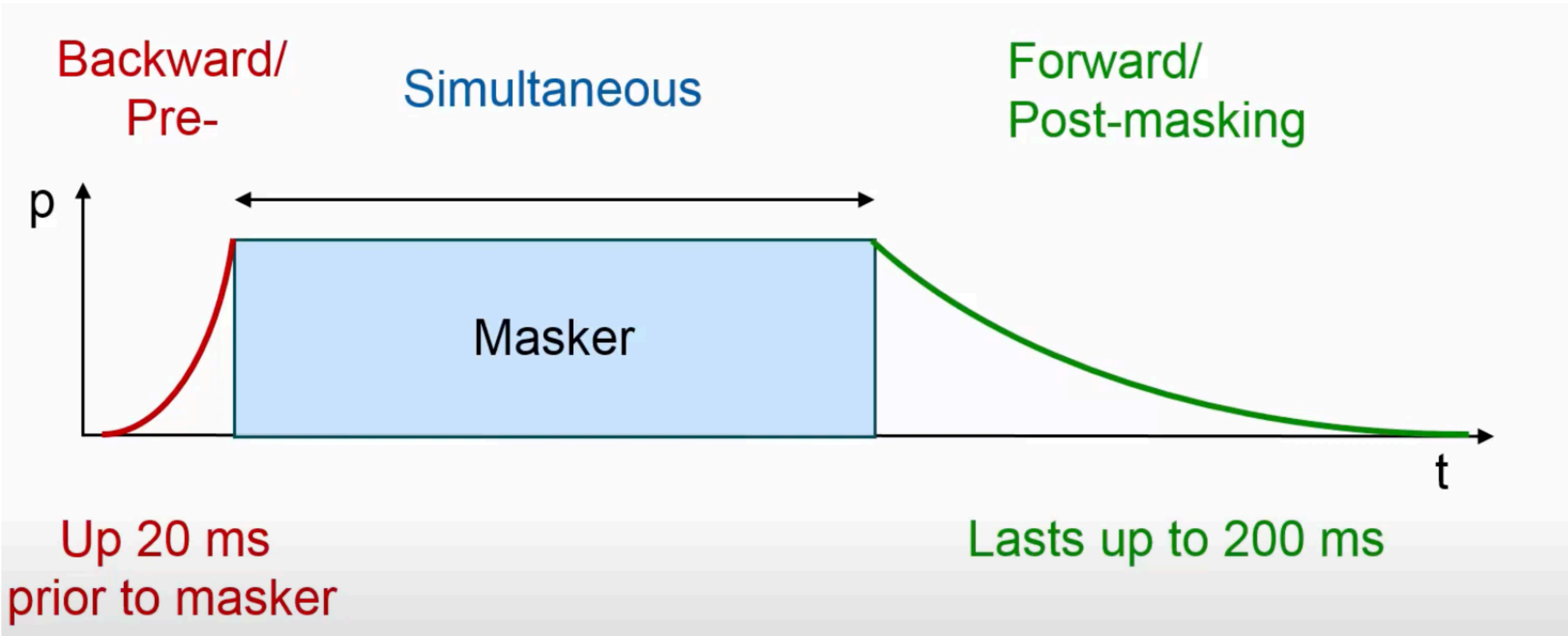
Audio Masking

- **Simultaneous masking vs temporal masking**



Audio Masking

- **Non-Simultaneous masking/temporal masking**
  - when both **do not** occur at same time

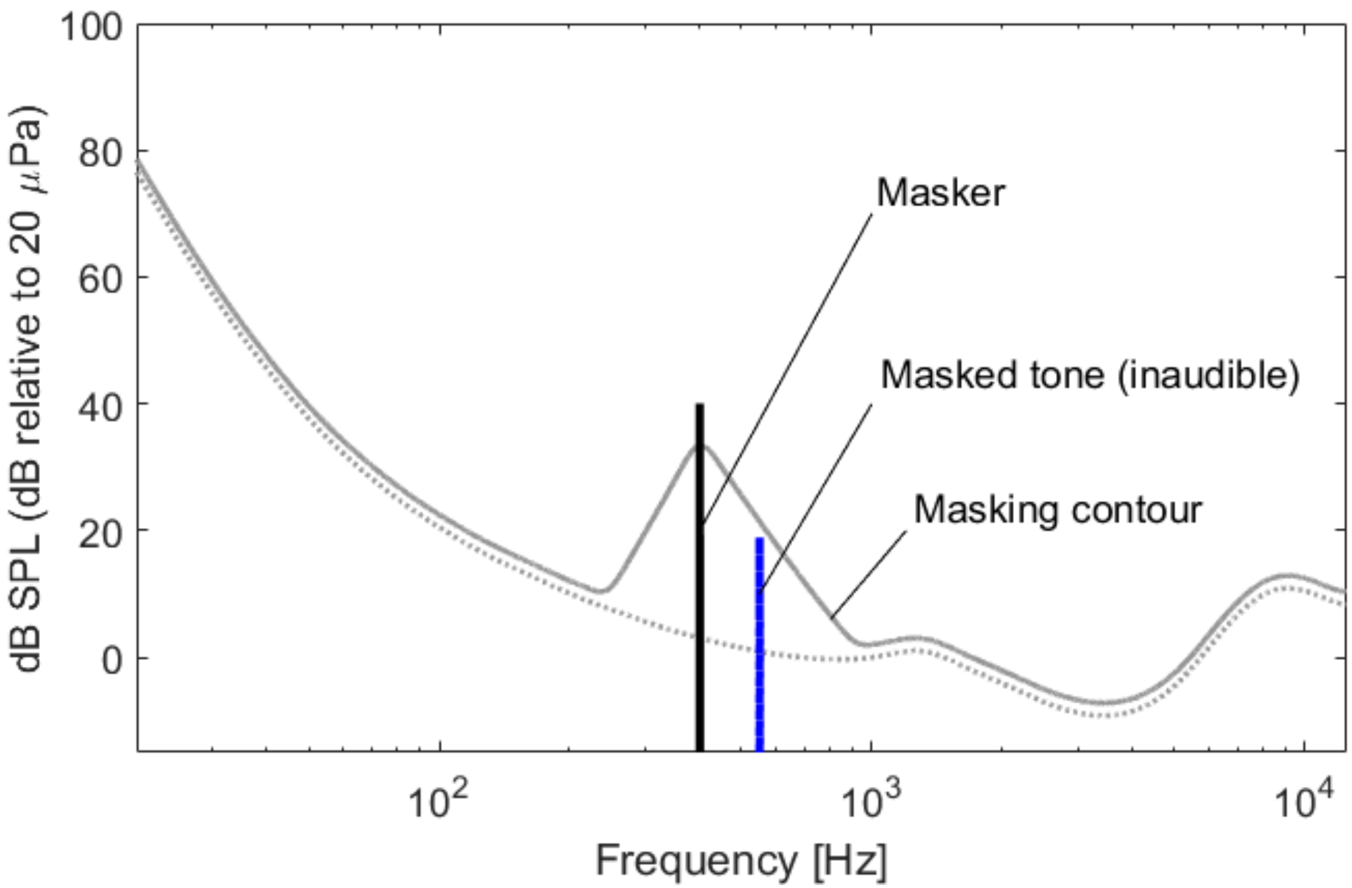


Audio Masking

- **Simultaneous masking/frequency masking**
  - when both occur at same time and close in frequency
  - a loud response on the basilar membrane will mask softer responses in the critical band around it

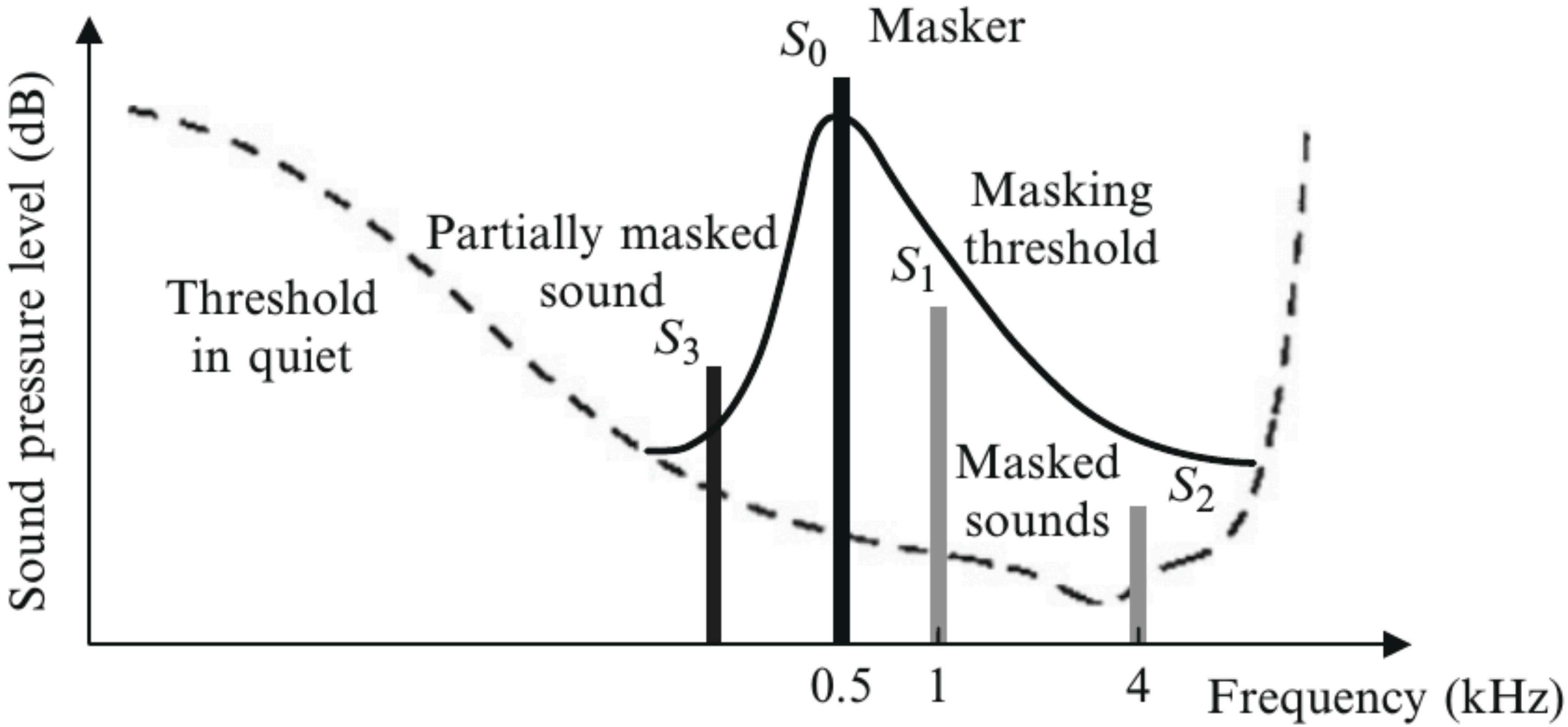
Audio Masking

- **Simultaneous masking/frequency masking**
  - when both occur at same time and close in frequency

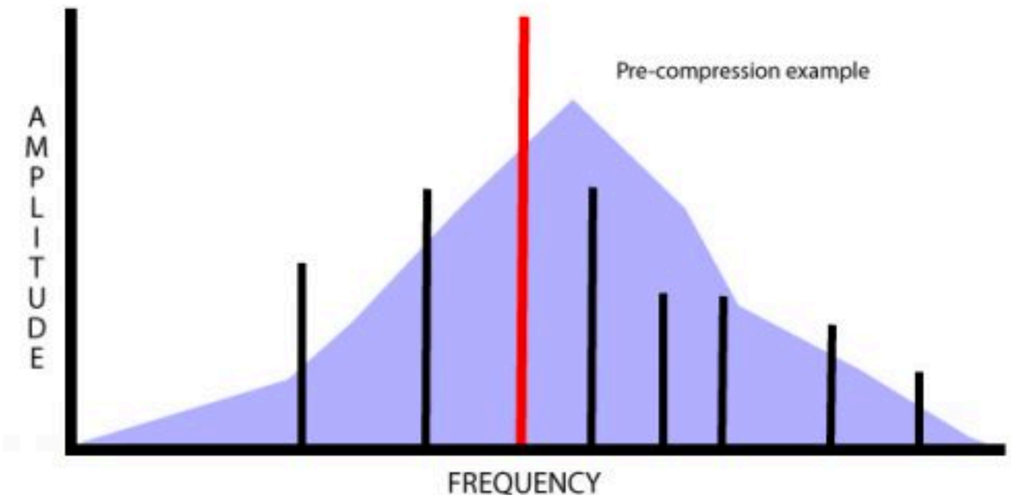


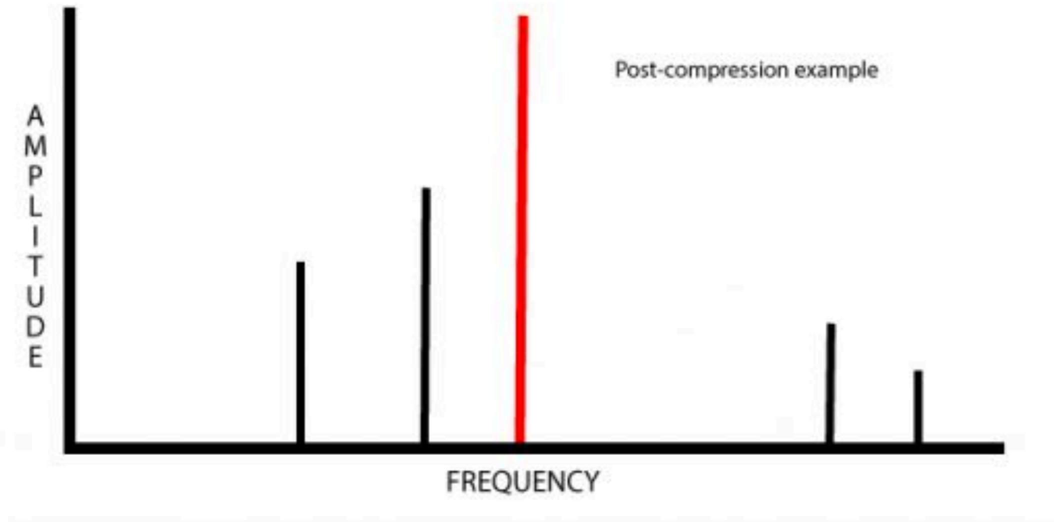
Audio Masking

- **Simultaneous masking/frequency masking**
  - when both occur at same time and close in frequency



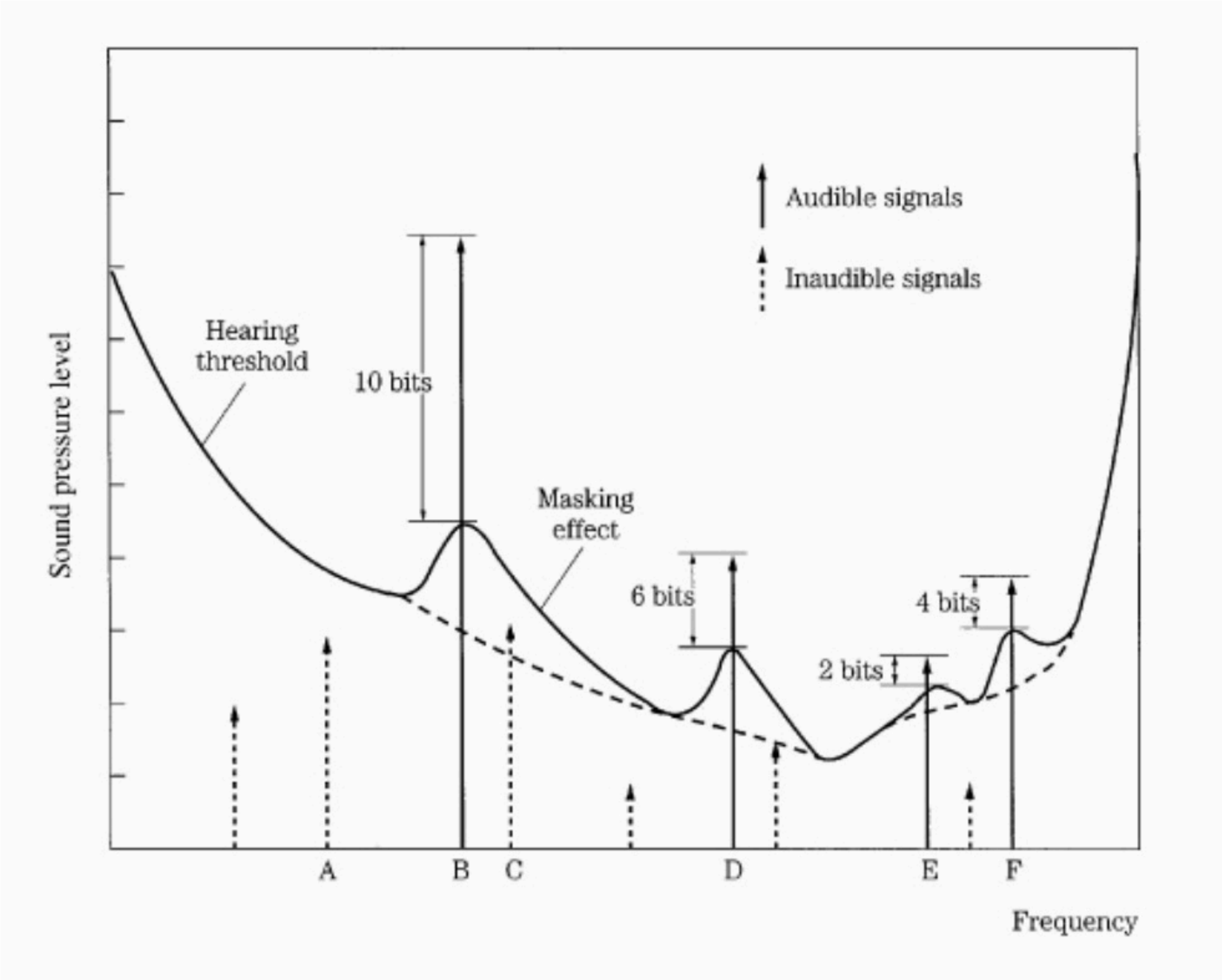
Audio Masking



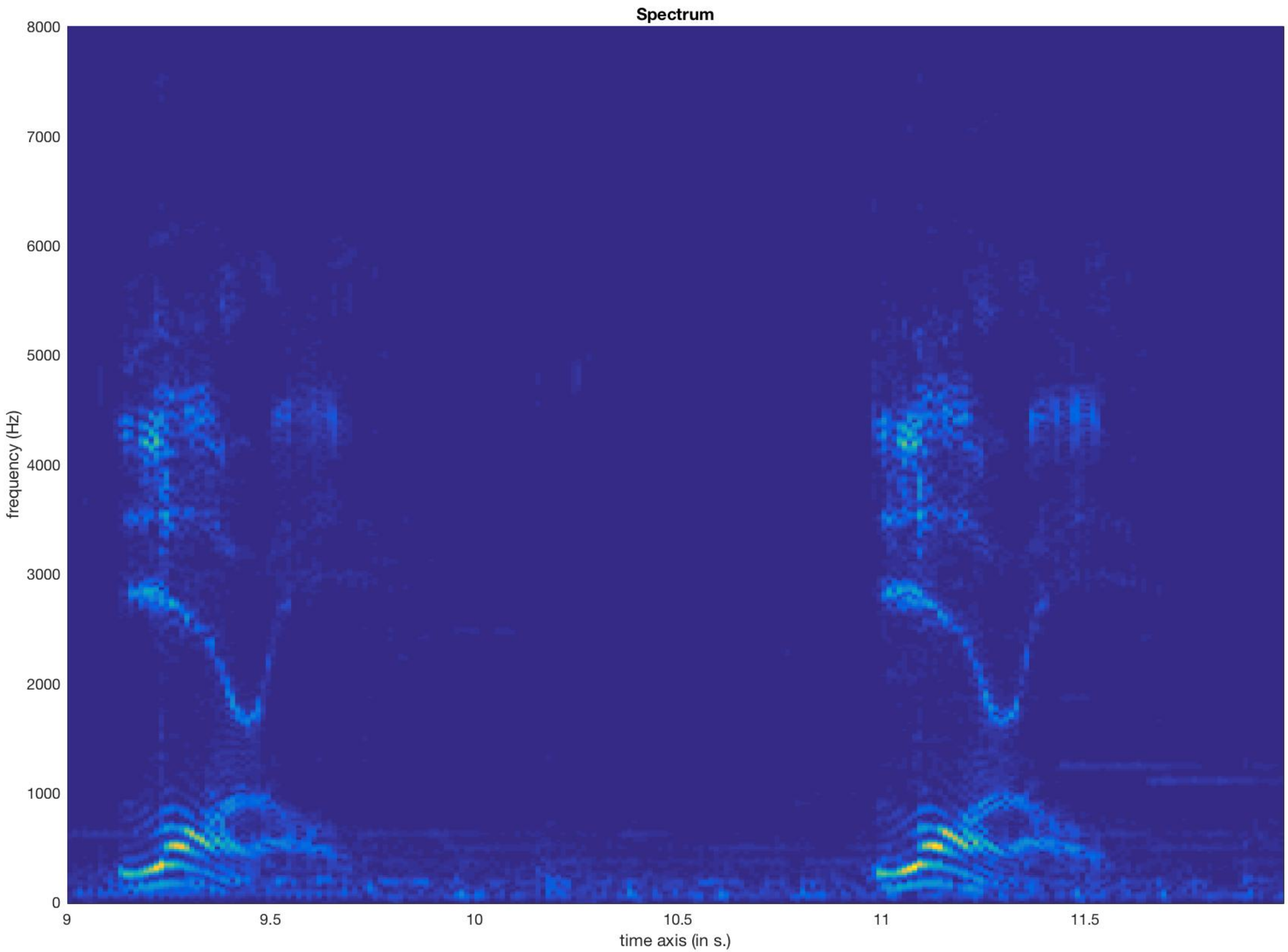


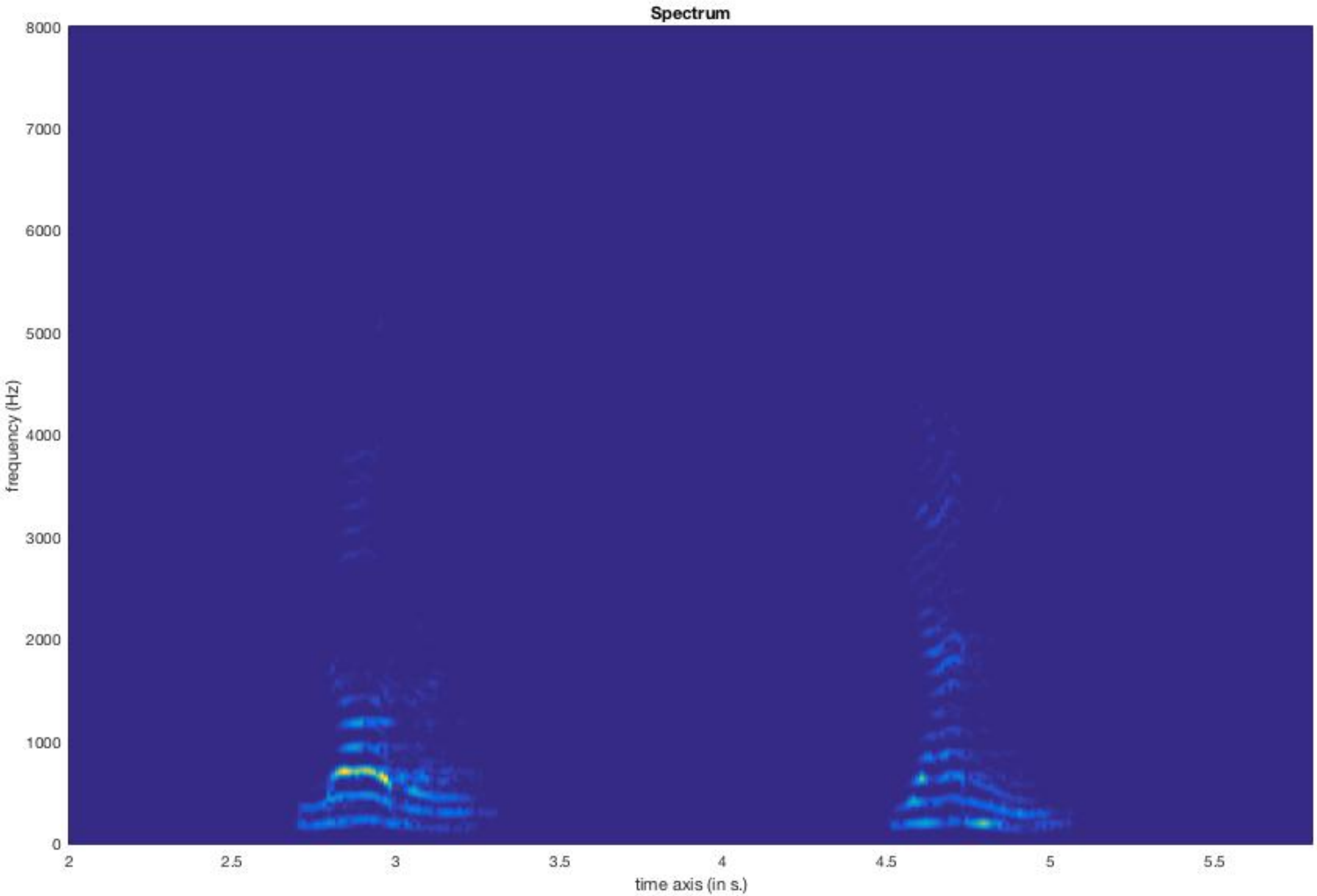
*illustrative purposes only

Audio Masking



sense ≠ perceive

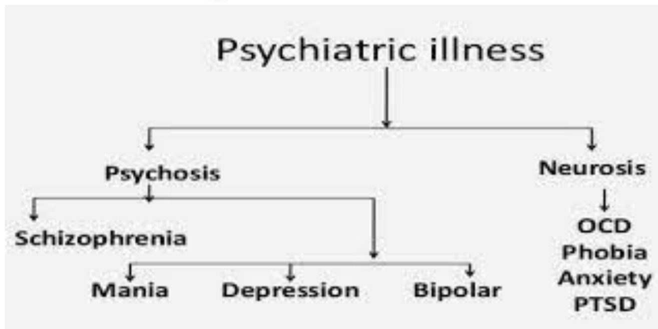

Spectrum
8000
7000
6000
5000
4000
3000
2000
1000
0
frequency (Hz)
9
9.5
10
10.5
11
11.5
time axis (in s.)



Spectrum
frequency (Hz)
time axis (in s.)
0
1000
2000
3000
4000
5000
6000
7000
8000
2
2.5
3
3.5
4
4.5
5
5.5

Auditory hallucinations

"the apparent, often strong, subjective perception of an object or event when no such stimulus or situation is present - perceiving sounds without auditory stimulus"



Auditory hallucinations

- Sensory depravation
  - brain compensation
- Deprivation of social interaction(human conversation)
  - brain more likely to produce hallucinated conversations (eg: sailors)
- Heightened emotional states
  - increase the propensity of the brain to produce corresponding verbal messages (eg: bereavement, abuse, bullying, PTSD)

Auditory hallucinations

- abnormal activation of normal auditory, language perception and production pathways
- activation of PAC, amygdala (emotion), hippocampus (memory), frontal (consolidation) and sensorimotor cortex in schizophrenics (Dirks et al., 1999; Lennox et al., 2000)
- increased blood flow in Broca's area in schizophrenics (McGuire set al., 1993)

"Broca's area is a surprise, since that's where you make sounds, not where you hear them," said Dr. Jerome Engel, a neurologist at the medical school of the University of California at Los Angeles. "I would have expected more brain activity in Wernicke's area, which is where you hear; the usual assumption is that people are listening to thoughts during auditory hallucinations. But this finding suggests that, in terms of unusual brain activity, auditory hallucinations have more to do with the generation of words in the brain than listening to them."

Tinnitus

- "tinnitus" - tinkling in Latin
- noises in the head, not related to any psychiatric condition
- ringing in the left/right ears or in the head (ex: after a loud concert)
- may vary in pitch
  - buzzing
  - hissing
  - humming
  - thumping
  - whistling
  - ticking
  - clicking

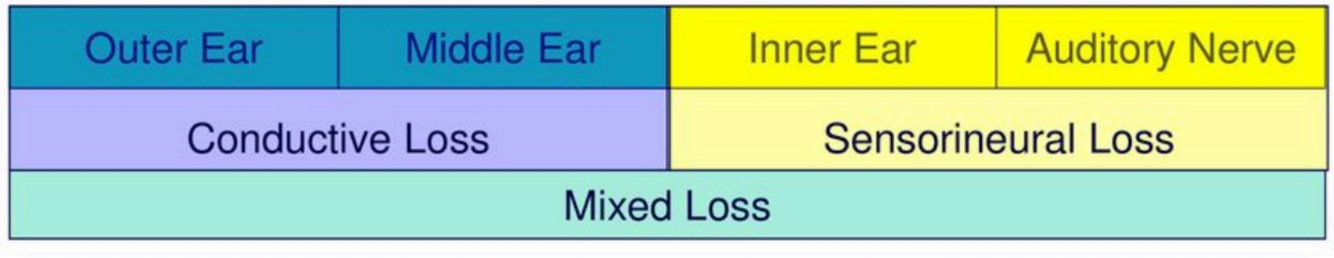
Tinnitus Causes

- damage to the microscopic endings of the hearing nerve in the inner ear due to acoustical trauma
- middle ear infections
- stiffening of the middle ear bones
- foreign object, or earwax touching the eardrum
- high or low blood pressure (blood circulation problems)
- certain types of tumors
- head trauma
- large doses of anti-inflammatories, antibiotics, sedatives, antidepressants, and aspirin
- age

Tinnitus Treatment

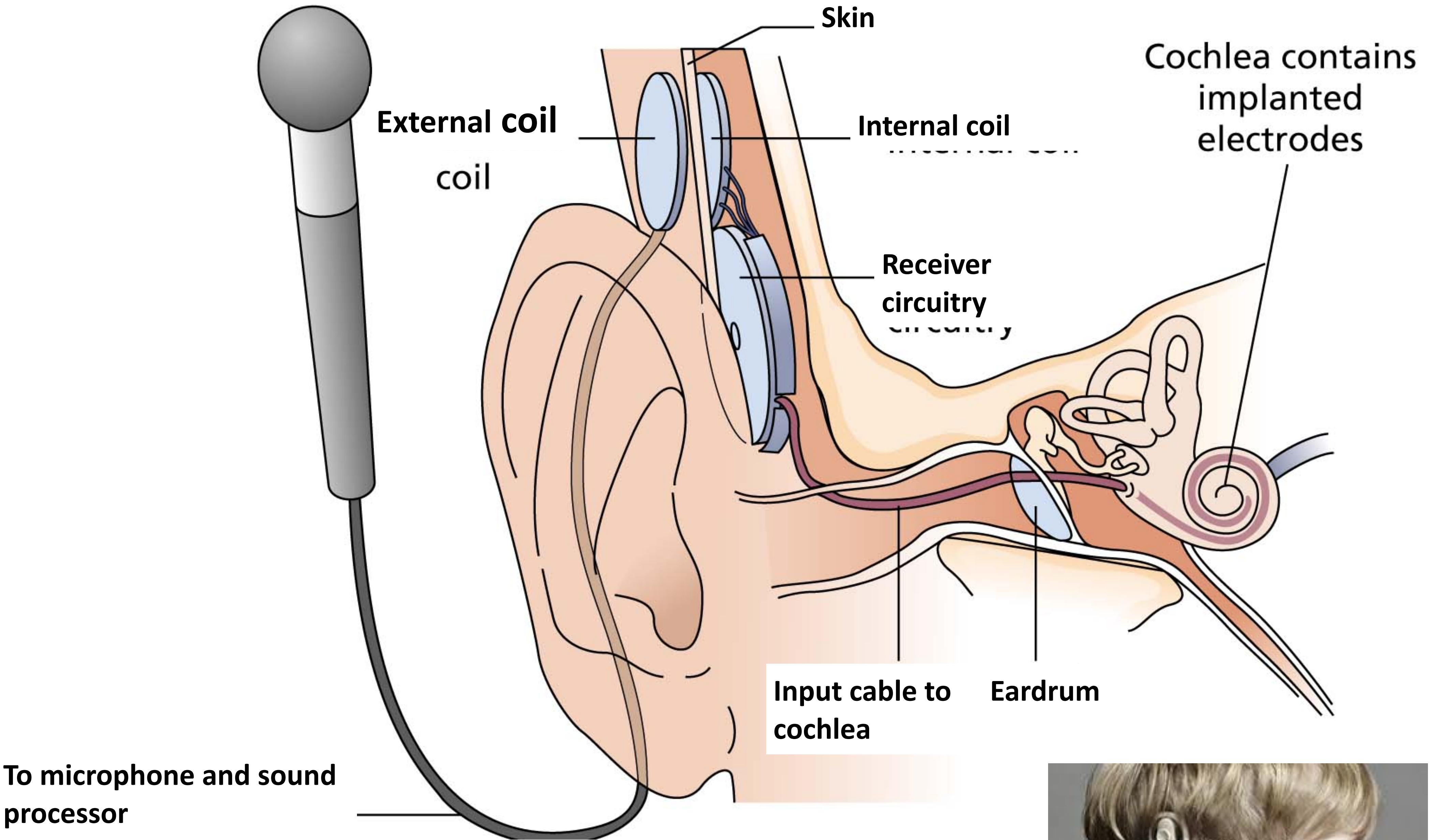
- sound therapy
  - broadband noise maskers (white noise) - (auditory masking)
  - living in india helps!!!!!!!!!
- use ear protection in noisy areas

Types of Hearing Loss

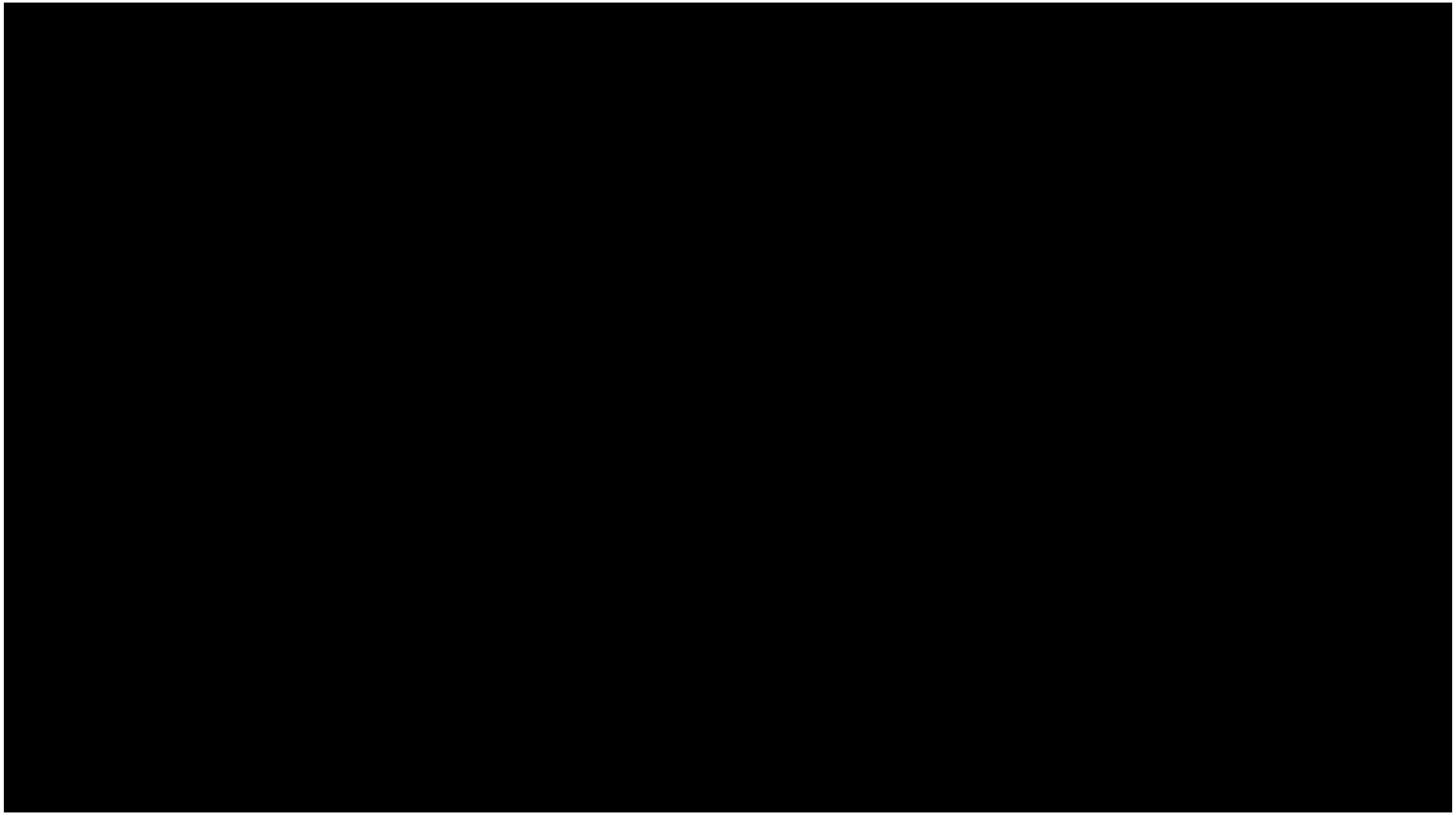


- **Conductive Hearing Loss**: Problems in transmitting sound waves to the cochlea
- **Sensorineural Hearing Loss**: Caused by damage to inner ear or auditory nerve
- **Mixed Hearing Loss**

Cochlear implant



One main cause of Hearing Loss



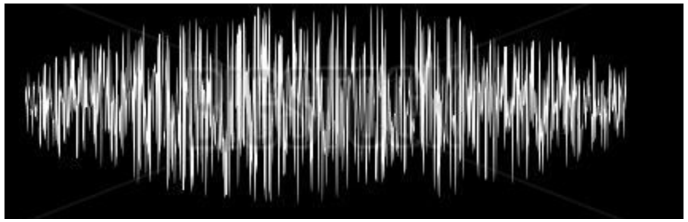
Timbre

Timbre

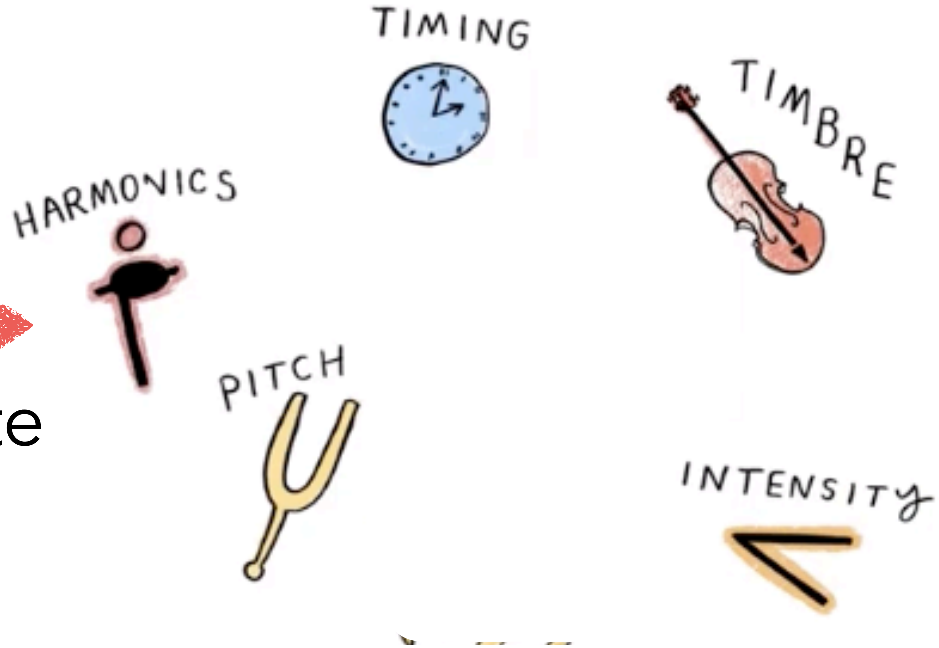

- ”Timbre is the miscellaneous category for describing the psychological attributes of sound, gathering into one bundle whatever was left over after pitch, loudness, and duration had been accounted for.” (Dowling and Harwood 1986)
  - a negative definition or what it is not
- “Psychoacoustician’s wastebasket category”
- timbre is complex - no generally agreed definition
- timbre is multidimensional

Why the fuss about Timbre?

- major structuring force in music and one of the most important and ecologically relevant features of auditory events (Menon, Levitin, Smith, Lembke, Krasnow, Glazer et al., 2002, p. 1742)



separate

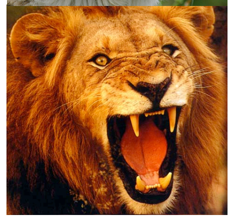


Cocktail Party Effect

"selective hearing" or "selective attention" - a phenomenon that refers to our ability to focus on one specific auditory stimuli while filtering out others

Why the fuss about Timbre?

- major structuring force in music and one of the most important and ecologically relevant features of auditory events (Menon, Levitin, Smith, Lembke, Krasnow, Glazer et al., 2002, p. 1742)



- vehicle for source identity (McAdams & Giordano, 2009; McAdams, 1993; Handel 1995
    - survival: escaping danger elicits feeling of reward, which can be thought of as a fundamental incentive to human action

Attempts to Define what Timbre is

- ”On analysis, the difference between tone-colors of instruments are found to correspond with **differences in the harmonics** represented in the sound.” (Jacobs 1991. Penguin Dictionary of Music)

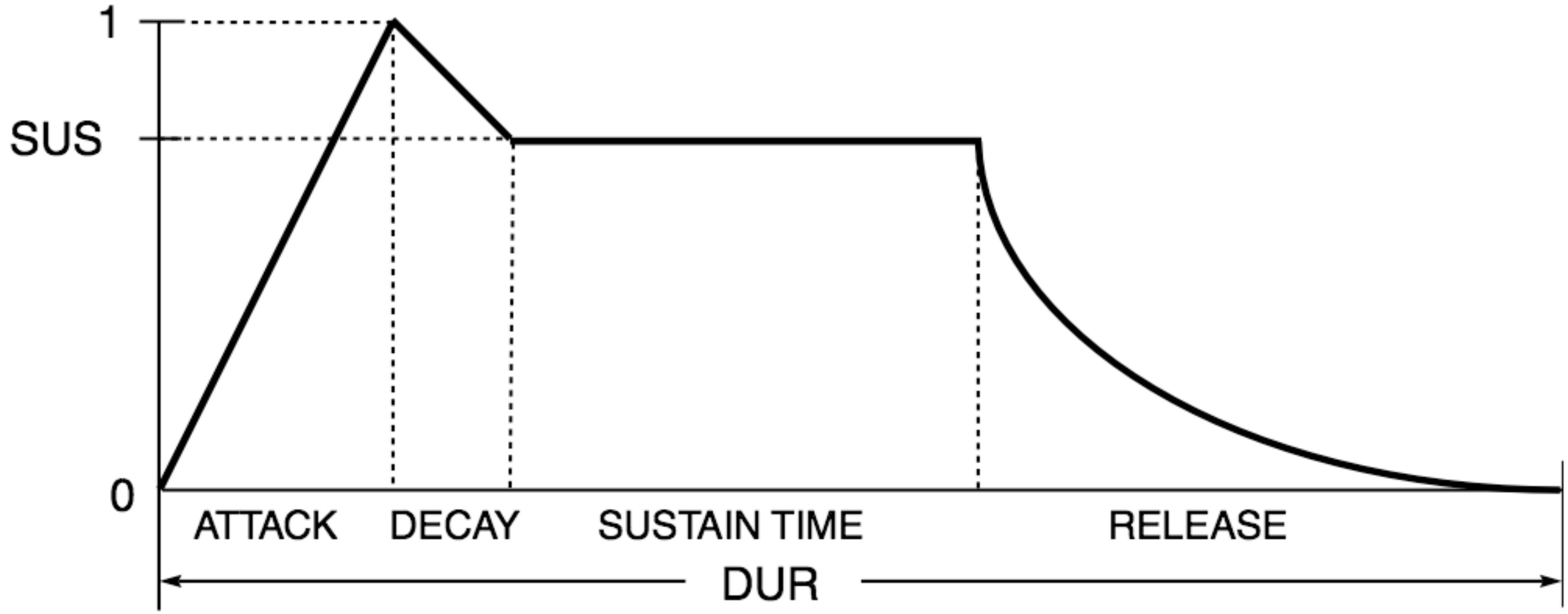
- ”[Sound] quality is determined by the overtones, the distinctive timbre of any instrument being the result of **the number and relative prominence of the overtones** it produces.” (Columbia Encyclopedia)

If timbre depends on ”relative prominence of overtones” only ...

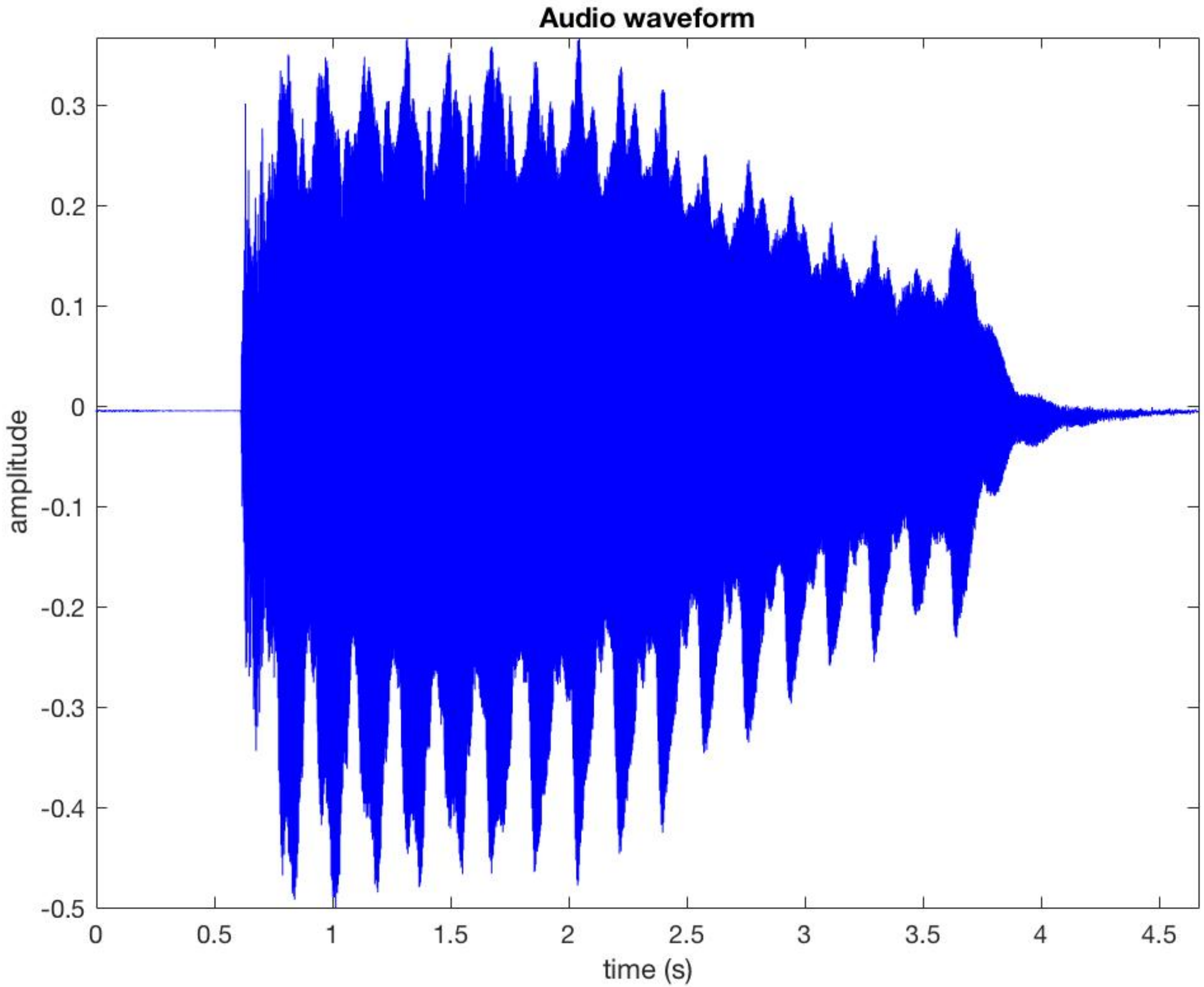
- 1. play some ragtime ...
- 2. play it backwards ...
- 3. reverse the sound file ...
- 1 & 3 have similar overtone structure -> same timbre?

Temporal Envelope

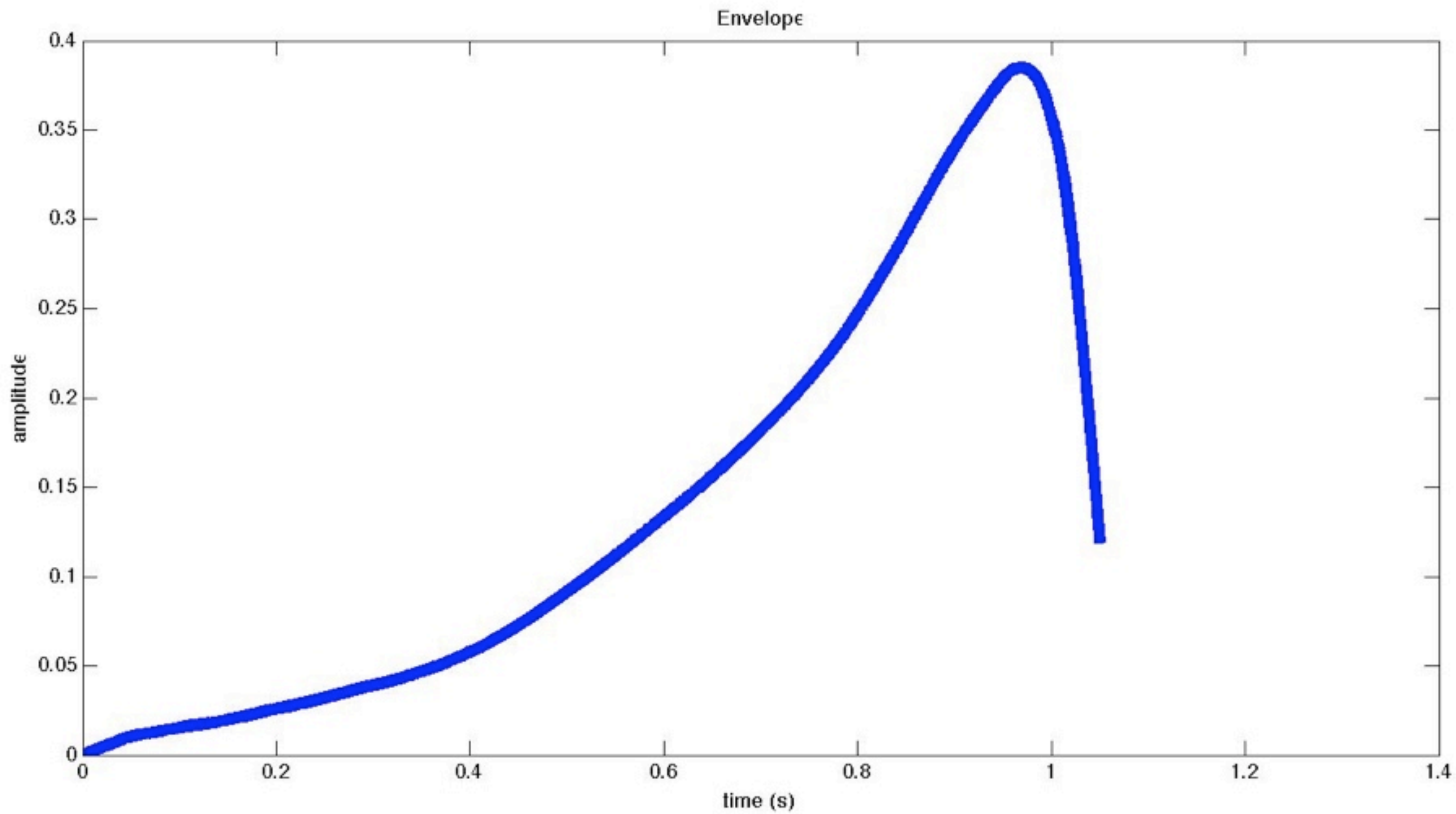
- Attack-Decay-Sustain-Release



Temporal Envelope

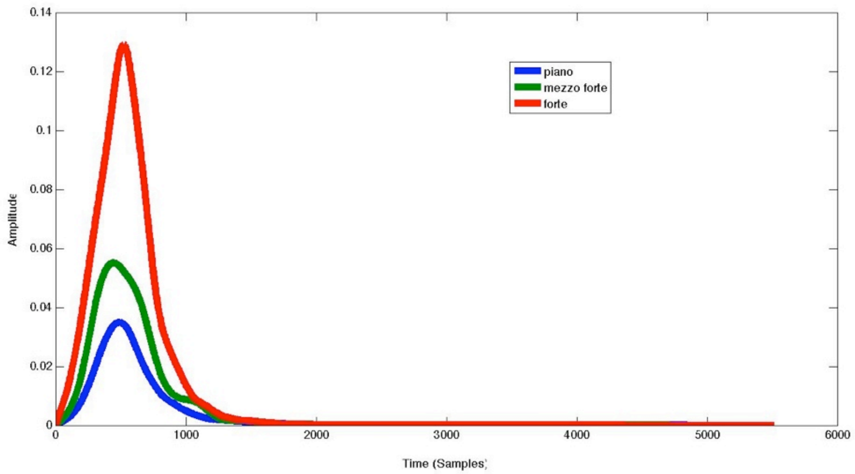


Attack Time & Attack Slope

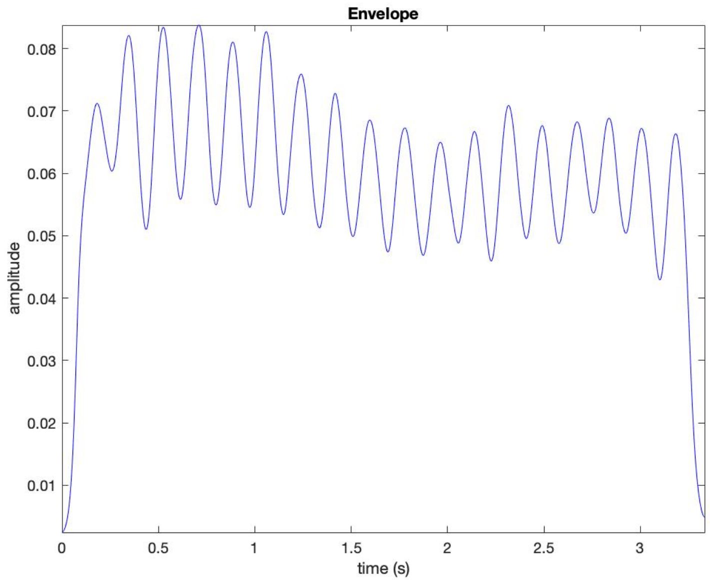


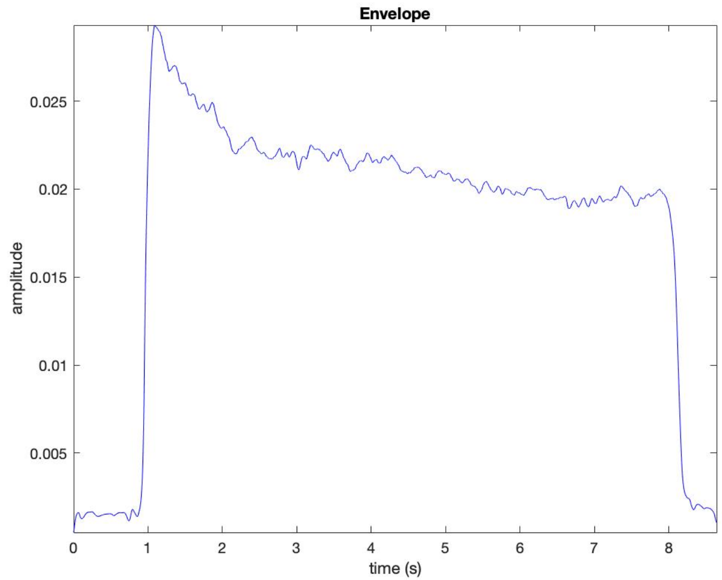
Attack Time & Attack Slope

- removal of attack?
  - difficult to identify the source
- Effect of Loudness
  - affects attack slope



Match the sound to temporal envelope

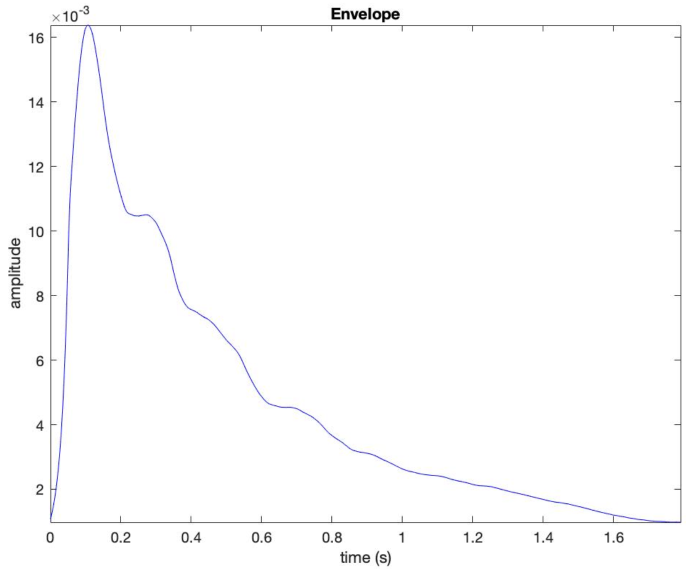




piano

flute

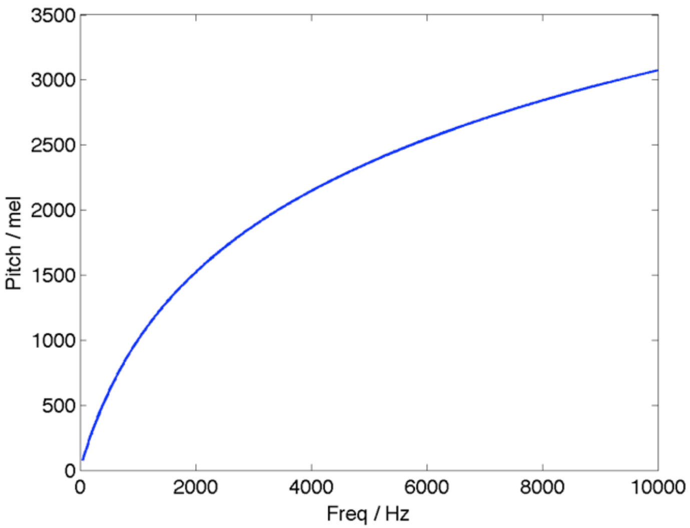
trumpet

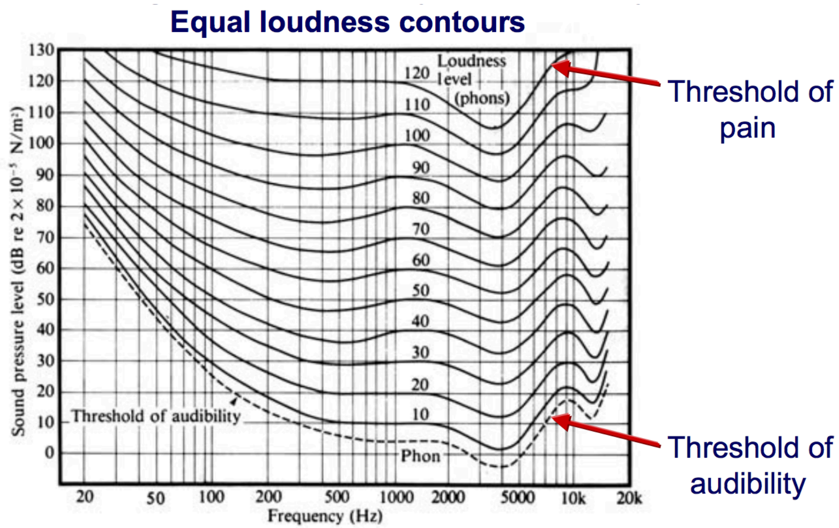


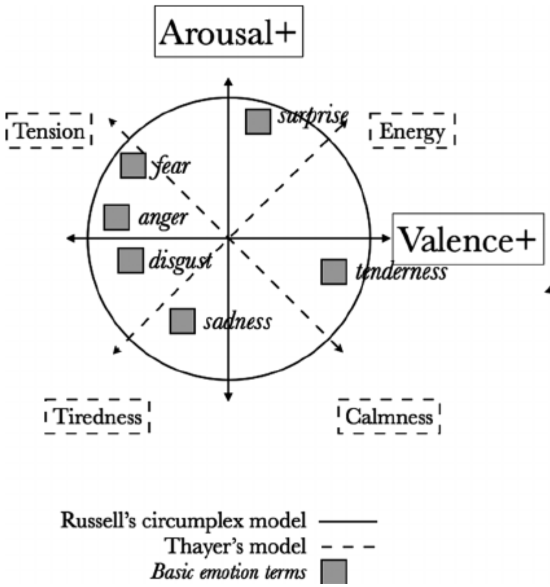
Physical Representations

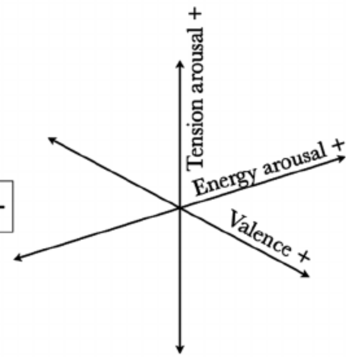
How would you investigate timbre perception?

Perceptual Representations



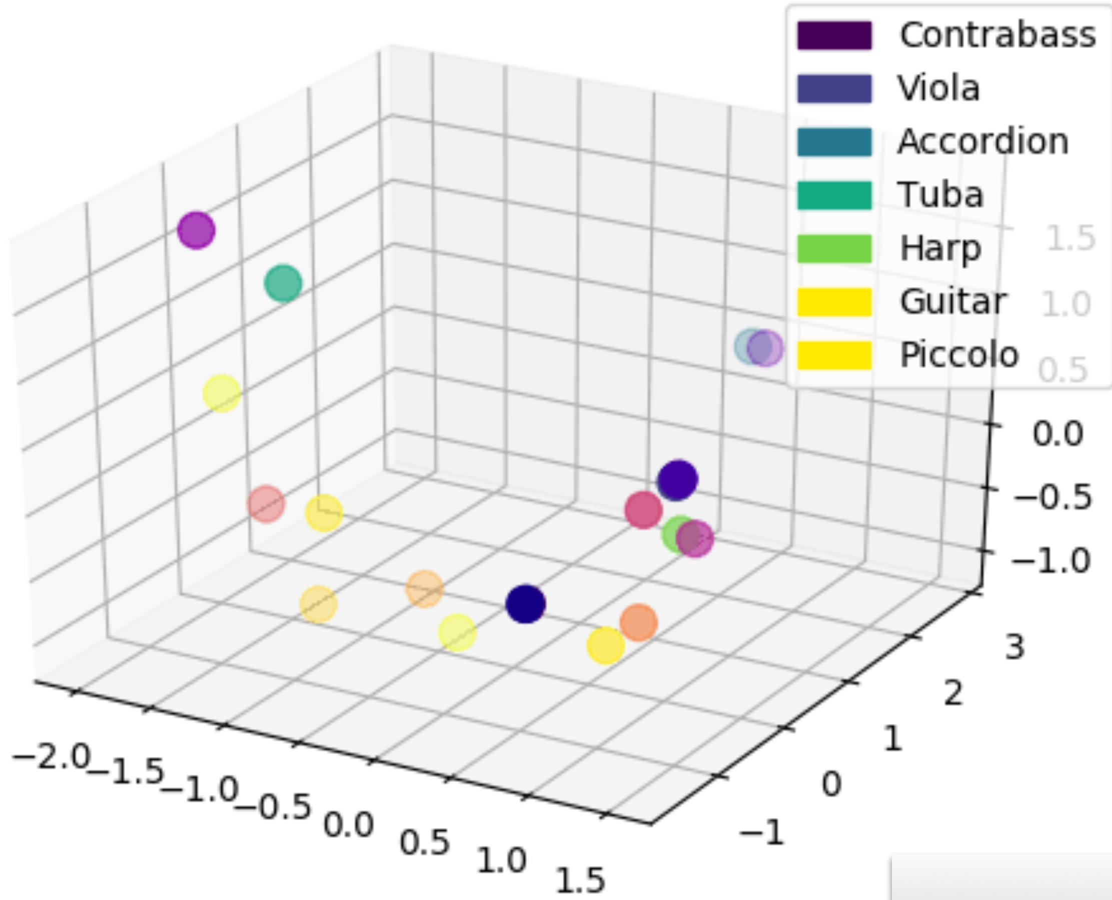






Schimmack & Grob model

Timbre Spaces....



Dimension 3?

Dimension 2?

Dimension 1?

Timbre Spaces....

- a model that tries to explain perceptual results such as timbral interval perception
- different sets of stimuli give rise to different sets of dimensions with differing acoustic correlates
- hence requires a meta-analytic approach

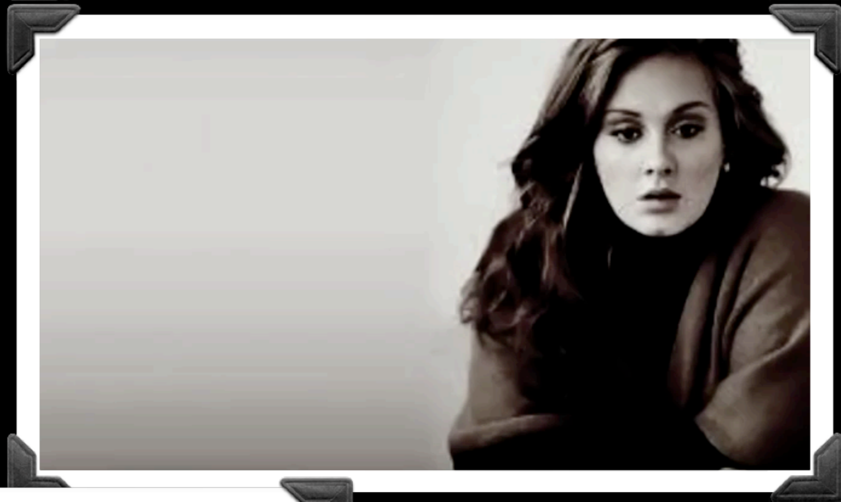


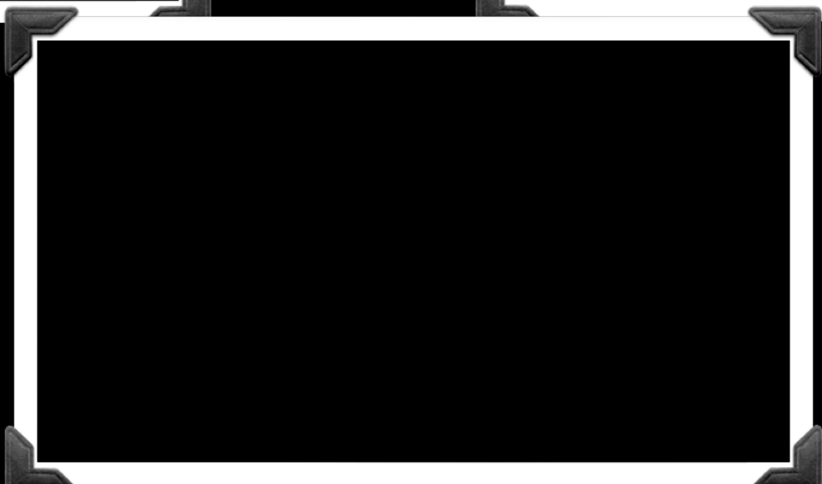






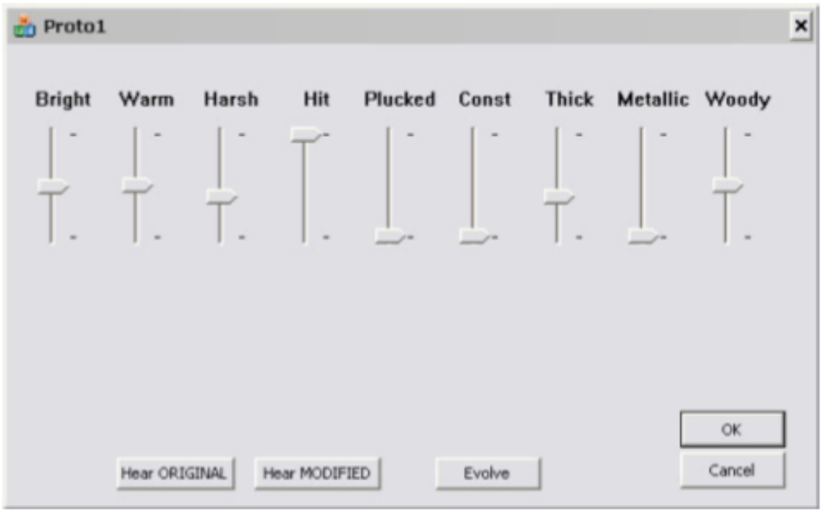

CARTASI, MASTERCARD, VISA NUMERO





Perceptual Descriptions?

- why is it important?
  – “common language” between artist and producer



Gounaropoulos & Johnson, 2006

Timbre
Semantic Studies

- Adjectives describing timbre
- Surveys, Semantic Differential, VAME
  - *harsh, bright, full, warm .....*
- Acoustic correlates of descriptors
  - *spectral centroid, spectral slope, roughness,....*

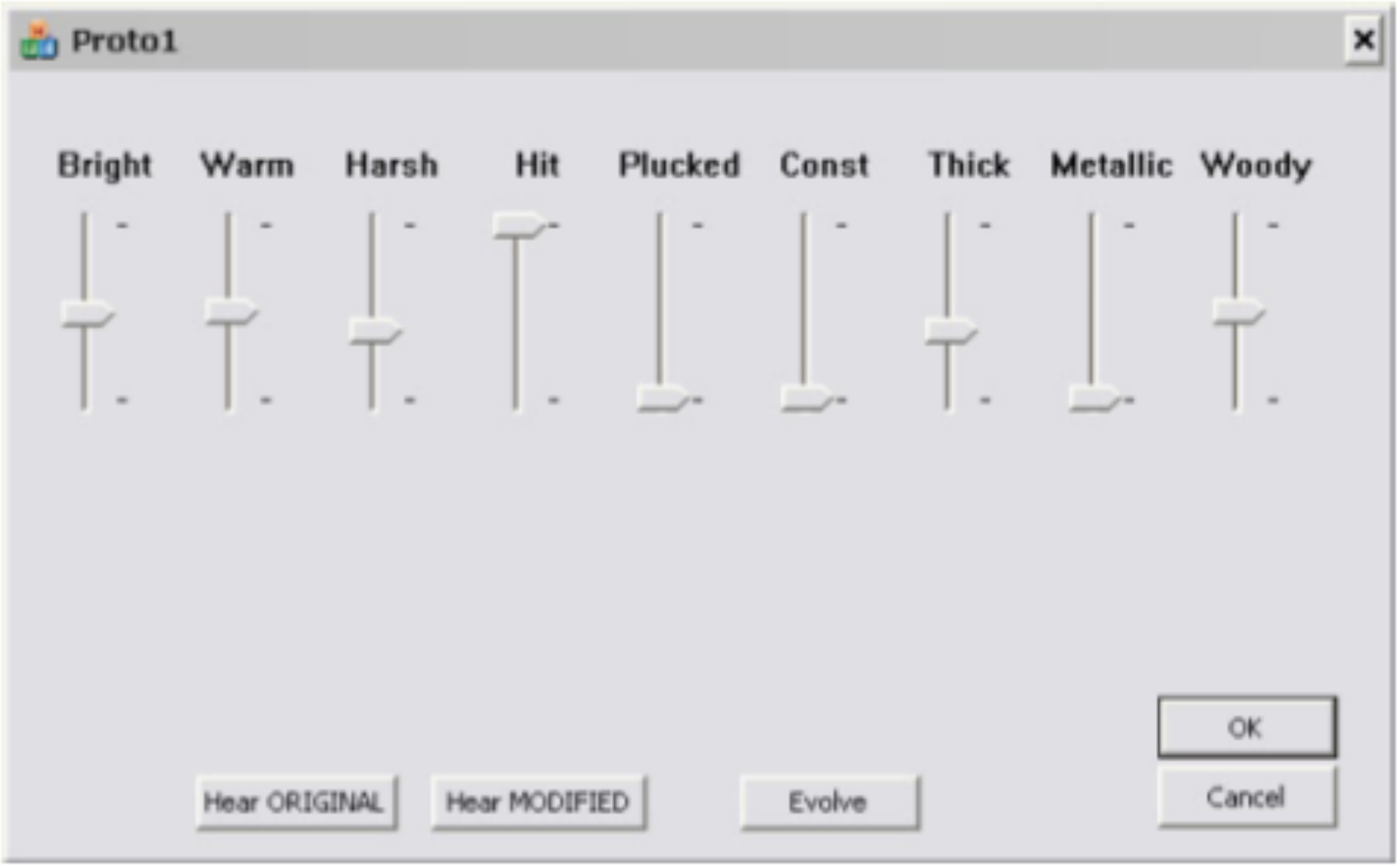
Perceptual Descriptions

- proposed rating scales (W. Sethares 1999) :
  - dull <-> sharp
  - cold <-> warm
  - soft <-> hard
  - pure <-> rich
  - compact <-> scattered
  - full <-> empty
  - static <-> dynamic
  - colorful <-> colorless

Physical and Perceptual Correlates of Timbre (Bolger 2005)

Name	Type	Physical Correlate	Perceptual Correlate	Description
Spectral centroid	Spectral	Energy concentration in low/high spectral area	Brightness/ Dullness	Balance of energy in spectrum.
Irregularity	Spectral	Fluctuating energy between adjacent partials	Richness	Amplitude variation of adjacent components.
Roughness	Spectral	Beating of overlapping partials	Harshness/ Smoothness	Inharmonic and noise components in spectrum.
Harmonicity	Spectral	Harmonic/ Inharmonic	Cohesive/ Diffuse	Ratio of harmonic to inharmonic spectral components.
Attack/ Decay times	Temporal	Slope of attack and decay	Instrument identification	Time taken to reach max. amp from 0 (attack).

Timbre Synthesis Interfaces



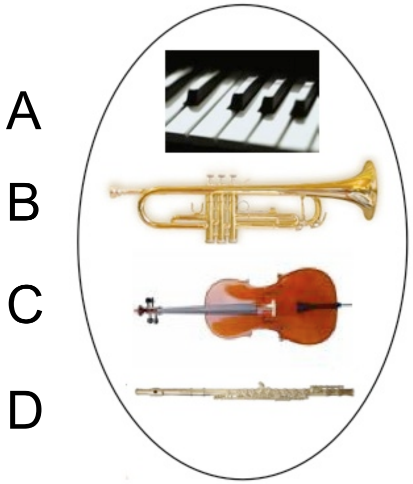
- Gounaropoulos & Johnson, 2006

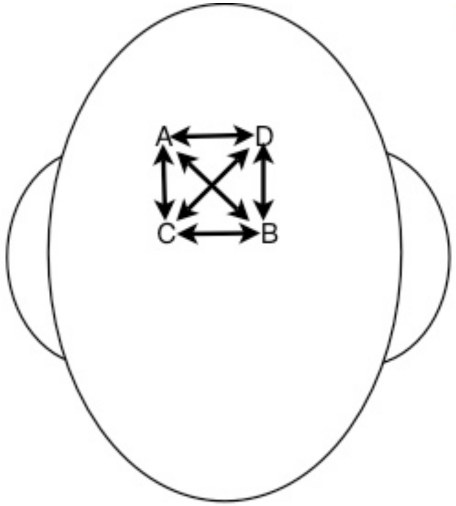
Perceptual and Acoustical Correlates of Timbre

– A few studies have aimed at creating ‘physical’ timbre spaces from features of the audio in order to find correlations with perceptual timbre spaces (De Poli et al. (1993), Cosi et al. (1994), Toiviainen et al. (1995), Loureiro et al. (2004), Teresawa et al.(2005))

– clustering of sounds in these studies has been found to be comparable to human similarity judgments

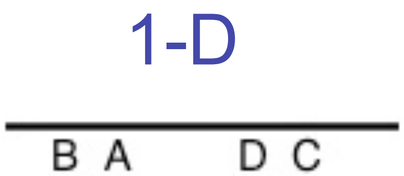
Timbre Spaces....

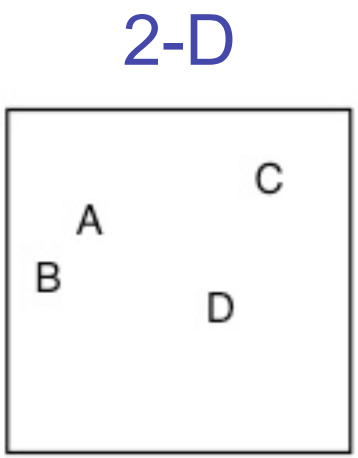


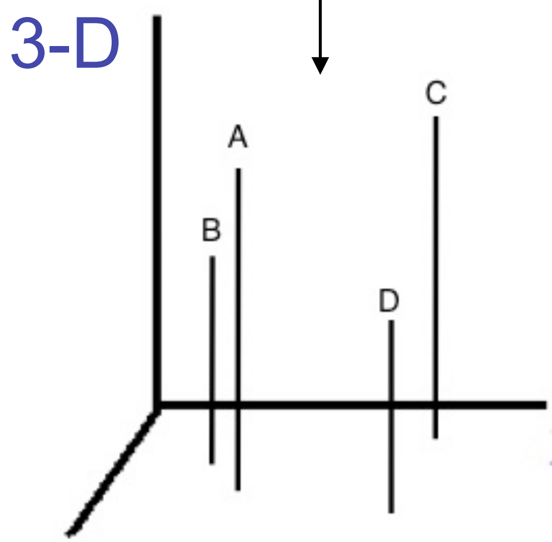


	A	B	C	D
A		2	5	6
B	2		3	4
C	5	3		2
D	6	4	2	

MDS



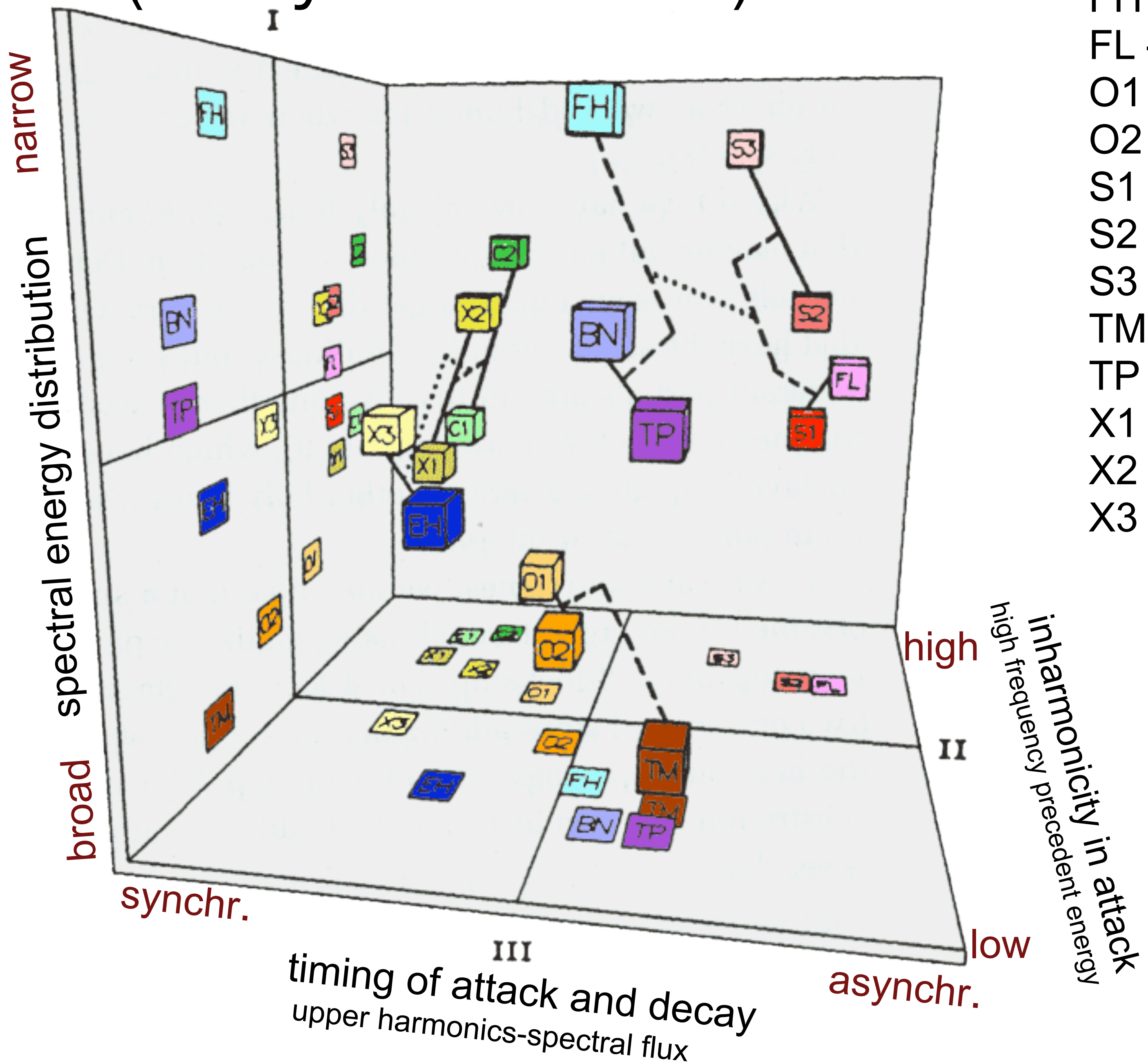




Correlating Timbre with Physical Attributes of Tones

- Grey & Moorer 1977:
  - 16 synthetic tones
    - equal pitch and duration
    - resembled acoustical instruments
  - similarity rating experiment
    - perceived similarity of all pairs of tones
  - Multidimensional Scaling solution
    - projection of timbres onto a 3-D space

Timbre Space (Grey & Moorer)



BN - Bassoon
C1 - E flat Clarinet
C2 - B flat Bass Clarinet
EH - English Horn
FH - French Horn
FL - Flute
O1 - Oboe
O2 - Oboe
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

Meta-analysis of Timbre Spaces

- Recorded musical instrument tones
  - Lakatos(2000-winds strings,percussion,combined), Iverson & Krumhansl (1993 - whole tones)
- Recorded and modified instrument tones
  - Grey(1977), Grey&Gordon(1978), Iverson & Krumhansl (1993 - attack and remainder portions of tones)
- FM-synthesized simulations of orchestral instrument tone
  - Krumhansl(1989), McAdams et al.(1995)

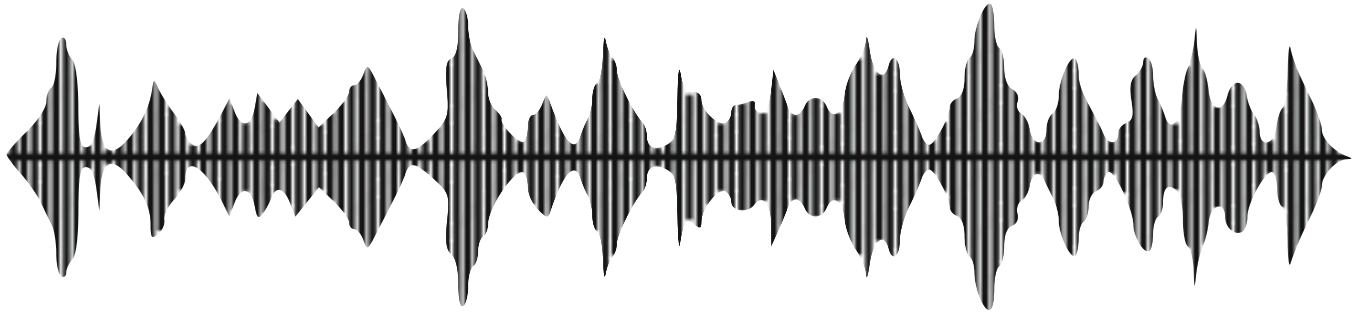
Meta-analysis of Timbre Spaces

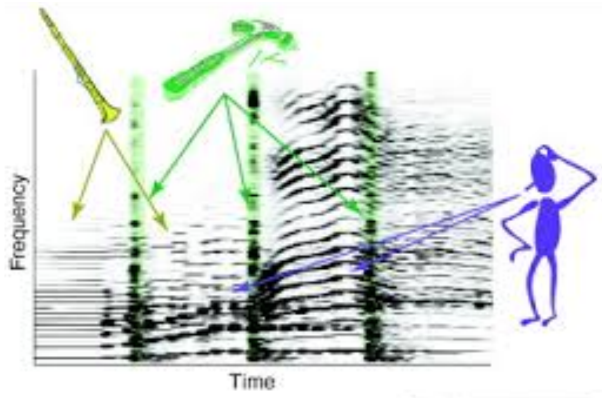
- Use same MDS technique on all data sets
- Use same set of acoustical descriptors derived from signals on all sound sets
- The meta-analytic approach seeks generalization of results across diverse experimental conditions

Interpretation of Dimensions

- Across studies, several different potential acoustic cues correspond to timbre dimensions
  - **Spectral centroid**
  - Spectral deviation
  - Spectral density
  - **Attack time**
  - Decay time
  - Amplitude envelope
  - **Spectral flux**
  - Pitch strength
  - Attack synchrony
  - Attack centroid
  - Noisiness

spectral
temporal
spectrotemporal



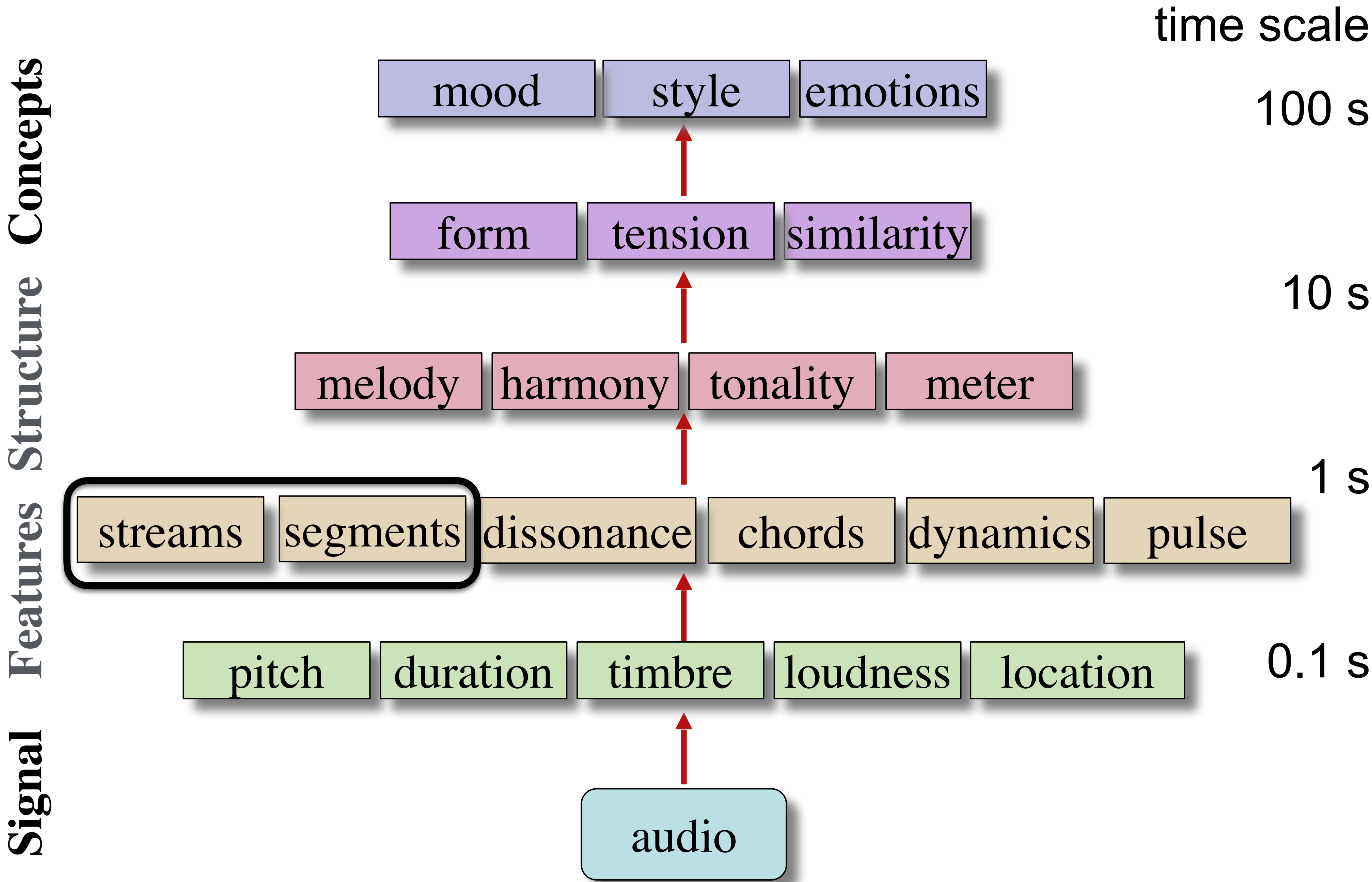


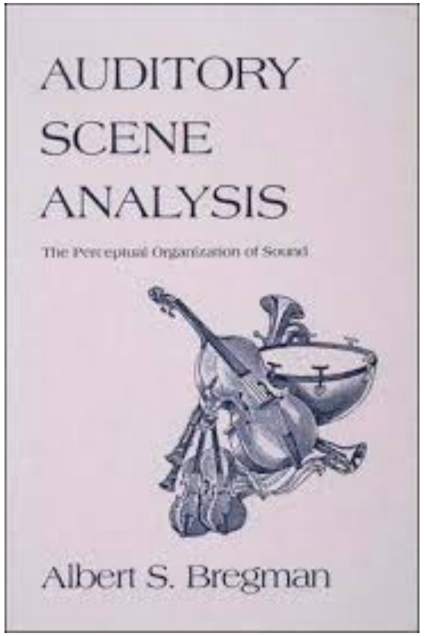
Frequency
Time

How many violins/violinists/streams do you hear?

3:40 4:29 6:56

Levels of Music Processing

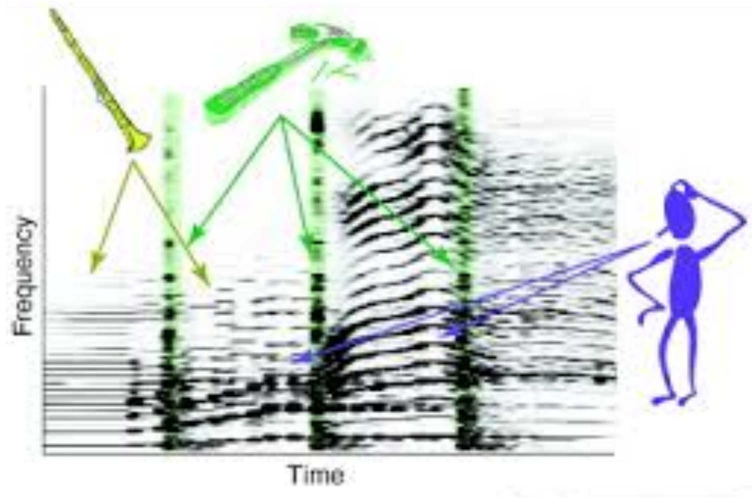


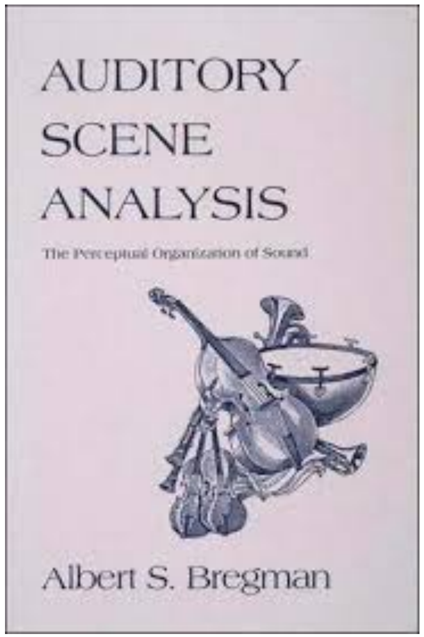


Perception and Perceptual Organisation

Auditory scene analysis

- coined by **Albert Bregman**
  - proposed model for the basis of auditory perception - the auditory system needs to make sense of the superposition of component sounds (**the auditory scene).**
  - human auditory system segregates and organizes sound into perceptually meaningful elements or **auditory streams** from an incoming mixed stream
  - it needs to group the components of the sound that come from the same sound source
  - ability to determine **location** and **distance** of sound
  - many of these can be explained by auditory analogues of the **Gestalt principles**





Gestalt Theory

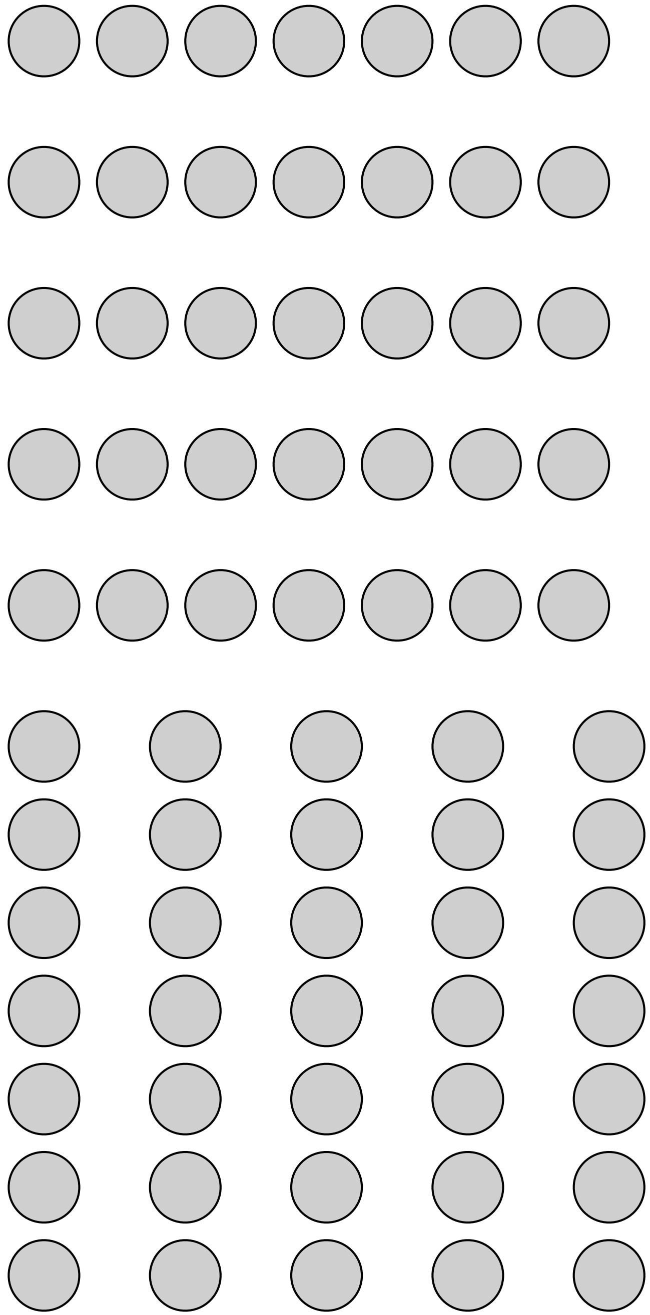
- Gestalt: shape or form
- started in 1890s
  - Kurt Koffka, Max Wertheimer, Wolfgang Köhler
- reaction to *atomism (*nature of things is absolute and not dependent on context)
- *holistic/gestalt*: **the whole is something else than the sum of its parts**

Gestalt principles of perceptual organization

- brain’s innate organising tendencies allow us to perceive things as organised wholes than individual elements
- main ideas:
  - we often experience things that are not a part of our simple sensations
  - perception of things is affected by the context
  - ”The whole is something other than the sum of its parts”
  - can be summarized as a set of principles



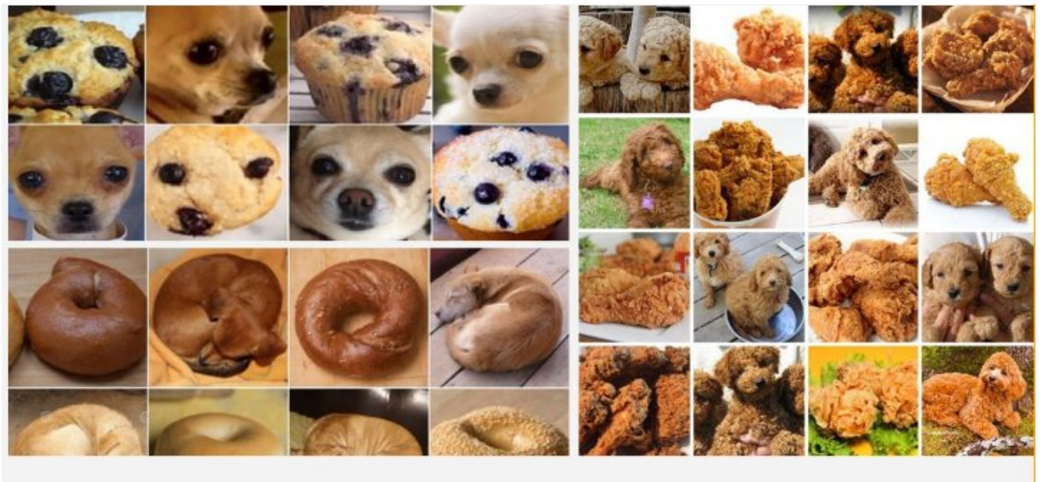
Principle of Proximity

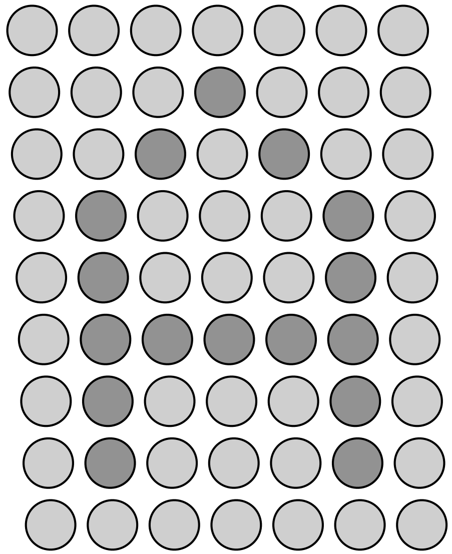


- objects that are close to each other tend to be grouped together

Principle of Similarity

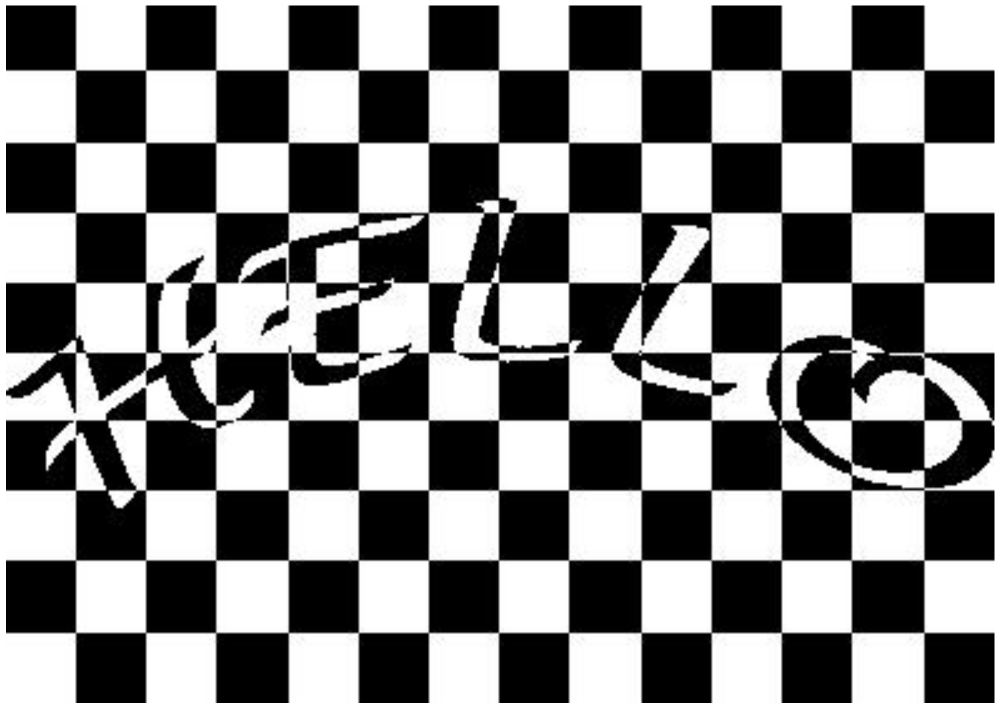
- objects that share visual characteristics such as colour, shape, size, texture, or orientation tend to be grouped together



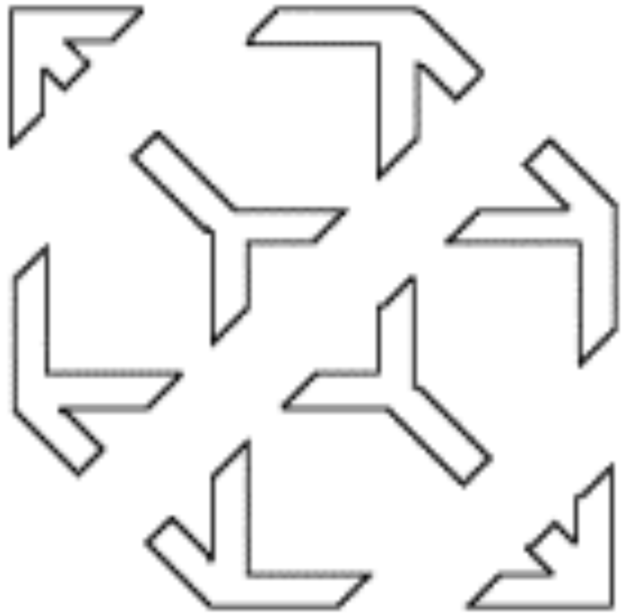


Principle of Continuity

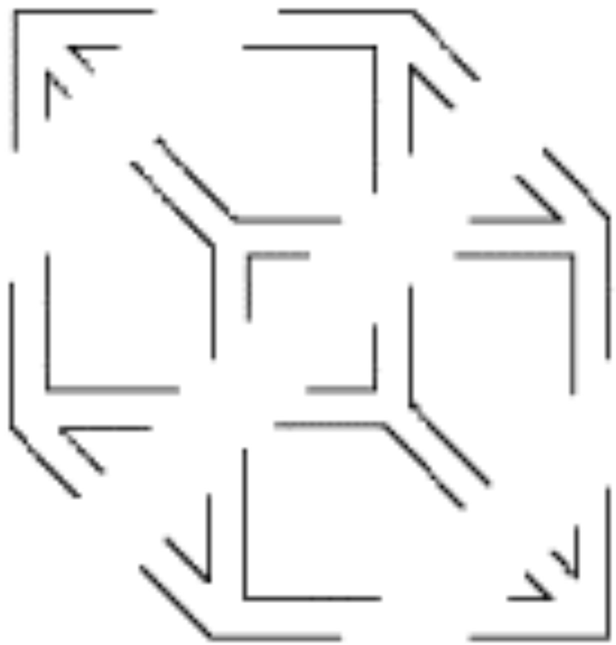
- there is a preference for continuous figures



Principle of Closure



- objects that seem to form closed entities tend to be grouped together

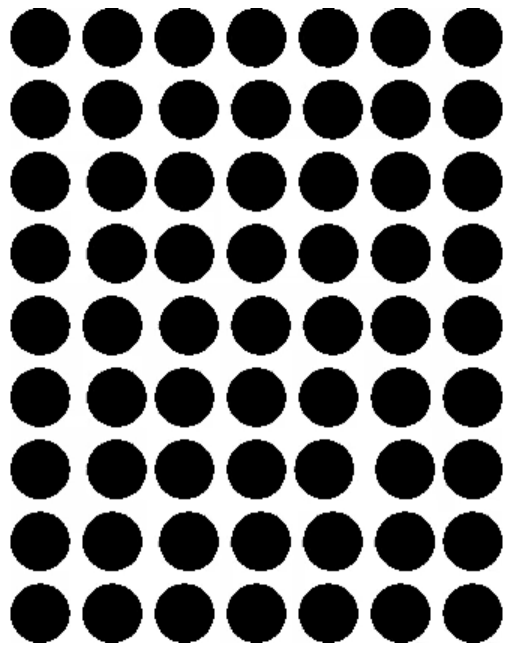


Principle of Closure, Continuity, & Figure-Ground

- the perceptual system fills in the gaps where evidence is incomplete

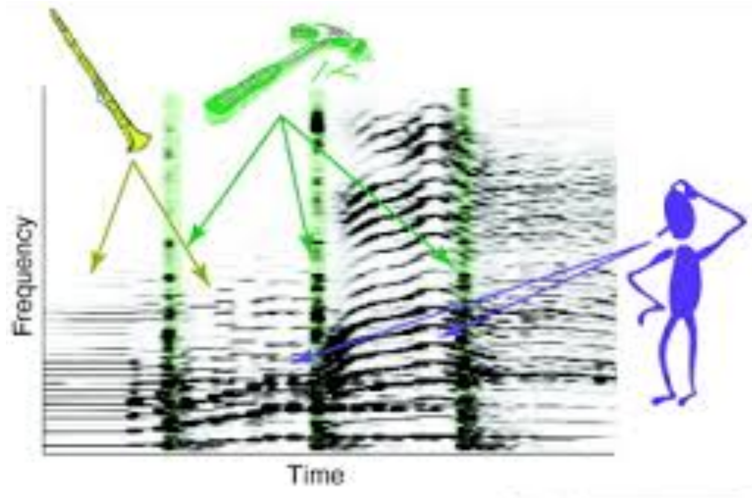
Principle of Common Fate

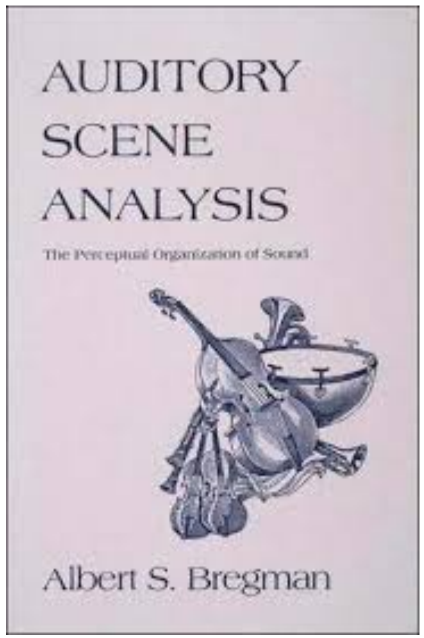
- objects that move together tend to be grouped together



Auditory scene analysis

- three main categories
  - **sequential integration** - putting together events that follow one another in time
  - **spectral integration** - integrating components that occur at the same time in different parts of the spectrum
  - **old-plus-new heuristic** - perceptual continuation of old sound at the presentation of a more complex sound



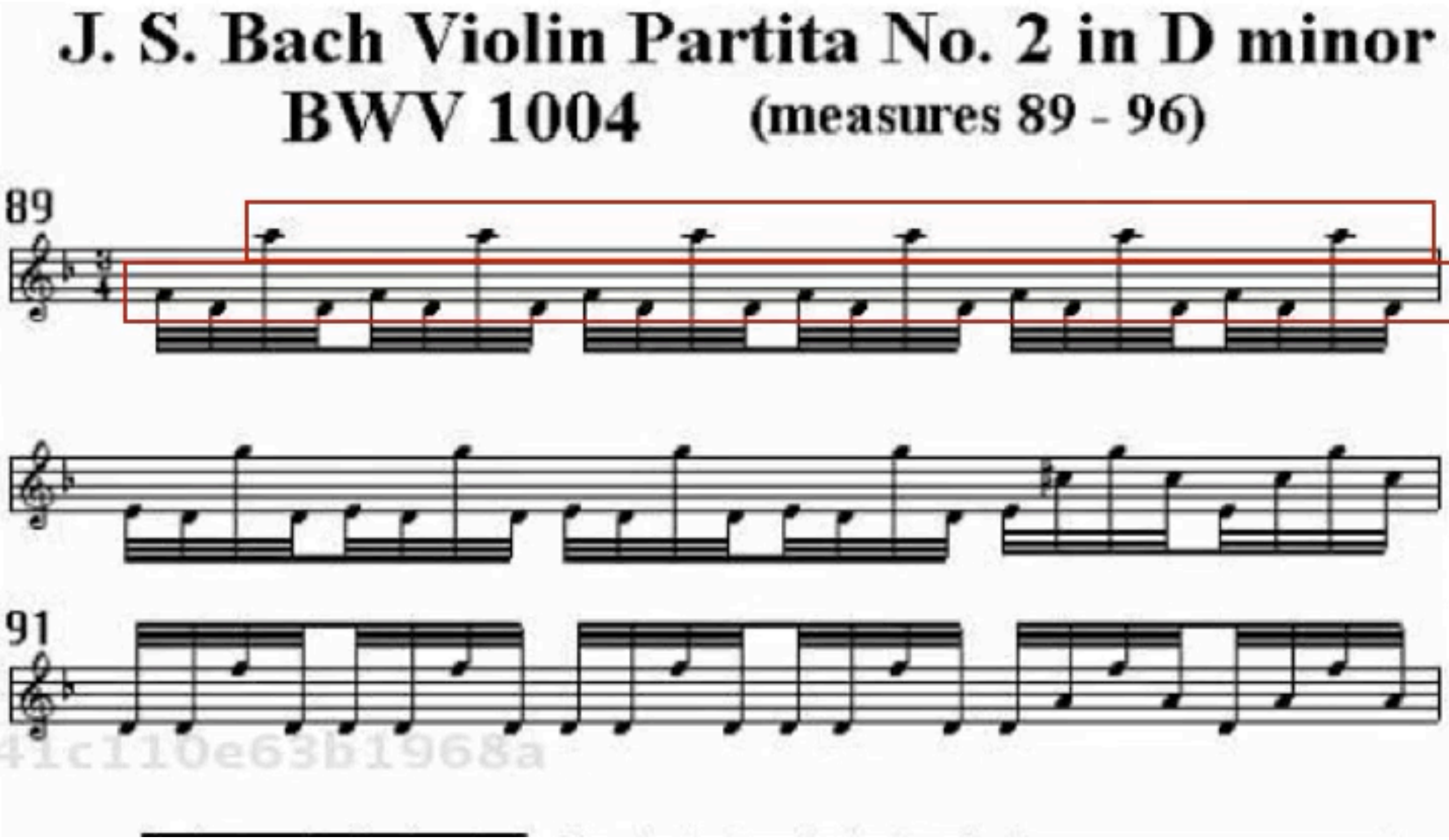


What do you hear?

- On what basis are you segregating them into streams?

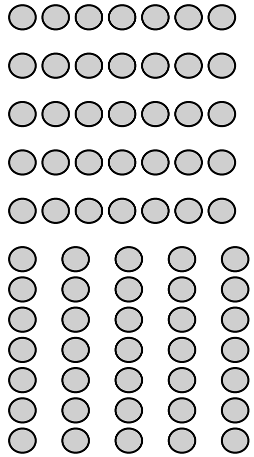
What do you hear?

- Bach: Partita No. 2 BWV 1004



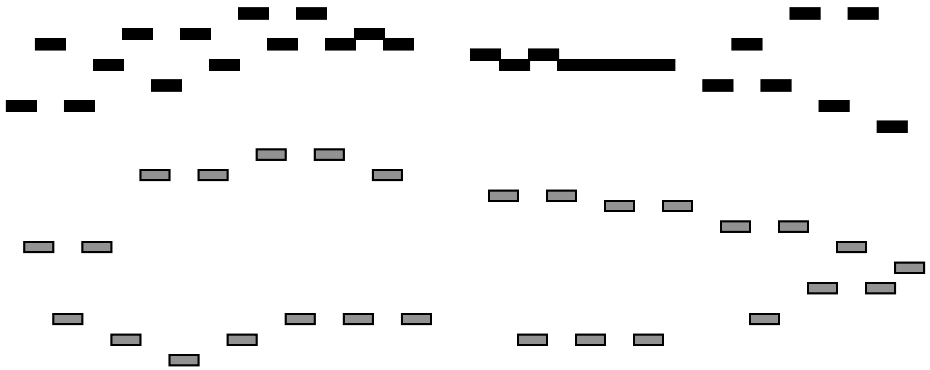
Sequential Integration

Streaming by Proximity



- proximity in auditory domain:
    - proximity in time
    - proximity in pitch
- auditory events that are proximal in pitch and/or time
    - are probably to be created by one and the same source
    - tend to be heard as a stream

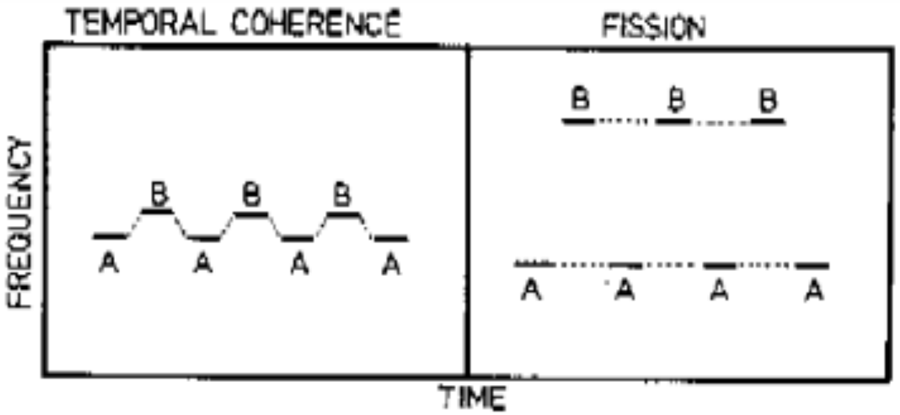
Streaming by Proximity



- interleaved melodies
    - increasing the height difference between melodies results in separation of auditory streams

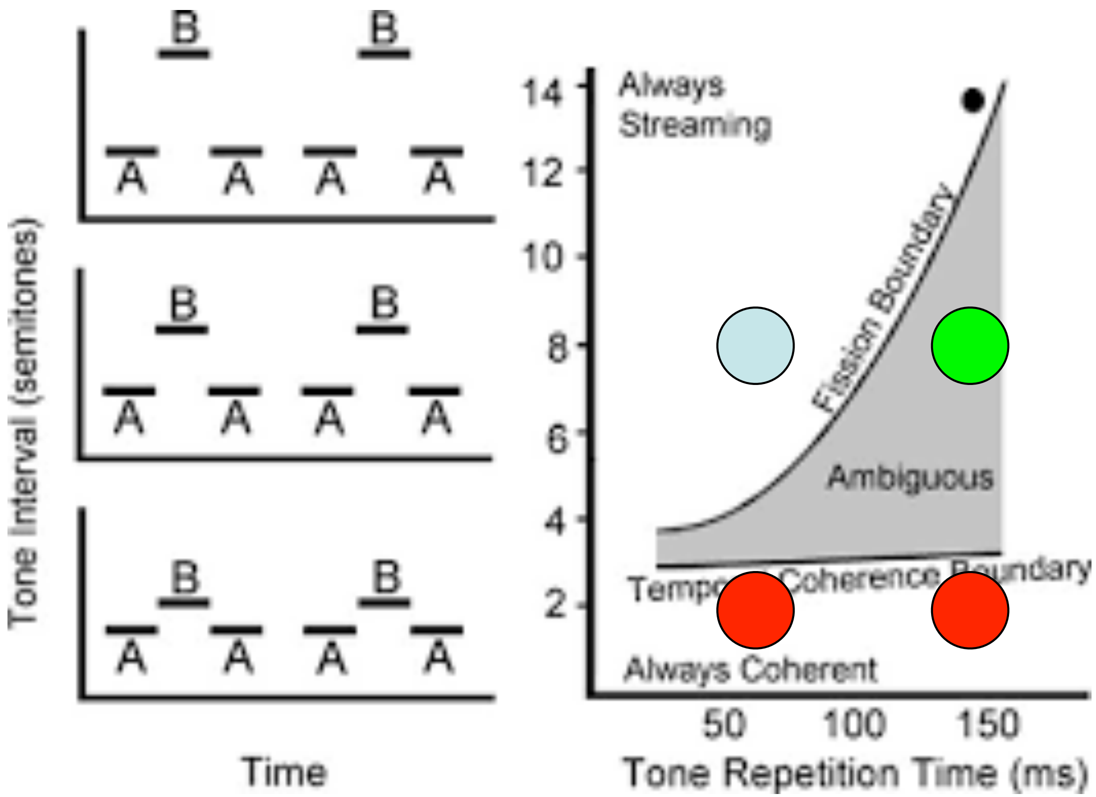
What is proximal enough for streaming to occur?

- van Noorden (1975): dependence of streaming on repetition time and pitch interval
  - Temporal Coherence and Fission
    - occurrence of temporal coherence and fission is determined by the musical interval between the tones

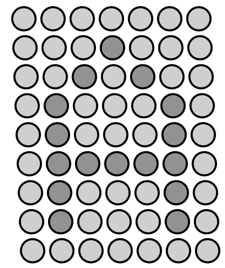


What is proximal enough for streaming to occur?

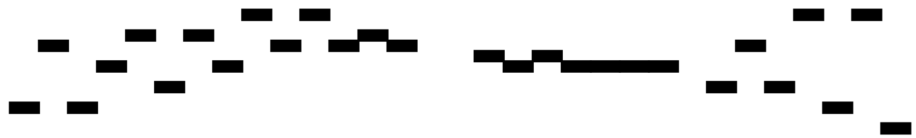
- A: 2 streams
- C: 1 stream
- B: 1-2 streams (depending on attention)



Streaming by similarity

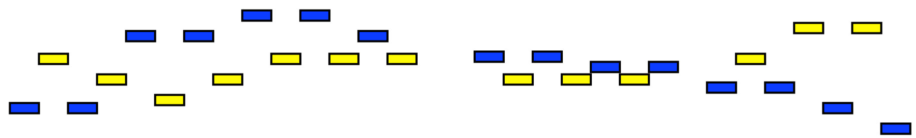


- auditory streaming can be based on
  - similarity in pitch



  - similarity in loudness

  - similarity in timbre



Sequential Integration: Summary

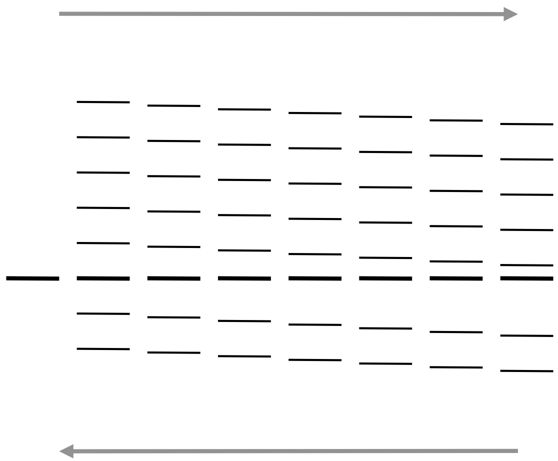
- depends on proximity of
  - time (onset)
  - pitch
- Temporal Coherence and Fission (interval & repetition time)
- depends on the similarity in
  - loudness
  - timbre

Spectral Integration

Spectral Integration

- what kind of principles govern the integration of simultaneous auditory components into perceptual units?

Spectral Relations

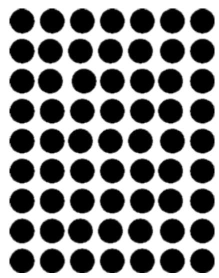


- grouping by harmonicity
  - frequency components sharing the same fundamental are probable to come from the same source and are grouped together

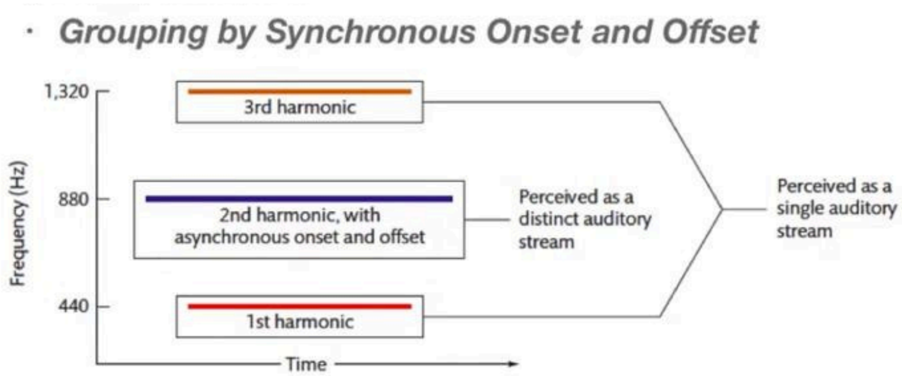
Common Fate

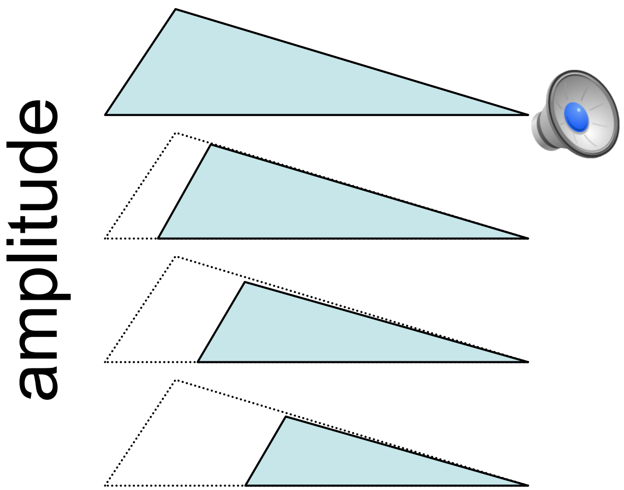
- partials that are in a common harmonic series are more likely to be treated as the spectrum of a single sound.
- when different partials in the spectrum undergo the same change at the same time, they are bound together into a common perceptual unit and segregated from partials whose time-varying behavior is different.
    - this principle applies both to changes in intensity and changes in frequency

Common Fate

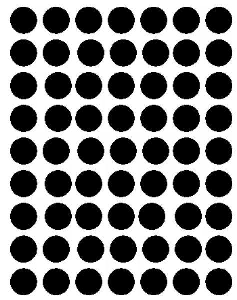


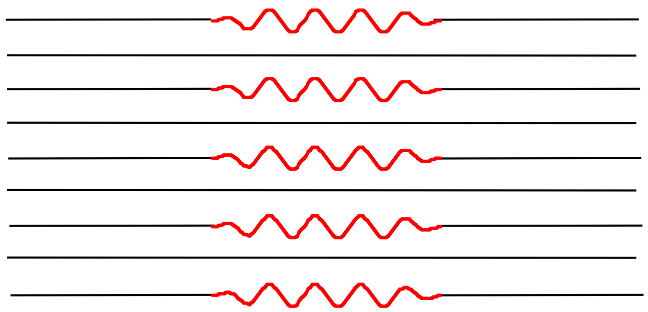
- grouping by onset (and offset)
  - frequency components having proximal onset times are probable to come from the same source and are grouped together





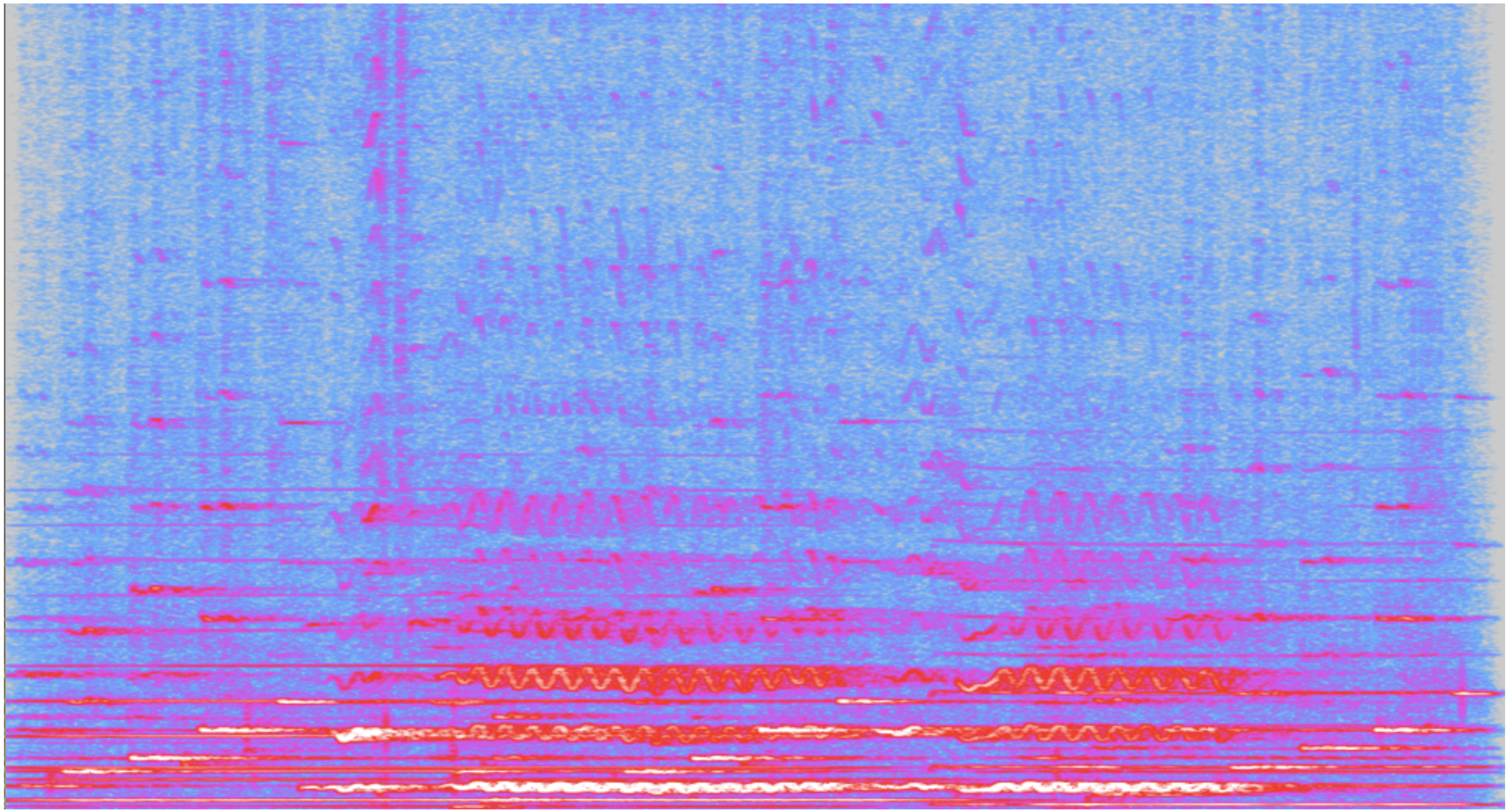
Common Fate: FM





- spectral components sharing the same FM pattern are probably associated to the same source and are grouped together

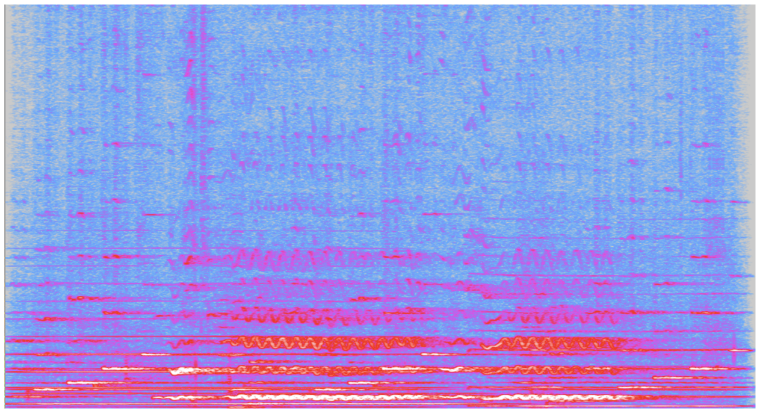
Why does the singer use vibrato?



Why does the singer use vibrato?

- auditory system looks for different sets of harmonic series and segregates them

- to maintain perceptual distinctness is to be producing pitches that are not the same as those produced at the same time by the accompaniment - and hence **stand out**



Old-Plus-New Heuristic

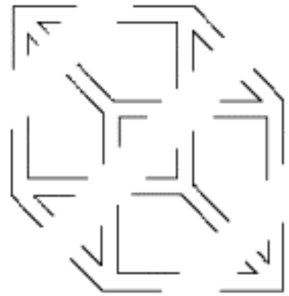
Old-Plus-New Heuristic

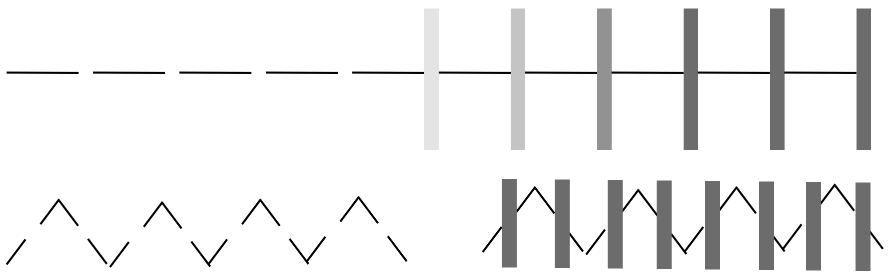
- perceptual continuation of old sound at the presentation of a more complex sound

Old-Plus-New Heuristic

“picket-fence effect”

Apparent continuity through perceptual completion





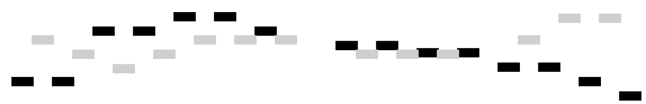
- if a spectrum suddenly becomes more complex or more intense, it is interpreted as the old sound continuing joined by a new one

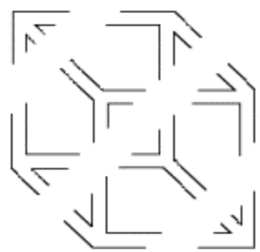
Summary Σ

- Auditory Scene Analysis
  - formation of auditory streams (perceptual grouping)
  - much of it can be explained by Gestalt principles
- main types of ASA processes
  - sequential integration
  - spectral integration
  - old-plus-new heuristic

Summary

- sequential integration
  - proximity in time and/or pitch
  - similarity in loudness/timbre
- spectral integration
  - proximal onset & offset times
  - spectral relations (harmonicity)
  - common fate
- old-plus-new heuristic
  - apparent continuity





Gestalt principles of perceptual organization

Bach was the master of auditory stream segregation

proximity **SIMILARITY**