Introduction lecture, Psychoacoustics

Václav Vencovský

Department of Radioelectronics, Faculty of Electrical Engineering, Czech Technical University in Prague, Czech Republic

Aim of the course (hearing part)

- Teach fundamentals of the physiology of hearing
- Give an opportunity for practical measurements (e.g., threshold of hearing, otoacoustic emissions)
- Present models of the inidividual parts of the auditory pathway

Applications:

- Communication acoustics: telephony, multimedia
- Clinical medicine: differential diagnostics of hearing loss, hearing aids

Practical information and teachers

- Moodle site of the course with materials
  `https://moodle.fel.cvut.cz/course/B0M37FAV`
- Lecturers
  - Václav Vencovský, vencovac@fel.cvut.cz, lectures on hearing, garant of the course
  - Petr Maršálek, petr.marsalek@lf1.cuni.cz, lectures on hearing and vision
  - Miloš Klíma, klima@fel.cvut.cz, lectures on vision
  - Karel Fliegel, fliegek@fel.cvut.cz, lectures on vision
- Training courses and laboratory exercises
  - Václav Vencovský, training courses on hearing
  - Karel Fliegel, training courses on vision
  - Jan Bednář, bednaja4@fel.cvut.cz, training courses on vision
  - Adam Zizien, zizieada@fel.cvut.cz, training courses on vision

Classification

- **B0M37FAV** - Physiology and modeling of hearing and vision
- Classes: 2P + 2C + 4D, grading: ungraded assessment + exam, 6 credits
- Attendence: You must attend training courses (two excused absences are allowed)
- It is neccessary to prepare for the practices beforehand, you will get points if you fulfill practical tasks
- To get the assessment, you must write a test (at least 60%) with fundamental questions from the practices and tutorials
- Final classification will be 20% from the semester (training courses) and 80% is exam
- For exam, you can choose three lecture topics provided by teachers at the end of semestr

Recommended study materials

- Syka J., Vrabec F., Voldřich L. Fyziologie a patofyziologie zraku a sluchu, Avicenum, 1981. (only in czech, cover both, hearing and vision, the book is old hence two more books are recommended)
- Pickles J.O. An introduction to the physiology of hearing. Third edition, Emerald Group Publishing Limited, 2008.
- Wandell B.A. Foundations of vision, Sinauer Associates, 1995,
- Additional study materials:
    - Manley G.A., Gummer A.W., Popper A.N., Fay R.R. (Eds.) Understanding the cochlea, Springer Handbook of Auditory Research, 2017,
    - Lyon R.F. Human and machine hearing. Cambridge University Press, 2018,
    - Kremers J. Human color vision, Springer Verlag, 2016,
    - Zhaoping L. Understanding vision: Theory, models, and data, Oxford, 2018.

Overview of the lecture

- Physical variables: sound intensity level, sound pressure level, frequency, spectral characteristics (e.g., bandwidth)
- Perceived sensations: pitch, loudness, timbre
- Outer and middle ear, anatomy and physiology

Sound

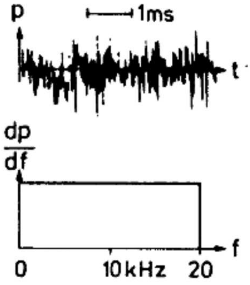
- Acoustic wave in the air: small (relative to the atmospheric pressure) pressure changes (acoustical pressure)
- The acoustic wave can be described by physical variables
- For perception, the relevant variables or physical parameters
  1. sound intensity
  2. frequency for pure tones, or frequencies of spectral components (e.g. harmonic or inharmonic complex tone)
  3. spectral parameters (shape of the spectrum, bandwidth)
  4. temporal parameters (shape of the temporal envelope)

Sound intensity

- "Sound energy transmitted per unit of time in the specified direction through a unit area normal to this direction at the point" (Fletcher and Munson, 1933)
- For plane or spherical free progressive waves (in the direction of the wave progression): $I = \frac{p^2}{\rho c}$ (W/m^2), where p is the acoustic pressure, ρ is the density, and c is the wave speed
- Sound intensity level: $10 \log_{10} \left(\frac{I}{I_{\text{ref}}} \right)$ (dB), where the reference $I_{\text{ref}} = 10^{-12}$ (W/m^2) is roughly a sound pressure (effective value) of 1-kHz tone at the threshold of hearing in normally hearing young humans
- Note: According to Fletcher and Munson (1933): the choice of such a reference was also due to the fact that it was "a simple number which was convenient as a reference for computation work. . . "
- p is the effective value of signal $s(t)$ $p = \sqrt{\frac{1}{T} \int_0^T s(t)^2 dt}$
- Sound pressure level (SPL): $20 \log_{10} \left(\frac{p}{p_{\text{ref}}} \right)$, where $p_{\text{ref}} = 2 \cdot 10^{-5}$

Sound density

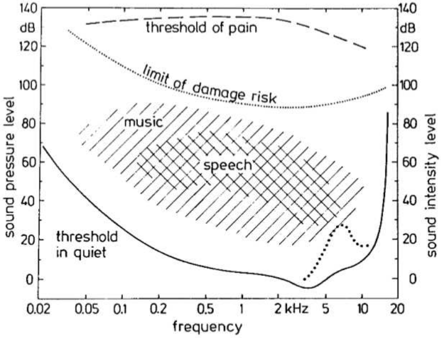
- For noises, it is advantageous to use *sound density* instead of *sound intensity*, i.e., the sound intensity within a bandwidth of 1 Hz
- Sound intensity density level (density level) l is the logarithmic correlate of the sound density
- Density level is independent of frequency for white noise and l and L are related by $L = [l + 10\log_{10}(\Delta f/\mathrm{Hz})]$, where Δf is the noise bandwidth in Hz.



Taken from Fastl and Zwicker (2007): Psychoacoustics: Facts and Models

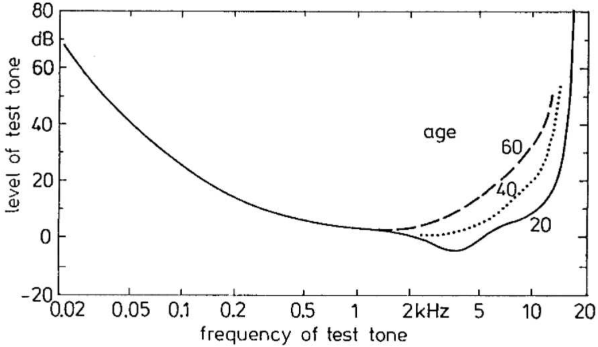
Hearing threshold, dynamic and frequency range

- Humans hear frequencies roughly between 20 Hz to 20 kHz
- Dynamic range of the human hearing system is about 120 dB, which is 10^6 change in the acoustic pressure



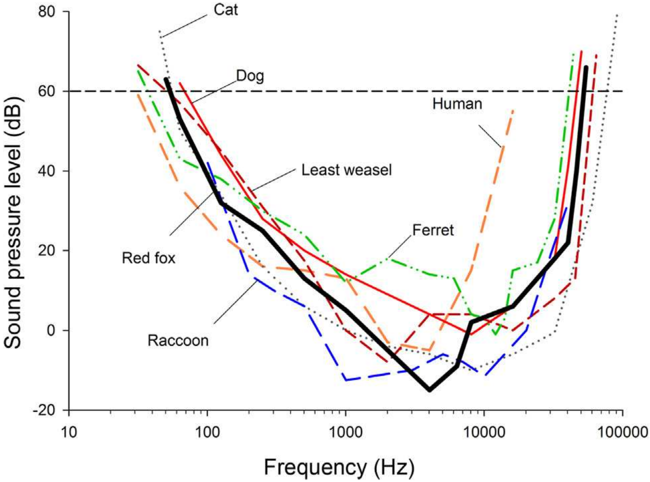
Taken from Fastl and Zwicker (2007) Psychoacoustics: Facts and Models

Hearing threshold and aging



Taken from Fastl and Zwicker (2007) Psychoacoustics: Facts and Models

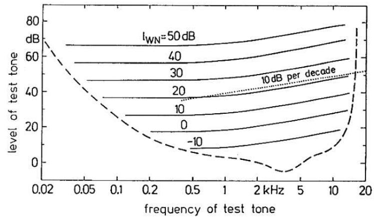
Hearing theshold of other mammals

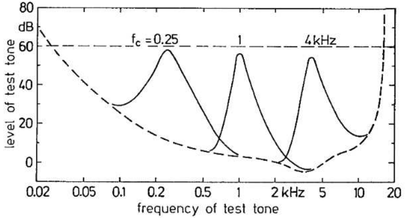


Taken from Malkemper et al. (2015)

Masking

- Hearing threshold of a sound is elevated if another audible (in rare cases also inaudible) sound is present
- This phenomenon is called masking
- Below are shown masked thresholds for a pure tone masked by white noise (broad band noise) and narrow-band noise (presented simultaneously)
- Masking is effective also if the masker preceed or follows a test tone

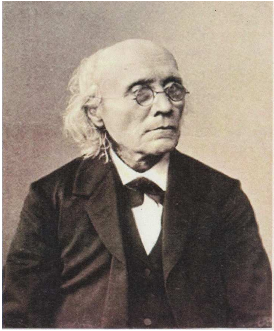




Taken from Fastl and Zwicker (2007) Psychoacoustics: Facts and Models

Psychoacoustics

”Psychoacoustics is a psychophysical study of acoustics” and the term psychophysics was defined by Gustav Fechner **“as the study of the relationship between sensory perception (psychology) and physical variables (physics), e.g. how does perceived loudness (perception) relate to sound pressure level (physical variable)”** (Yost, 2015)



“Father of Psychophysics” Gustav Fechner (1801–1887)

Taken from Yost (2015) Acoustics Today, 11:46–53

History of psychoacoustics

- The relationship between physical attributes of sound (music) and perception was studied by early Greeks
- Pythagoras and others were fascinated by music and tried to understand the physical/mathematical bases of musical scales, consonance nad dissonance
- Aristotle (350 B.C.) suggested that sound is carried by air movement; Leonardo da Vinci ($\approx$ 1500) realized that sound is carried by waves
- Ernst Heinrich Weber (1795–1878) (Fechner's teacher) developed a concept of just noticable difference in psychophysics (the concepts holds for variaty of sensations, not just hearing)
- Weber's law: $\Delta I/I = k$, where ΔI is just-noticable intensity difference, I is the intensity, and k is a constant (Weber's fraction)

History of psychoacoustics

Fechner's law

- Fechner assumed that
  1. the Weber's law is correct, i.e. $\Delta I/I = k$
  2. the just-noticable difference represents basic unit of perceived magnitude (perceived sensation)
- With these assumptions, the perceived magnitude $P = k \int \frac{dI}{I} \propto k \ln I$

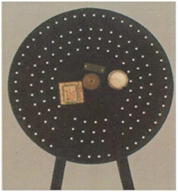
Steven's power law

- Stanley Smith Stevens (1906–1973) proposed that instead of the Fechner's law, the perceived sensation is given by a power law: $P = kI^n$
- Steven's experimentally determined exponents n for different types of sensations

Taken from Yost (2015)

History of psychoacoustics

- Pitch and sound source localization were the main interests of the first researchers who studied sound sensation
- Hermann von Helmholtz (1821–1894) was a commanding scientist in the field (he was influenced by Ohm's "acoustic law" stating that "the ear performs limited Fourier analysis"
- Pitch and timbre were studied by tuning forks and special "sirens"



Taken from Yost (2015)

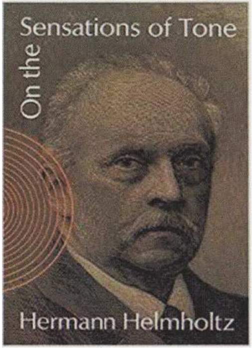
- Time and intensity cues combinged to allow sound source localization were first described by James William Strutt (Lord Rayleigh, 1842–1919)
- For detailed history, please see the review Yost (2015) Acoustics Today 11:46–53

Pitch

- Pitch was most probably the first hearing sensation atribute which was studied (recall already mentioned Greeks and their studies of musical intervals)
- Our ability to distinguish individual tones allowed us to create musical instruments and produce music
- August Seebeck in 1844 constructed a siren which when rotated let the air flow through it and produced a pitch based on serier of harmonic tones
- Harmonic tones had a frequency equal to an integer multiple of a fundamental tone whose frequency coresponded to the perceived pitch
- Also a siren which did not generate fundamental frequency but only the higher harmonics produced the same sensation of pitch

Pitch

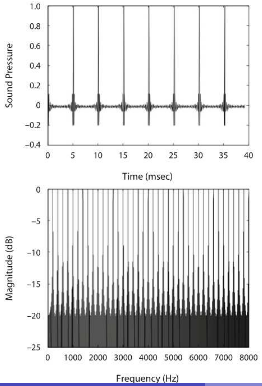
- Pitch of “missing fundamental” sound posed a challange for Helmholtz theory of pitch perception
- Based on Ohm’s acoustics law, Helmholtz suggested that the ear resonates at frequencies present in the harmonic signal and that those partials with largest amplitude determine the resulting pitch
- Helmohltz resonators were more highly tuned at low frequencies which suggested that the lowest frequencies determine the perceived pitch

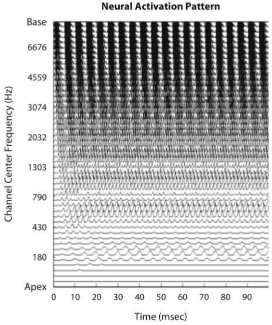


Taken from Yost (2015)

Pitch

- Missing fundamental is present in the temporal envelope of the harmonic signal and hence Shouten (1940) formulated “residue theory” stating that the missing fundamental stimuli exist in the temporal pattern of high frequency filtered signal
- The debate how the ear codes pitch (spectral vs. temporal approach, or mixture of both) still continues (Yost, 2009)





Adapted from Yost (2009)

Ratio pitch

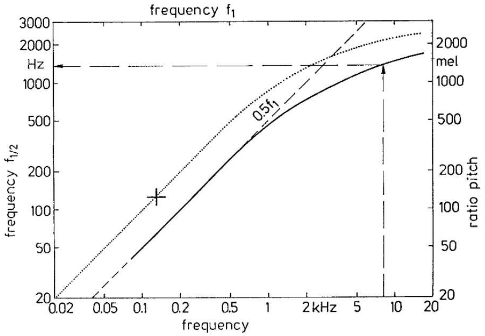
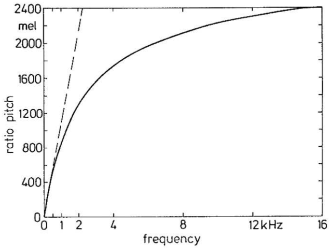


Fig. 5.1. Frequency and ratio pitch. The relationship between the frequency f_1 and the frequency $f_{1/2}$ producing "half pitch" (*solid*). Ratio pitch as a function of frequency (*dotted*). *Cross*: reference 125 Hz = 125 mel. *Dashed lines with arrows:* indication that 1300 Hz corresponds to the "half pitch" of 8 kHz

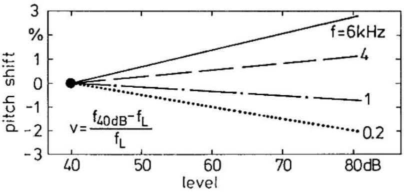


Adapted from Fastl and Zwicker (2007) Psychoacoustics: Facts and Models

- Ratio pitch is related to our sensation of melodies, hence the given unit of ratio pitch is *mel*

Effect of level on pitch

- Pitch of a pure tone is mainly determined by its frequency, however, the tone intensity plays a small role too
- Figure below shows the largest effect at very low and very high frequencies



Adapted from Fastl and Zwicker (2007) Psychoacoustics: Facts and Models

Just noticeable differences in frequency

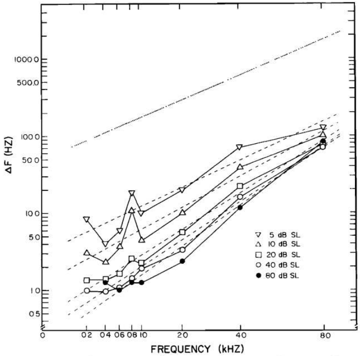


FIG. 1. Mean frequency DL's averaged across subjects. The dotted line represents fit to critical band data from Zwicker, Flottorp, and Stevens (1957).

Taken from Wier et al.(1977)

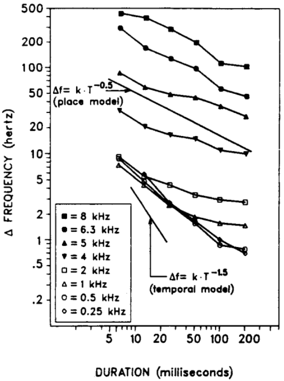


FIGURE 5 Just noticeable frequency differences as a function of stimulus duration. Sinusoidal tones had a constant loudness level of 60 phons. Line segments represent predictions of Siebert's (1970) place model (top, slope = −0.5) and temporal model (bottom, slope = −1.5). (Data from Moore, 1973.)

Taken from Houtsma (1995) Hearing, ed. B.C.J. Moore

Loudness, Loudness level

- Loudness belongs to the category of intensity sensations
- In addition to the sound intensity, loudness depends on bandwidth, frequency content, and duration of the sound signal

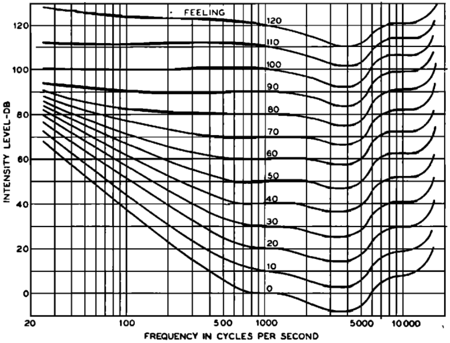
- Accurate studies on loudness were allowed by invention of loudspeakers and microphones which allowed controled presentation of signals
- A bunch of work towards our understanding of loudness sensation was done in the first half of the 20th century at Bell Telephone Laboratories (research organization of AT&T) and is connected with the work of Harvey Fletcher

Harvey Fletcher (1884–1981), a physicist, fundamental achievements in communication acoustics: loudness, masking, critical bandwidth, speech (articulation index), artificial laryngs, audiometry, hearing aids, stereophonic sound

Yost (2015)

Loudness level

- Different sounds can be matched in loudness, i.e., intensity of one sound is altered in order to make it as loud as a refence
- This technique is called "loudness matching" and can be used to measure so called loudness level (introduced about 100 years ago by Barkhausen)
- "Loudness level of a sound is equal to the sound pressure level of a 1-kHz tone in a plane wave and frontal incident that is as loud as the sound" (Fastl and Zwicker, 2008)



Taken from Fletcher and Munson (1933)

Revised measurement of equal-loudness contours

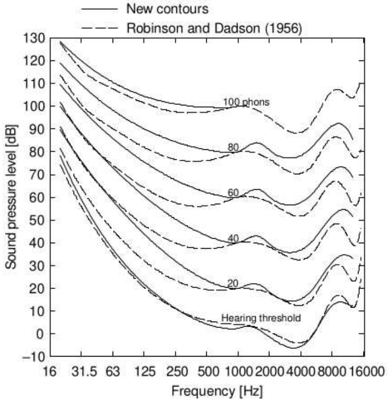


Fig. 4: Estimated new contours and the contour proposed by Robinson and Dadson[1] (former standard)

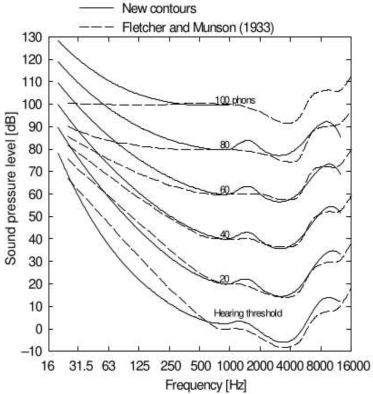
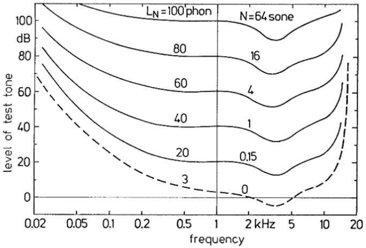


Fig. 5: Estimated new contours and the contour proposed by Fletcher and Munson [3]

Taken from ISO 226:2003

Loudness

- *Loudness matching* experiments provide indirect measure of sensation, i.e. loudness level does not say how much louder is a certain sound than a reference
- Direct measurement can be achieved for example by the method of *magnitude estimation* which asks the listeners to assign number to a stimuli of different intensities
- Variations of the method exists, e.g. a listener is presented a sound and asked to adjust the intensity of another sound to achive required loudness ratio
- Stevens (1957) used such techniques to develop *loudness scales*, which show the relationship between perceived loudness and intensity
- Stevens gave the loudness a unit called *sone* and defined that the standard is 1-kHz, 40-dB SPL tone, which gives the loudness sensation of 1 sone.



- For pure tones, Stevens obtained a mathematical relationship between loudness and intensity which followed power law $S = kI^{0.3}$
- The relationship holds for levels above about 40 dB above hearing threshod

Fastl and Zwicker (2007) Psychoacoustics: Facts and Models

Effect of level on loudness

- Fletcher and Munson also found how an increase of intensity changes the loudness
- At 1 kHz, above 40 dB SPL the pure tone loudness is proportional to the cube root of the signal intensity: p scaled by $2^{3/2} \approx 9\,\mathrm{dB}$
- Below 40 dB SPL, Fletcher previously assumed that loudness was found to be proportional to the intensity p scaled by $2^{1/2} \approx 3\,\mathrm{dB}$, however, the measured data show larger variablity

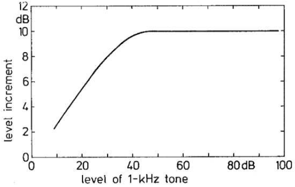


Fig. 8.3. Level increment (or decrement) necessary to produce a doubling (or halving) of the loudness of a 1-kHz tone as a function of its level; Track 30

Fastl and Zwicker (2007) Psychoacoustics: Facts and Models

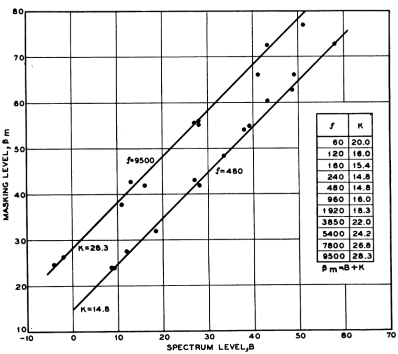
Additivy of loudness

- Fletcher and Munson (1933) studied how the perceived loudness of multiple stimuli combine
- This concept is called *loudness additivity*
- When equally loud tones were presented together their loudness was twice as loud
- Loudness additivity holds also for more than two tones, until the tones mask each other
- Loudness additivity also holds for tones played to different ears
- Exceptions exists, e.g. for speech in noise

Critical bands

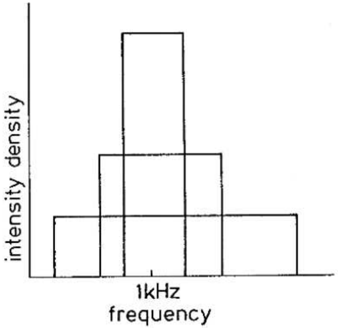
- During his loudness experiments, Flatcher noticed that the hearing system is frequency selective
- Fletcher noticed that broadband noise masks differently tones of different frequencies

- *critical ratio*: between the pure tone level and spectral density of the noise
- *critical bandwidth*: equivalent rectangular bandwidth for an auditory filter (nicely summarized in Allen (1996))
- The similar results were obtained by Fletcher when a bandpass noise was used as a masker and its bandwidth was varied
- Bandwidths of noise larger than the *critical bandwidth* did not increase masked threshold



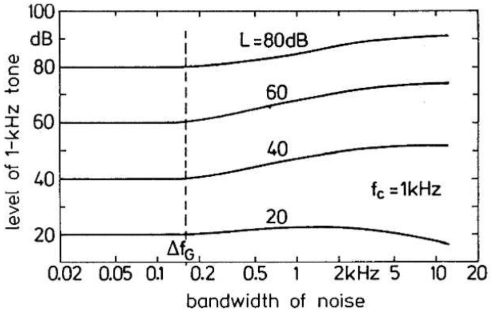
Taken from Fletcher (1938)

Loudness of bandpass noise



Spectrum of bandpass noise whose sound pressure (intensity) level is constant

Taken from Fastl and Zwicker (2007) Psychoacoustics: Facts and Models



Level of a 1-kHz tone which evoked the same loundess as a band-pass noise with center frequency 1 kHz

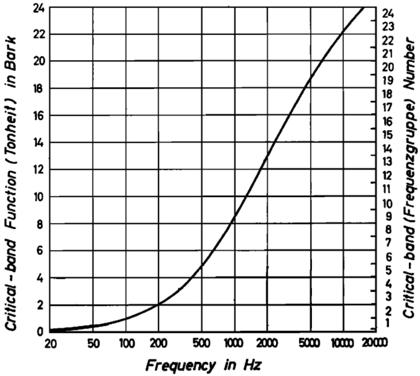
Taken from Fastl and Zwicker (2007) Psychoacoustics: Facts and Models

Critical bandwidth, Barks

Karl Eberhard Zwicker (1924–1990) measured critical bandwidth and determined a scale with units *Barks* (according to Heinrich Georg Barkhausen)

Number	Center frequencies Hz	Cut-off frequencies Hz	Bandwidth Hz
		20	
1	50	100	80
2	150	200	100
3	250	300	100
4	350	400	100
5	450	510	110
6	570	630	120
7	700	770	140
8	840	920	150
9	1000	1080	160
10	1170	1270	190
11	1370	1480	210
12	1600	1720	240
13	1850	2000	280
14	2150	2320	320
15	2500	2700	380
16	2900	3150	450
17	3400	3700	550
18	4000	4400	700
19	4800	5300	900
20	5800	6400	1100
21	7000	7700	1300
22	8500	9500	1800
23	10 500	12 000	2500
24	13 500	15 500	3500

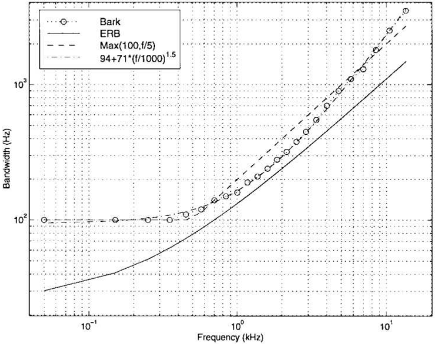
Taken from Zwicker (1961)



Taken from Zwicker (1961)

Bark vs Cams scale

- Glasberg and Moore (1983) derived auditory filter shapes and their equivalent rectangular bandwidth (ERB) from notched-noise masking data
- Their data and resulting scale (Cams scale due to their affiliation to the University of Cambridge) is comapred with bandwidths of Zwicker



Taken from Smith and Abel (1999)

Just noticeable level differences

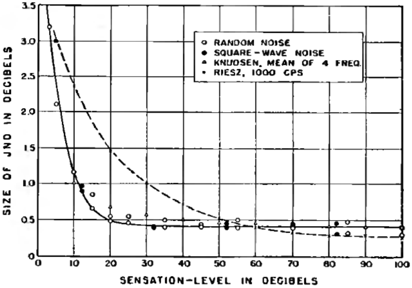
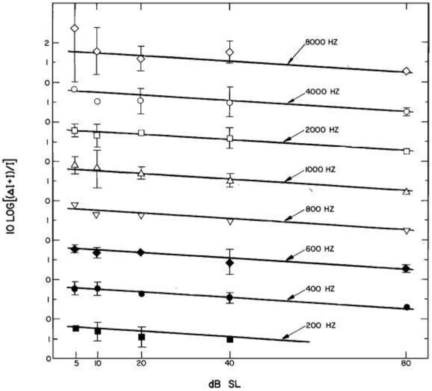


FIG. 3. Increments in intensity heard 50 percent of the time are plotted as a function of the intensity of the noise in decibels above the threshold of hearing. Data for tones are presented for purposes of comparison. The solid line represents Eq. (2).

Taken from Miller (1947)



Taken from Jesteadt et al. (1977)

Díla, je jeho povinností zdržet se užívání Díla a je povinen toto Dílo trvale odstranit včetně veškerých kopií (elektronické, tiskové, vizuální, audio a zhotovených jiným způsobem) z elektronického zařízení a všech záznamových zařízení, na které jej Držitel díla umístil.