THE ABSOLUTE SEMITONE SCALE

KULLO VENDE

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,

$$P\ (st) = 12 \log_2 F_o\ (Hz),$$

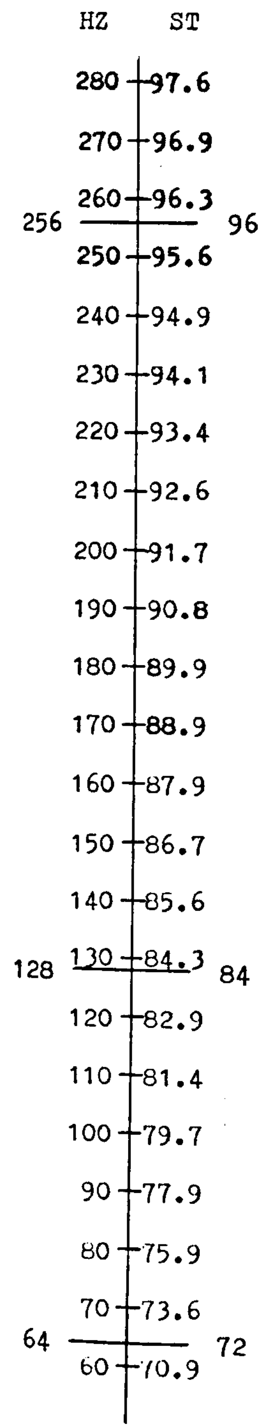
to relate fundamental frequency to its correlate perceptual units of pitch, viz., semitones above 1 Hz (1 Hz = 0 st). The absolute pitch units are much more convenient than Hz for the presentation, comparison (i.e., calculation of relative pitch differences) and other processing of raw data obtained in instrumental prosodic research.

In prosodic research the presentation of fundamental frequency in cycles per second is admissible only as far as raw measurement data are concerned. Any further manipulation and discussion or interpretation of the data should be carried out in units of perception. Even the graphs of F_o movement applying the linear frequency 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 case of one speaker, let alone speakers with different F_o 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 useless to state, in a discussion of their perception, that their difference is 50 Hz: the perception of the 50 Hz difference here is quite unlike the perception of a 50 Hz difference, say, between 250 Hz and 200 Hz. Instead, one can state that the r a t i o of the first pair of frequencies is 3:2 whereas that of the second pair is 5:4. For an untrained imagination, however, it is still clearer to state that the (musical) i n t e r v a l between the first two frequencies is 7 semitones, or a fifth, and that between the other two frequencies is 4 semitones, or a third.
But calculating average F_o values in Hz is of very doubtful value, as is drawing conclusions from differences between such values or such averages. To say that one F_o contour individually or on an average differs from another by a 10 Hz difference between their peaks is quite meaningless.

As long as we believe that the perception of the fundamental frequency of speech is logarithmic in the same way as it is for pure tones, the only possibility to process F_o data mathematically is in linear units on the logarithmic scale to which the data should be converted. The basic unit of pitch is the octave. The convenient unit for the analysis of fundamental pitch is the semitone. Proceeding from FLETCHER (1929) who introduced a scale of octaves and centioctaves above 1 kHz for the whole of the audible pitch range, it is possible to modify Fletcher's formula for calculating the pitch of the voice fundamental in a b s o l u t e s e m i t o n e s a b o v e 1 Hz:

$$P\ (st) = 12 \log_2 F_o\ (Hz \approx cps).$$

According to this formula,
1 Hz = 0 st, 2 Hz = 12 st, 4 Hz = 24 st, 64 Hz = 72 st, 512 Hz = 108 st (Fig. 1).
That is, instead of operating with figures in the F_o range of (roughly) 64..512 Hz, we can operate in the pitch range of 72....108 st. The figures of the latter scale are suited for any kind of mathematical processing without notably violating the perceptual reality. Considering these two pairs of numerical data (depicted in Fig. 2), we can easily find the average pitch of the latter pair to be 90 st; the difference (interval) between the lower pitch and the average as well as between the higher pitch and the average is 18 st (1.5 octaves). The result is perceptually informative, unlike the average of the two former figures, 288 Hz, where the lower interval would be about 26 st as against 10 st 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



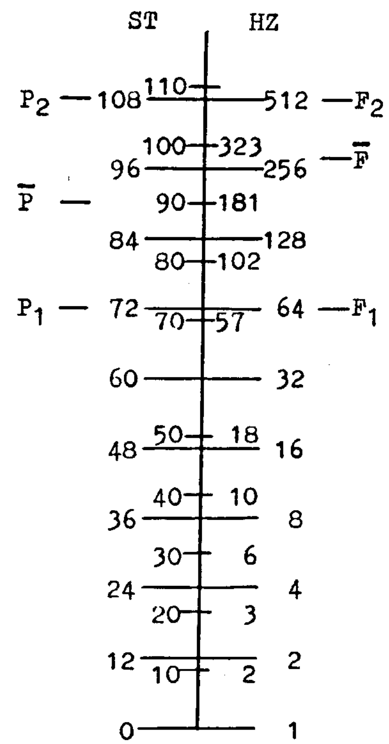


Fig. 1. (Left.) Linear frequency scale in Hertz (left) and the correlate pitch values in semitones (right).

Fig. 2. (Right.) Linear pitch scale in st (left) and the correlate frequency values in Hz (right). Plotting of two fundamental frequencies, F_1 = 64 Hz and F_2 = 512 Hz, on the linear pitch scale and the correlate logarithmic frequency scale. $\bar{F}$ is the mean frequency, $\bar{P}$ is the mean pitch of the two signals.

value in st with the accuracy of .1 st (higher precision is unnecessary in phonetics). This table is printed on the 4th page of the present paper. In computer-assisted F_o 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

$$P = 12 \times 1.442695 \times \ln F_o.$$

It would be highly advisable to present even raw data in these absolute perceptual units. The investigator himself could immediately estimate the perceivable differences between the measured parameters of pitch and carry out 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 into ratios and then back into the traditional but unnecessary Hz if he wants to publish his results of comparison. One could combine and process the results and initial data for further generalizations, compute averages of pitch contours of different authors (including one's own), etc.
For example, the paper of LIN et al. (1984) includes two tables representing the average pitch in 2- and 3-syllable tone groups of Chinese. Pitch is expressed in Hz. Let us consider a line of their Table 1 - tone 4 + tone 3:

	1st syll.	2nd syll.
male speaker	196-110	104-82-114
female speaker	242-152	143-82-156

The figures are given as averages. Although it is wrong in phonetics to average hertzes (what can be averaged is their logarithms), let us regard these figures as representing single speech acts. All we 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 hertzes into semitones:

	1st syll.	2nd syll.
male speaker	91.4-81.4	80.4-76.3-82.0
female speaker	95.0-87.0	85.9-76.3-87.4

Here the extent of pitch movement is at once obvious: the male speaker appears to make a 10-st fall in the 1st syllable against the female speaker's 8 st; the 2nd syllable starts 1 st below the end of the 1st; the subsequent falls in the 2nd syllable are 4 st and nearly 10 st, respectively, and the final rises about 6 and 11 st, respectively. 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 st arithmetically,
(m+f)/2   93.2-84.2   83.2-76.3-84.7,
and draw the resulting contour of the same shape on square paper.

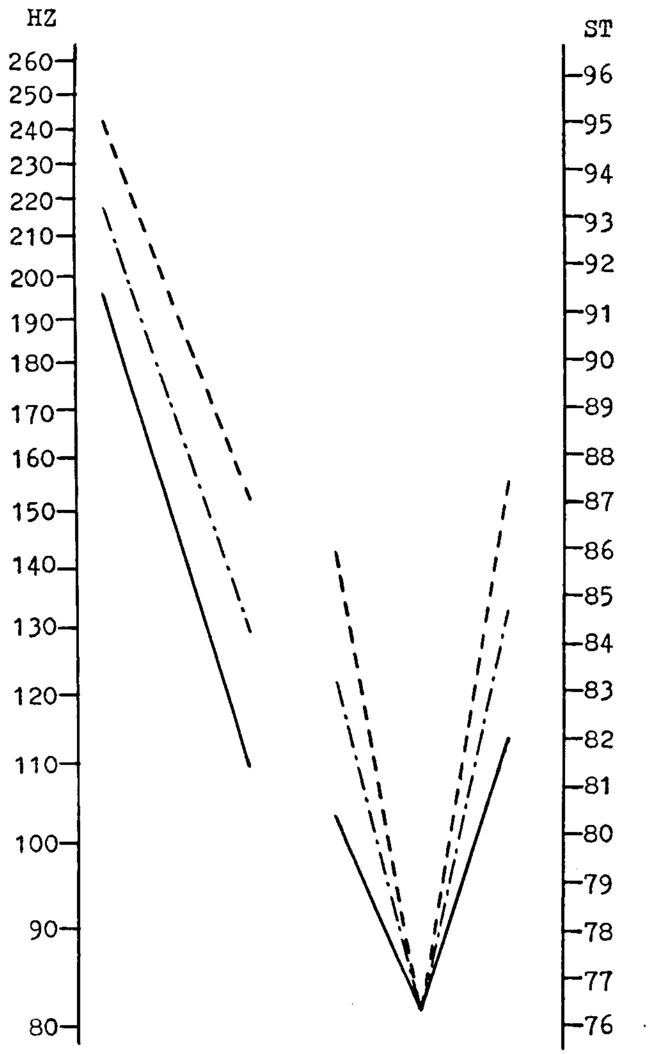


Fig. 3. Plotting and averaging of two contours of a Chinese tone group on the logarithmic frequency scale and absolute semitone scale.
——— male speaker
------ female speaker
-.-.-. the average contour

When synthetic speech is used in prosodic research, it is expedient, with a view to their subsequent mathematical/statistical analysis, to make 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 was first introduced in Tallinn in 1972 (VENDE 1972) and has since been successfully applied here (e.g., PIIR 1985).
The other existing pitch scales, such as the mel scale or the Bark 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 likewise both physical and perceptual, will gradually break through the hitherto dominating hertz tradition, take root and spread in prosodic research.

REFERENCES

Fletcher, H. 1929. Speech and Hearing. London.
Lin, M.C., Yan, T.Zh., Sun, G.H. 1984. The stress pattern and its acoustic correlates in Beijing Mandarin. - In: The Proceedings of the Tenth International Congress of Phonetic Sciences. - Dordrecht: Foris Publications, p. 504-514.
Piir, H. 1985. Acoustics of the Estonian diphthongs. - Estonian Papers in Phonetics: EPP 1982-1983. - Tallinn, p. 5-96.
Vende, K. 1972. Intrinsic pitch of Estonian vowels: measurement and perception. - Estonian Papers in Phonetics. - Tallinn, p. 44-108.

TABLE

Conversion of Hertzes into Semitones

Hz	st	Hz	st	Hz	st	Hz	st
1	0.0	60	70.9	119	82.7	178	89.7
2	12.0	61	71.2	120	82.9	179	89.8
3	19.0	62	71.5	121	83.0	180	89.9
4	24.0	63	71.7	122	83.2	181	90.0
5	27.9	64	72.0	123	83.3	182	90.1
6	31.0	65	72.3	124	83.5	183	90.2
7	33.7	66	72.5	125	83.6	184	90.3
8	36.0	67	72.8	126	83.7	185	90.4
9	38.0	68	73.0	127	83.9	186	90.5
10	39.9	69	73.3	128	84.0	187	90.6
11	41.5	70	73.6	129	84.1	188	90.7
12	43.0	71	73.8	130	84.3	189	90.7
13	44.4	72	74.0	131	84.4	190	90.8
14	45.7	73	74.3	132	84.5	191	90.9
15	46.9	74	74.5	133	84.7	192	91.0
16	48.0	75	74.7	134	84.8	193	91.1
17	49.0	76	75.0	135	84.9	194	91.2
18	50.0	77	75.2	136	85.0	195	91.3
19	51.0	78	75.4	137	85.2	196	91.4
20	51.9	79	75.6	138	85.3	197	91.5
21	52.7	80	75.9	139	85.4	198	91.6
22	53.5	81	76.1	140	85.6	199	91.6
23	54.3	82	76.3	141	85.7	200	91.7
24	55.0	83	76.5	142	85.8	201	91.8
25	55.7	84	76.7	143	85.9	202	91.9
26	56.4	85	76.9	144	86.0	203	92.0
27	57.1	86	77.1	145	86.2	204	92.1
28	57.7	87	77.3	146	86.3	205	92.2
29	58.3	88	77.5	147	86.4	206	92.2
30	58.9	89	77.7	148	86.5	207	92.3
31	59.5	90	77.9	149	86.6	208	92.4
32	60.0	91	78.1	150	86.7	209	92.5
33	60.5	92	78.3	151	86.9	210	92.6
34	61.0	93	78.5	152	87.0	211	92.7
35	61.6	94	78.7	153	87.1	212	92.7
36	62.0	95	78.8	154	87.2	213	92.8
37	62.5	96	79.0	155	87.3	214	92.9
38	63.0	97	79.2	156	87.4	215	93.0
39	63.4	98	79.4	157	87.5	216	93.1
40	63.9	99	79.6	158	87.6	217	93.1
41	64.3	100	79.7	159	87.8	218	93.2
42	64.7	101	79.9	160	87.9	219	93.3
43	65.1	102	80.1	161	88.0	220	93.4
44	65.5	103	80.2	162	88.1	221	93.5
45	65.9	104	80.4	163	88.2	222	93.5
46	66.3	105	80.6	164	88.3	223	93.6
47	66.7	106	80.7	165	88.4	224	93.7
48	67.0	107	80.9	166	88.5	225	93.8
49	67.4	108	81.1	167	88.6	226	93.8
50	67.7	109	81.2	168	88.7	227	93.9
51	68.1	110	81.4	169	88.8	228	94.0
52	68.4	111	81.5	170	88.9	229	94.1
53	68.7	112	81.7	171	89.0	230	94.1
54	69.1	113	81.8	172	89.1	231	94.2
55	69.4	114	82.0	173	89.2	232	94.3
56	69.7	115	82.1	174	89.3	233	94.4
57	70.0	116	82.3	175	89.4	234	94.4
58	70.3	117	82.4	176	89.5	235	94.5
59	70.6	118	82.6	177	89.6	236	94.6

237	94.7	306	99.1	375	102.6	444	105.5
238	94.7	307	99.1	376	102.7	445	105.6
239	94.8	308	99.2	377	102.7	446	105.6
240	94.9	309	99.3	378	102.7	447	105.6
241	95.0	310	99.3	379	102.8	448	105.7
242	95.0	311	99.4	380	102.8	449	105.7
243	95.1	312	99.4	381	102.9	450	105.8
244	95.2	313	99.5	382	102.9	451	105.8
245	95.2	314	99.5	383	103.0	452	105.8
246	95.3	315	99.6	384	103.0	453	105.9
247	95.4	316	99.6	385	103.1	454	105.9
248	95.5	317	99.7	386	103.1	455	106.0
249	95.5	318	99.8	387	103.2	456	106.0
250	95.6	319	99.8	388	103.2	457	106.0
251	95.7	320	99.9	389	103.2	458	106.1
252	95.7	321	99.9	390	103.3	459	106.1
253	95.8	322	100.0	391	103.3	460	106.1
254	95.9	323	100.0	392	103.4	461	106.2
255	95.9	324	100.1	393	103.4	462	106.2
256	96.0	325	100.1	394	103.5	463	106.3
257	96.1	326	100.2	395	103.5	464	106.3
258	96.1	327	100.2	396	103.6	465	106.3
259	96.2	328	100.3	397	103.6	466	106.4
260	96.3	329	100.3	398	103.6	467	106.4
261	96.3	330	100.4	399	103.7	468	106.4
262	96.4	331	100.4	400	103.7	469	106.5
263	96.5	332	100.5	401	103.8	470	106.5
264	96.5	333	100.6	402	103.8	471	106.6
265	96.6	334	100.6	403	103.9	472	106.6
266	96.7	335	100.7	404	103.9	473	106.6
267	96.7	336	100.7	405	103.9	474	106.7
268	96.8	337	100.8	406	104.0	475	106.7
269	96.9	338	100.8	407	104.0	476	106.7
270	96.9	339	100.9	408	104.1	477	106.8
271	97.0	340	100.9	409	104.1	478	106.8
272	97.0	341	101.0	410	104.2	479	106.8
273	97.1	342	101.0	411	104.2	480	106.9
274	97.2	343	101.1	412	104.2	481	106.9
275	97.2	344	101.1	413	104.3	482	107.0
276	97.3	345	101.2	414	104.3	483	107.0
277	97.4	346	101.2	415	104.4	484	107.0
278	97.4	347	101.3	416	104.4	485	107.1
279	97.5	348	101.3	417	104.4	486	107.1
280	97.6	349	101.4	418	104.5	487	107.1
281	97.6	350	101.4	419	104.5	488	107.2
282	97.7	351	101.5	420	104.6	489	107.2
283	97.7	352	101.5	421	104.6	490	107.2
284	97.8	353	101.6	422	104.7	491	107.3
285	97.9	354	101.6	423	104.7	492	107.3
286	97.9	355	101.7	424	104.7	493	107.3
287	98.0	356	101.7	425	104.8	494	107.4
288	98.0	357	101.8	426	104.8	495	107.4
289	98.1	358	101.8	427	104.9	496	107.5
290	98.2	359	101.9	428	104.9	497	107.5
291	98.2	360	101.9	429	104.9	498	107.5
292	98.3	361	102.0	430	105.0	499	107.6
293	98.3	362	102.0	431	105.0	500	107.6
294	98.4	363	102.0	432	105.1	501	107.6
295	98.5	364	102.1	433	105.1	502	107.7
296	98.5	365	102.1	434	105.1	503	107.7
297	98.6	366	102.2	435	105.2	504	107.7
298	98.6	367	102.2	436	105.2	505	107.8
299	98.7	368	102.3	437	105.3	506	107.8
300	98.7	369	102.3	438	105.3	507	107.8
301	98.8	370	102.4	439	105.3	508	107.9
302	98.9	371	102.4	440	105.4	509	107.9
303	98.9	372	102.5	441	105.4	510	107.9
304	99.0	373	102.5	442	105.5	511	108.0
305	99.0	374	102.6	443	105.5	512	108.0