

Psychoacoustics: How Your Brain Decodes Sound

Authored by
mohammad looti

June 14, 2026

RECOMMENDED CITATION

mohammad looti (2026). *Psychoacoustics: How Your Brain Decodes Sound*.
PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=38198>

Psychoacoustics is the scientific study of sound perception. More specifically, it is the 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. Psychoacoustics received its name from a field within psychology--i.e., recognition science--which deals with all kinds of human perceptions. It is an interdisciplinary field of many areas, including psychology, acoustics, electronic engineering, physics, biology, physiology, and computer science.

Background

Hearing is not a purely mechanical phenomenon of wave propagation, but is also a sensory and perceptual event; in other words, when a person hears something, that something arrives at the ear as a mechanical sound wave traveling through the air, but within the ear it is transformed into neural action potentials. These nerve pulses then travel to the brain where they are perceived. Hence, in many problems in acoustics, such as for audio processing, it is advantageous to take into account not just the mechanics of the environment, but also the fact that both the ear and the brain are involved in a person's listening experience.

The inner ear, for example, does significant signal processing in converting sound waveforms into neural stimuli, so certain differences between waveforms may be imperceptible. Data compression techniques, such as MP3, make use of this fact. In addition, the ear has a nonlinear response to sounds of different intensity levels; this nonlinear response is called loudness. Telephone networks and audio noise reduction systems make use of this fact by nonlinearly compressing data samples before transmission, and then expanding them for playback. Another effect of the ear's nonlinear response is that sounds that are close in frequency produce phantom beat notes, or intermodulation distortion products.

The term "psychoacoustics" also arises in discussions about cognitive psychology and the effects that personal expectations, prejudices, and predispositions may have on listeners' relative evaluations and comparisons of sonic aesthetics and acuity and on listeners' varying determinations about the relative qualities of various musical instruments and performers. The expression that one "hears what one wants (or expects) to hear" may pertain in such discussions.

Limits of perception

An equal-loudness contour. Note peak sensitivity around 2-4 kHz, the frequency around which the human voice centers.

The human ear can nominally hear sounds in the range 20 Hz (0.02 kHz) to 20,000 Hz (20 kHz). The upper limit tends to decrease with age; most adults are unable to hear above 16 kHz. The

lowest frequency that has been identified as a musical tone is 12 Hz under ideal laboratory conditions. Tones between 4 and 16 Hz can be perceived via the body's sense of touch.

Frequency resolution of the ear is 3.6 Hz within the octave of 1000 - 2000 Hz. That is, changes in pitch larger than 3.6 Hz can be perceived in a clinical setting. However, even smaller pitch differences can be perceived through other means. For example, the interference of two pitches can often be heard as a repetitive variation in volume of the tone. This amplitude modulation occurs with a frequency equal to the difference in frequencies of the two tones and is known as beating.

The semitone scale used in Western musical notation is not a linear frequency scale but logarithmic. Other scales have been derived directly from experiments on human hearing perception, such as the mel scale and Bark scale (these are used in studying perception, but not usually in musical composition), and these are approximately logarithmic in frequency at the high-frequency end, but nearly linear at the low-frequency end.

The intensity range of audible sounds is enormous. Human ear drums are sensitive to variations in the sound pressure, and can detect pressure changes from as small as a few micropascals (μPa) to greater than 100 kPa. For this reason, sound pressure level is also measured logarithmically, with all pressures referenced to 20 μPa (or 1.97385×10^{-10} atm). The lower limit of audibility is therefore defined as 0 dB, but the upper limit is not as clearly defined. The upper limit is more a question of the limit where the ear will be physically harmed or with the potential to cause noise-induced hearing loss.

A more rigorous exploration of the lower limits of audibility determines that the minimum threshold at which a sound can be heard is frequency dependent. By measuring this minimum intensity for testing tones of various frequencies, a frequency dependent absolute threshold of hearing (ATH) curve may be derived. Typically, the ear shows a peak of sensitivity (i.e., its lowest ATH) between 1-5 kHz, though the threshold changes with age, with older ears showing decreased sensitivity above 2 kHz.

The ATH is the lowest of the equal-loudness contours. Equal-loudness contours indicate the sound pressure level (dB SPL), over the range of audible frequencies, that are perceived as being of equal loudness. Equal-loudness contours were first measured by Fletcher and Munson at Bell Labs in 1933 using pure tones reproduced via headphones, and the data they collected are called Fletcher-Munson curves. Because subjective loudness was difficult to measure, the Fletcher-Munson curves were averaged over many subjects.

Robinson and Dadson refined the process in 1956 to obtain a new set of equal-loudness curves for a frontal sound source measured in an anechoic chamber. The Robinson-Dadson curves were standardized as ISO 226 in 1986. In 2003, ISO 226 was revised as equal-loudness contour using

data collected from 12 international studies.

Sound localization

Sound localization is the process of determining the location of a sound source. The brain utilizes subtle differences in loudness, tone and timing between the two ears to allow us to localize sound sources. Localization can be described in terms of three-dimensional position: the azimuth or horizontal angle, the zenith or vertical angle, and the distance (for static sounds) or velocity (for moving sounds). Humans, as most four-legged animals, are adept at detecting direction in the horizontal, but less so in the vertical due to the ears being placed symmetrically. Some species of owls have their ears placed asymmetrically, and can detect sound in all three planes, an adaptation to hunt small mammals in the dark.

Masking effects

Audio masking graph

Suppose a listener cannot hear a given acoustical signal under silent condition. When a signal is playing while another sound is being played (a masker) the signal has to be stronger for the listener to hear it. The masker does not need to have the frequency components of the original signal for masking to happen. A masked signal can be heard even though it is weaker than the masker. Masking happens when a signal and a masker are played together. It also happens when a masker starts after a signal stops playing. The effects of backward masking is weaker than forward masking. The masking effect has been widely used in psychoacoustical research. With masking you can change the levels of the masker and measure the threshold, then create a diagram of a psychophysical tuning curve that will reveal similar features. Masking effects are also used for audio encoding. The masking effect is used in lossy encoders. It can eliminate some of the weaker sounds, so the listener can not hear the difference. This technique has been used in MP3's.

Missing fundamental

When presented with a harmonic series of frequencies in the relationship $2f$, $3f$, $4f$, $5f$, etc. (where f is a specific frequency), humans tend to perceive that the pitch is f .

Software

Perceptual audio coding uses psychoacoustics-based algorithms.

The psychoacoustic model provides for high quality lossy signal compression by describing which parts of a given digital audio signal can be removed (or aggressively compressed) safely--that is, without significant losses in the (consciously) perceived quality of the sound.

It can explain how a sharp clap of the hands might seem painfully loud in a quiet library, but is hardly noticeable after a car backfires on a busy, urban street. This provides great benefit to the overall compression ratio, and psychoacoustic analysis routinely leads to compressed music files that are 1/10th to 1/12th the size of high quality masters, but with discernibly less proportional quality loss. Such compression is a feature of nearly all modern lossy audio compression formats. Some of these formats include Dolby Digital (AC-3), MP3, Opus, Ogg Vorbis, AAC, WMA, MPEG-1 Layer II (used for digital audio broadcasting in several countries) and ATRAC, the compression used in MiniDisc and some Walkman models.

Psychoacoustics is based heavily on human anatomy, especially the ear's limitations in perceiving sound as outlined previously. To summarize, these limitations are:

High frequency limit

Absolute threshold of hearing

Temporal masking

Simultaneous masking

Given that the ear will not be at peak perceptive capacity when dealing with these limitations, a compression algorithm can assign a lower priority to sounds outside the range of human hearing. By carefully shifting bits away from the unimportant components and toward the important ones, the algorithm ensures that the sounds a listener is most likely to perceive are of the highest quality.

Music

Psychoacoustics includes topics and studies that are relevant to music psychology and music therapy. Theorists such as Benjamin Boretz consider some of the results of psychoacoustics to be meaningful only in a musical context.

Irv Teibel's Environments series LPs (1969-79) are an early example of commercially available sounds released expressly for enhancing psychological abilities.

Applied psychoacoustics

Psychoacoustics Model

Psychoacoustics has long enjoyed a symbiotic relationship with computer science, computer engineering, and computer networking. Internet pioneers J. C. R. Licklider and Bob Taylor both completed graduate-level work in psychoacoustics, while BBN Technologies originally specialized in consulting on acoustics issues before it began building the first packet-switched computer networks.

Licklider wrote a paper entitled "A duplex theory of pitch perception".

Psychoacoustics is applied within many fields from software development, where developers map proven and experimental mathematical patterns; in digital signal processing, where many audio compression codecs such as MP3 and Opus use a psychoacoustic model to increase compression ratios; in the design of (high end) audio systems for accurate reproduction of music in theatres and homes; as well as defense systems where scientists have experimented with limited success in creating new acoustic weapons, which emit frequencies that may impair, harm, or kill. It is also applied today within music, where musicians and artists continue to create new auditory experiences by masking unwanted frequencies of instruments, causing other frequencies to be enhanced. Yet another application is in the design of small or lower-quality loudspeakers, which can use the phenomenon of missing fundamentals to give the effect of bass notes at lower frequencies than the loudspeakers are physically able to produce (see references).