

Vox.

78
in: Manual of Phonetics ed. L. Kaiser
1957 Amsterdam S. 78-79
c. Masking

Simultaneous Sounds. In general, the simultaneous presence of two sounds reduces the sensitivity of the ear to one or both of the sounds. In some cases all subjective traces of one of them may even be eliminated. This phenomenon is known as masking. It can be evaluated by measuring the pressure level of single frequency tones when they just become detectable in the presence of the masking sound and again under quiet conditions. A plot of these differences, or threshold shifts, against frequency is called a masking pattern or masking spectrum.

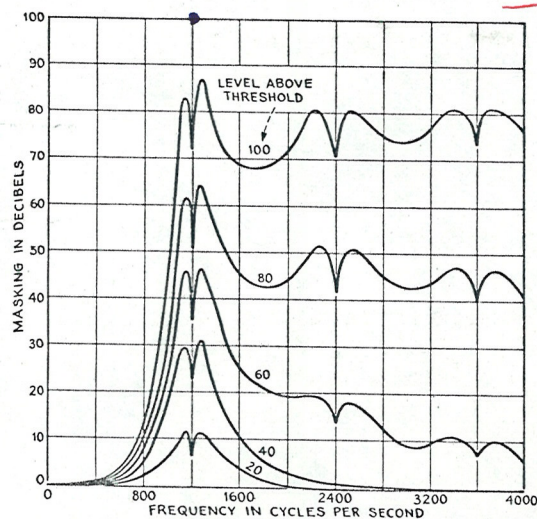


Fig. 9. Masking effect of a 1200 cycle tone on tones of other frequencies (WEGEL and LANE [47]).

broad bands of noise. Given the pressure spectra of such sounds, their masking patterns can be derived. The necessary relationships were established by FLETCHER and MUNSON [48] and supported by data reported later by HAWKINS and STEVENS [49].

Residual Masking. Instead of ceasing abruptly at the cessation of a sound stimulus, masking continues for a further period [50], [51], [52]. Diminishing with time, the magnitude of this residual masking or auditory fatigue, also depends upon other factors, including the frequency, intensity and duration of the stimulus and the frequency at which the masking is measured. With short sounds of moderate intensity the effect becomes negligible within a few tenths of a second. However, it may be large enough to be a factor in the recognition of

The masking patterns produced by a 1200 cycle tone, as measured by WEGEL and LANE [47], are shown on Fig. 9. These patterns display two features of masking that are particularly worthy of note: (1) the chief effect is on sounds of higher frequency than the frequency of the masking sound, and (2) high level sounds can cause large amounts of masking over a broad frequency range.

The masking sounds of ordinary life are usually

low intensity sounds in speech, where such sounds follow immediately after more intense sounds.

d. Pitch

Pitch is a term used to designate one of the attributes of auditory sensations. Although this attribute is hard to describe, it is generally associated with subjective impressions of highness or lowness.

Pure Tones. With pure tones, high pitch is commonly associated with tones of high frequency, low pitch with tones of low frequency. STEVENS and VOLKMANN [53] have shown that this attribute is not directly proportional to frequency, i.e., the sensation produced by a tone of say 2000 cycles does not appear to be twice as high as the sensation caused by a 1000 cycle tone. This lack of proportionality led them to establish a pitch scale. The unit chosen for designating pitch was called the mel and the pitch of a 1000 cycle tone 40 db above threshold was arbitrarily assigned a value of 1000 mels. Using this tone as a starting point, two procedures were used to establish the scale. In one, observers were asked to adjust the frequency of a second tone (also 40 db above threshold) until its pitch seemed to them to be "half as high" as the pitch of the reference tone. This test thus gave the frequency of a tone which has a pitch of 500 mels.

Similar tests gave the frequencies corresponding to other pitches such as 250 and 2000 mels. In another type of test, the frequency of a tone was adjusted until its pitch seemed "mid-way" between the pitches of two other tones. These two procedures gave consistent results which were combined to give the relationship between pitch and frequency shown in Fig. 10.

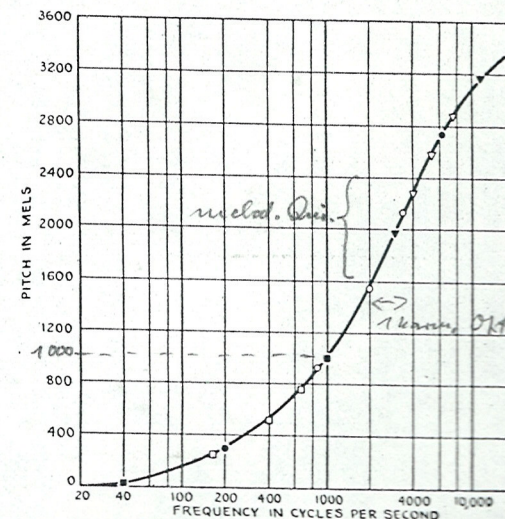


Fig. 10. Relation between pitch and the frequency of pure tones 40 db above threshold (STEVENS and VOLKMANN [53]).

low intensity sounds in speech, where such sounds follow immediately after more intense sounds.

d. Pitch

Pitch is a term used to designate one of the attributes of auditory sensations. Although this attribute is hard to describe, it is generally associated with subjective impressions of highness or lowness.

Pure Tones. With pure tones, high pitch is commonly associated with tones of high frequency, low pitch with tones of low frequency. STEVENS and VOLKMANN [53] have shown that this attribute is not directly proportional to frequency, i.e., the sensation produced by a tone of say 2000 cycles does not appear to be twice as high as the sensation caused by a 1000 cycle tone. This lack of proportionality led them to establish a pitch scale. The unit chosen for designating pitch was called the mel and the pitch of a 1000 cycle tone 40 db above threshold was arbitrarily assigned a value of 1000 mels. Using this tone as a starting point, two procedures were used to establish the scale. In one, observers were asked to adjust the frequency of a second tone (also 40 db above threshold) until its pitch seemed to them to be "half as high" as the pitch of the reference tone. This test thus gave the frequency of a tone which has a pitch of 500 mels. Similar tests gave the frequencies corresponding to other pitches such as 250 and 2000 mels. In another type of test, the frequency of a tone was adjusted until its pitch seemed "mid-way" between the pitches of two other tones. These two procedures gave consistent results which were combined to give the relationship between pitch and frequency shown in Fig. 10.

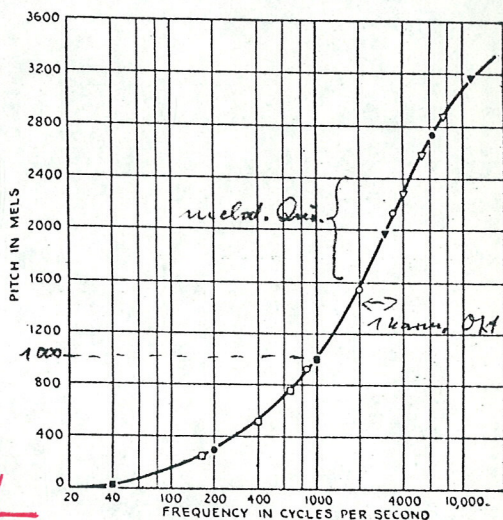


Fig. 10. Relation between pitch and the frequency of pure tones 40 db above threshold (STEVENS and VOLKMANN [53]).

Dia
4/28
vpl.
15/12

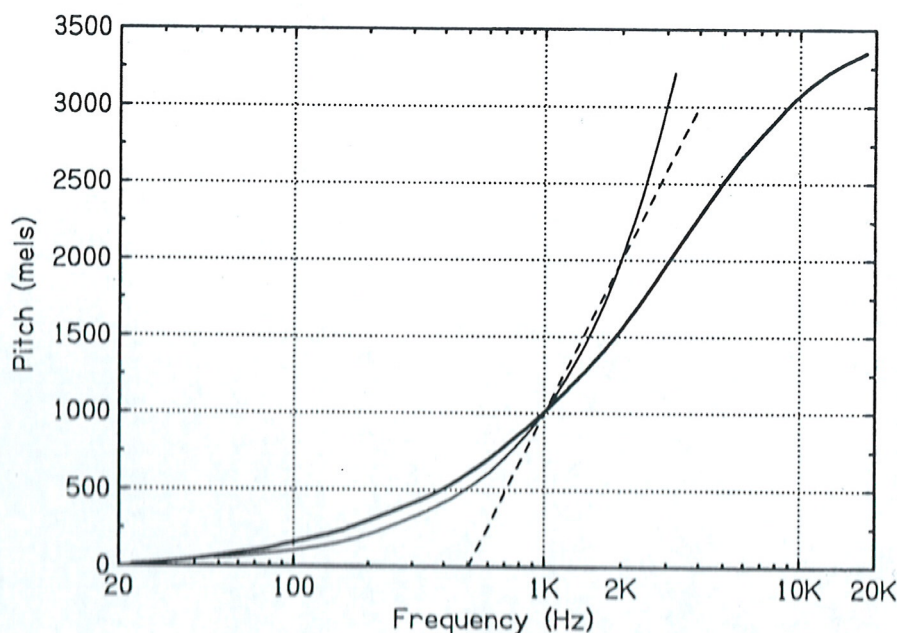


FIGURE 12.7. The heavy line is the mel scale from Stevens and Volkman (1940). The thinner line corresponds to pitch given by frequency in Hz, and the dashed line shows pitch proportional to octave number (the musical scale).

to a difference of $1550 - 1000 = 550$ mels.

Zwicker and his colleagues (Zwicker and Fastl, 1990) invented a different mel scale, with 125 Hz set equal to 125 mels and spanning about 2400 mels so that 100 mels corresponded to one Bark. A comparison with physiological observations suggested that both scales might correspond to distances along the basilar membrane for points of maximum excitation, such that $100 \text{ mels} = 1 \text{ Bark} \approx 1 \text{ mm}$ from apex to stapes. A similar calculation by Zwislocki (1965) concluded that one mel corresponded to 12 primary auditory neurons and that one critical band corresponded to 1300.

Because the critical band number (Bark) is a measure of auditory resolution, usually measured by masking in some form, the correspondence between critical band and cochlear distances seems natural. Masking occurs when excitation patterns collide. The correspondence between pitch and place, while appealing, is less necessary on logical grounds; it is more a matter of faith.

The mel scale has frequently been criticized because it contradicts the use of pitch in musical practice (e.g. Attneave and Olson, 1971). The review chapter on pitch scales by Burns and Ward (1982) does not even mention the mel scale. It has also been criticized on psychoacoustical and physiological grounds by Greenwood (1990, 1991, 1995) who argued that the cochlear map, the critical band function, and the scale of pitch should all be closer to logarithmic, with less flattening at low frequency. The Cam scale comes closer to Greenwood's (1961) critical band than does the Bark scale (see Chapter 10).

The scaling of pitch that leads to the mel scale is sensitive to many experimental

Erpingyuan Dia 4/28, 4/29

*Hartmann, William H. (1998)
Speed, Sound and Satisfaction, (N.Y.C.)
New York: H.W. Lee.*

~ nonsense!