THE ABSOLUTE SEMITONE SCALE

KULLO VENDS

Department of English
Tallinn Teacher Training Institute
Tallinn, Estonia, USSR 200102

ABSTRACT

The absolute semitone scale is a scale combining the properties of both physical and perceptual units. It is derived from a modified Fletcher's formula,

\[ P (at) = 12 \log_2 F_0 (Hz), \]

modified to relate fundamental frequency to its correlate perceptual units of pitch, viz., semitones above 1 Hz (1 Hz = 0 at). The absolute pitch units are much more convenient than Hz for the presentation, comparison, and perceptual units. It is derived from a' rectangular scale (by analogy with pitch differences) and other processing of raw data obtained in instrumental prosodic research.

In prosodic research the presentation of absolute pitch units may be far more convenient than Hz for the analysis of fundamental pitch, viz., the perception of a 20 Hz difference, say, between two measured frequencies, 150 Hz and 200 Hz. Instead, one can correlate perceptual units of pitch viz., the perception of a 50 Hz difference, say, between the lower two frequencies is 4 semitones above 1 Hz (1 Hz = 0 at). The absolute pitch unit is more convenient than Hz for the presentation, comparison, and perceptual units of pitch. Proceeding from Fletcher (1929) who introduced a scale of semitones and centioctaves above 1 Hz for the perception of a tone, it is possible to modify Fletcher's formula for calculating the pitch of the voice fundamental in semitones above 1 Hz:

\[ P (st) = 12 \log_2 F_0 (Hz) \text{ (in cps)} \]

According to this formula, 1 Hz = 0 at, 2 Hz = 12 at, 4 Hz = 24 at, 6 Hz = 32 at, 12 Hz = 60 at = 100 at = 1 (Fig. 1). That is, instead of operating with figures above 1 Hz, one can operate in the pitch range of 64 Hz to 512 Hz. The figures of the latter scale are perceptually misleading; they give a wrong idea of an extensive pitch movement which is never perceived by listeners as such. The logarithmic scale is a solution for graphs, although not very convenient for plotting unless one has special charts where every cps (Hz) can be plotted accurately.

The comparison and statistical processing of raw data in Hz in terms of perception leads to distortions even in cases of one speaker, two speakers, and four speakers with different pitch ranges. The perceptually relevant comparison of two tones can be carried out by calculating their ratio, which further may be converted into semitones. Thus, given two measured frequencies, 150 Hz and 100 Hz, it is possible to calculate the difference (interval) between the lower pitch and the average as well as between the higher pitch and the average is 15 Hz, 1 octave in the (1/3 octave intervals). Therefore, the scale in Hz is absolutely informative, unlike the scale of the two former figures, 288 Hz, where the lower interval would be about 26 Hz as against 10 Hz of the upper interval. Data in Hz can easily be converted into absolute semitones by means of a table where every Hz is given its correlate value in st with the accuracy of .1 at (higher precision is unnecessary in phonetics). This table is printed on the 4th page of the present paper. In computer-assisted P. extraction the conversion can be done automatically, applying the above formula. For a programming language applying natural logarithms (such as BASIC used for computing the given conversion table) the formula will be

\[ P = 12 \times \log_2 (F_0) \]

It would be highly advisable to present even raw data in these absolute perceptual units. The investigator himself could immediately easily perceive differences between the measured parameters of pitch. Only then can all kinds of mathematical operations with the data without the ad hoc calculation of ratios or finding of logarithms. Intervals could be calculated by simple subtraction. Pitch contours and other graphs can easily be drawn on ordinary square paper.

The reader, too, could at once see what the measured pitches, intervals and ranges mean in terms of perception. Also, a reader of publications applying the absolute semitone scale could easily compare the data of different authors without the need to convert the Hz frequencies and then back into the traditional but unnecessary Hz if he wants to publish his own comparison of pitch contours and other graphs of the results and initial data for further generalizations, compute averages of pitch contours of different authors (including one's own), etc.

P. Lehiste and I. L. I. (1984) includes two tables representing the average pitch in 2- and 3-syllable tone groups. Pitch syllables are expressed in Hz. Let us consider a line of Table 1: male speaker 150-152, 143-146 Hz.

Fig. 1. (Left) Linear frequency scale in Hz (left) and the correlate pitch values in semitones (right). Table 1 shows the data in Hz. Let us consider a line of Table 1: 1st syll. 2nd syll.

male speaker 150-152, 143-146 Hz. The figures are given as averages. Although it is wrong in phonetics to average heart sounds, it is possible to apply logarithms, let us regard these figures as representing single speech acts. As one can see is that both speakers pronounce the first syllable with falling pitch and the second with a fall-rise which is steeper for the female speaker. Now let us convert the heights of semitones into Hz. Male speaker 95-90, 80-82 Hz. Female speaker 95-91, 80-82 Hz.

Fig. 2. (Right.) Linear pitch scale in at (left) and the correlate pitch values in semitones (right). Table 2 shows the data in Hz. Let us consider a line of Table 2: 1st syll. 2nd syll. Male speaker 95-90, 80-82 Hz. Female speaker 95-91, 80-82 Hz.

Here the extent of pitch movement is at once obvious: the male speaker appears to make a steeper pitch movement against the female speaker's 8 at; the 2nd syllable is much steeper in the female speaker than in the male speaker. Further comparison with the other tone groups in the table
may show to what extent these findings are relevant.

Another aspect. In order to average the two pitch contours of the above tone group, with their parameters expressed in Hz, we would have to draw both of them on logarithmic paper and calculate the average contour geometrically (Fig. 3). Yet it is much simpler to average the parameters expressed in Hz, as we have to do in this paper and to draw the resulting contour of the logarithmic paper and calculate the average contour geometrically.

The absolute semitone scale was first introduced in Tallinn in 1972 (VERDE 1972) and has since been successfully applied here (e.g., Piir 1995).

The other existing pitch scales, such as the model scale or the Ear’s scale, are efficient for plotting psychoacoustic data for frequencies above 500 Hz, i.e., for spectrum analysis, but apparently not sensitive enough and too clumsy to handle (otherwise why should prosodists have avoided them?) in the range of fundamental frequency. It remains to hope that the absolute semitone scale, which is both physical and perceptual, will gradually replace the hitherto dominating hertz tradition, take root and spread in prosodic research.

REFERENCES

Fletcher, H. 1929. Speech and Hearing. London.


Fig. 3. Plotting and averaging of two contours of a Chinese tone group on the logarithmic frequency scale and absolute semitone scale.

male speaker
--- female speaker
-.-- the average contour

When synthetic speech is used in prosodic research, it is expedient with a view to their subsequent mathematical/statistical analysis, to take up the tonal contours for the synthetic stimuli in absolute pitch units, varying the pitch of certain contour points by steps of m st instead of n Hz. It will considerably facilitate, for instance, correlation analysis between the input pitch data and the listeners' responses when the former are expressed in semitones on the absolute linear pitch scale; and it is equally easy to interpret the results of such analysis.

The absolute semitone scale has since been successfully applied in Tallinn in 1972 (VERDE 1972) and has since been successfully applied here (e.g., Piir 1995).

The other existing pitch scales, such as the model scale or the Ear's scale, are efficient for plotting psychoacoustic data for frequencies above 500 Hz, i.e., for spectrum analysis, but apparently not sensitive enough and too clumsy to handle (otherwise why should prosodists have avoided them?) in the range of fundamental frequency. It remains to hope that the absolute semitone scale, which is both physical and perceptual, will gradually replace the hitherto dominating hertz tradition, take root and spread in prosodic research.

REFERENCES

Fletcher, H. 1929. Speech and Hearing. London.

