EESTI NSV TEADUSTE AKADEEMIA TOIMETISED. 26. KÖIDE
ÜHISKONNATEADUSED. 1977, NR. 3

ИЗВЕСТИЯ АКАДЕМИИ НАУК ЭСТОНСКОЙ ССР. ТОМ 26
ОБЩЕСТВЕННЫЕ НАУКИ. 1977, № 3

https://doi.org/10.3176/hum.soc.sci.1977.3.06

U. LIPPUS, M. REMMEL, J. ROSS

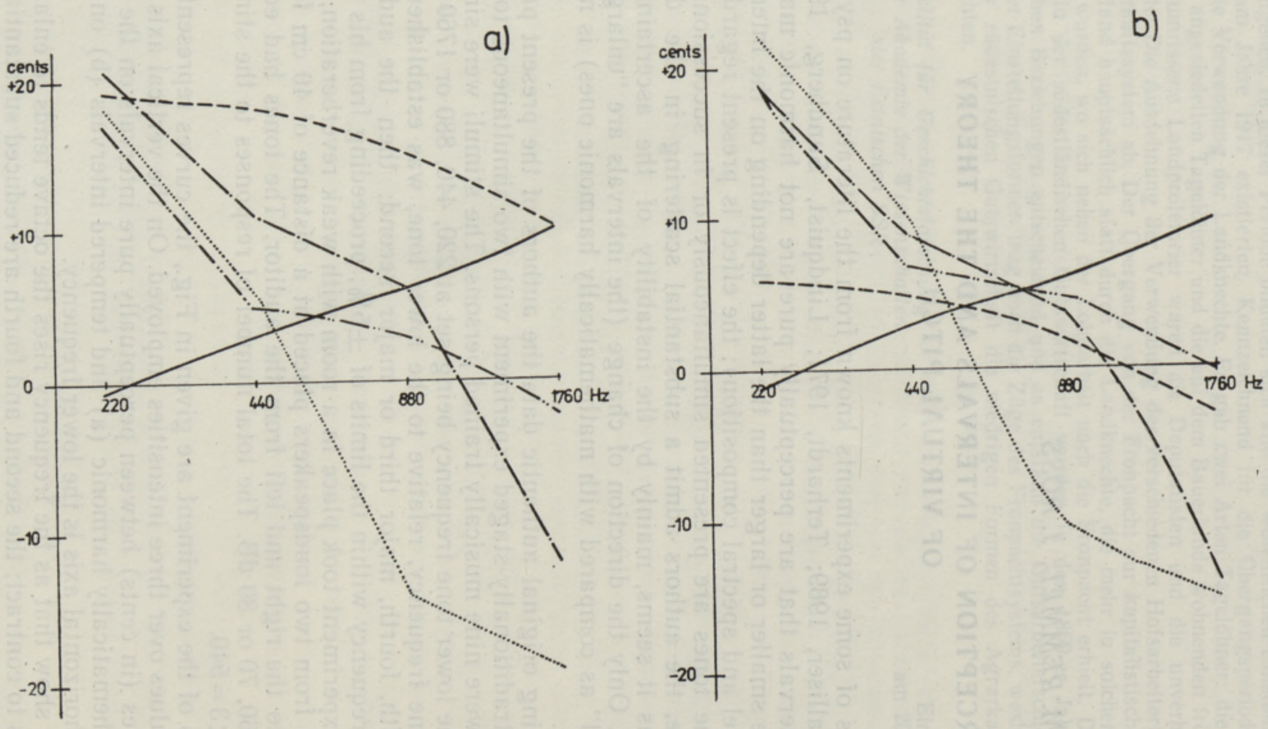
PERCEPTION OF INTERVALS AND THE THEORY OF VIRTUAL PITCH

The results of some experiments known from the literature on psychoacoustics (Walliser, 1969; Terhardt, 1971; Lindquist, Sundberg, 1971) show that intervals that are perceptually pure are not harmonic mathematically (are smaller or larger than the latter depending on the interval, frequency, level and spectral composition); the effect is present regardless of whether the tones are presented simultaneously or in succession. At the same time, the authors admit a substantial scattering in the data, conditioned, as it seems, mainly by the instability of the ascertainable local extrema. Only the direction of change (the intervals are "enlarged" or "contracted" as compared with mathematically harmonic ones) is more or less stable.

For obtaining original authentic data the authors of the present paper carried out a traditionally-staged experiment with two simultaneous tones. The auditors were nine musically trained persons. The stimuli were sinusoidal tones, the lower tone frequency being set at 220, 440, 880 or 1760 Hz. The higher tone frequency, relative to the lower tone, was established at the octave, fifth, fourth, major third or major second; then the auditor adjusted the frequency within the limits of ±5%, proceeding from his perception. The experiment took place in a room with weak reverberation, the stimuli issued from two loudspeakers placed at a distance of 40 cm from one another, to the right and left from the auditor. The tones had equal intensities of 60, 70 or 80 dB. The total number of responses to the stimuli was $9\times4\times5\times3=540$.

The results of the experiment are given in Fig., the curves representing the average values over three intensities employed. On the vertical axis are the divergencies (in cents) between perceptually pure intervals on the one hand and mathematically harmonic (a) and tempered intervals (b) on the other. On the horizontal axis is the lower frequency.

These data show that as the frequency rises the octave tends to enlarge, other intervals to contract; the second and fourth are reduced substantially. By means of rough interpolation, the zero-crossings of the curves are given the values presented in the left-hand column of Table 1 (the case of tempered tuning); the frequencies in the right-hand column are obtained by a similar extrapolation — at those frequencies the presented tones occur as higher harmonics in case the fundamental frequency is represented by the value of the zero-crossing.



Divergences (in cents) between perceptually pure and mathematically harmonic (a), and perceptually pure and tempered (b) intervals, on the horizontal axis lower tone frequency. Octave —, fifth —.., fourth —. , major third — — — , major second .....

Table 1

octave	260	520
fifth	1760	5280
fourth	1100	(5500)
major third	1100	5500
major second	660	5280

The system of values derived in this way has an interesting property, viz. the zero-crossings of intervals define two regions: 0.5 . . kHz and 5. . . kHz. Those regions, as is well known, have a special role in speech acoustics: they designate the upper limits of the fundamental frequency and of the relevant part of the vowel spectrum.

The established regularity is not entirely unexpected, especially in view of the model developed by Terhardt relating the basic characteristics of speech acoustics to the intervallic content of music (Terhardt, 1974). The model is based on the concept of virtual pitch and its gist is as follows.

In order that a human being could acquire the ability to reconstruct a signal with correct phase relations, as well as other abilities to respond to the acoustic "Gestalt", the sufficient precondition in the stage of learning is to have at one's disposal a starting mechanism synchronizing the synphasic complex tone with the lowest partial. In man (child) the role of the corpus during the learning process is played by natural speech — correct phase relations promote the prediction of (also imaginary) harmonic partials, which together with the above-mentioned shifts in the reception of pitch and a special mechanism for summing up the "pitch impulses" and "synchroimpulses" produces the pattern of virtual pitch. The learning matrix, with the pattern of virtual pitch being developed at one of its outputs, is located, due to the above-mentioned connections with speech acoustics, lower than 0.5 kHz in the dimension of synchronization and under 5. . . kHz in the dimension of spectral components.

So far one question is left open in Terhardt's model, namely the mechanism of the pitch shift (i. e. the problem of on what is it still that the enlargement/contraction of intervals is based). Before proceeding to the hypothesis that also the pitch shift is a product of learning, let us consider two factors which, as is known from the literature, are related to the issue in question. These are the sound pressure level and the presence/absence of adjacent partials. It was the first that we studied directly on the basis of our material, but without finding any explicit correlation between the sound pressure level and the extent of the shift (as a matter of fact, the difference of levels was relatively small, with limit values of 60 and 80 dB). No separate experiment was staged as to the presence/absence of adjacent partials; however, the authors notated by means of a digital oscilloscope some Estonian runic tunes based on the non-tempered ("pure") scale (four authentic performances, two performers). The averaged values (in cents) of enlargement/contraction of intervals are given in Table 2 (a total of 487 intervals was measured).

The Table shows that nearly all the intervals have become enlarged, besides, in the given frequency area the enlargement change is convex, thus coinciding in the case of the fifth with an earlier result, see Walliser, 1969. Such enlargement of intervals is by no means limited to Estonian runic tunes, in some musical traditions it occurs repeatedly during a lengthy song. A suitable example here is Yakut folk music (see Алексеев, 1976) where during a song the intervals between supporting tones may grow by several semitones, the principal system of supporting tones being retained; at the same time, the rise in the pitch between adjacent verses

Table 2

Interval	Approximate frequency of the lower tone, Hz				
	160	185	205	235	260
fifth	27	29	36	26	41
fourth	38	—	—	42	37
major third	—	24	44	40	19
minor third	42	27	36	—	12
major second	−14	—	27	19	42
minor second	—	19	15	16	—16

in Alexeyev's examples is of the same order as the values given in Table 2 or Fig. In the given aspect, the difference between Estonian and Yakut song consists in the fact that in a Yakut song, each following verse refers to the preceding one and in an Estonian song to a stable initial pattern.

Also the second regularity pointed out by Alexeyev ("as the pitch rises the distances between adjacent stages have a tendency to contract proportionally" — op. cit., p. 91) agrees with the results described in the present paper, the ratio of the contraction presented in his examples being likewise of the same order as presumed by the convexity in Table 2.

Returning now to the generating mechanism of the enlarged intervals let us take closer look at a scheme for speech analysis as described recently by M. Rohtla. At the output of the scheme a pattern is produced different from classical conceptions — its formant (peak) frequencies do not keep to fixed values during a pitch period, instead, they fluctuate around a conditional central value (Rohtla, 1976; Remmel, Rohtla, 1977). The scheme itself consists of four wide band filters (the bands in Rohtla's variant are 80—320 Hz, 320—720 Hz, 720—2160 Hz and 2160—6480 Hz) and four period meters connected to the filter outputs. The period meters are interconnected in such a manner that they can measure only those time intervals as are smaller than the ones registered by the period meter located behind the lower frequency filter. In order to define more precisely the statistical properties of peak fluctuations, Rohtla's scheme was first simulated on the computer (the result was a similar pattern) and then subjected to usual digital filtration at intervals smaller than the pitch periods, the beginning of the pitch period being established at the zero-crossing before the energy started to rise. Isolated vowels from one male and one female speaker were fed into the computer directly from the microphone by means of a 11-bit converter with a sampling frequency of 30 kHz. 14-order elliptic filters were used. The filtration was carried out in such a manner that the locations of energy peaks were slidingly searched for in the vicinity of harmonic partials. Then the period mean values, square deviations and asymmetries were calculated for established fluctuations. The results for one typical period of the vowels (*a, i, u*) are given in Table 3.

The Table indicates the dominance of negative asymmetry which in its turn means that for most of the time the peak frequency is located higher than would be the case with the corresponding multiple of fundamental frequency.

It is possible that the described phenomenon may (through "Gestaltic" transformations) be the natural source of stretched musical intervals, and, as one can see, it can in all its details be easily connected with Terhardt's model. At the same time, the causes of asymmetry remain as yet an open question.

Table 3

Mean values of peak frequencies, standard deviations and asymmetries in the vicinity of harmonic partials

The number of harmonic partial, i	Sex	[a]			[i]			[u]		
		f_i/f_1	$\sigma(f_i)$	$S_k(f_i)$	f_i/f_1	$\sigma(f_i)$	$S_k(f_i)$	f_i/f_1	$\sigma(f_i)$	$S_k(f_i)$
2	♂	2.01	7.1	−1.2	1.99	5.5	−1.3	1.97	10.6	−1.1
	♀	2.02	10.2	−1.1	2.01	7.1	−1.7	2.00	9.9	− .7
3	♂	3.04	8.7	−1.2	3.02	7.1	−2.4	2.99	8.5	− .9
	♀	3.04	13.1	−1.8	3.04	4.2	−2.3	3.01	12.6	− .2
4	♂	4.04	9.0	−2.7	4.05	8.3	−2.0	4.02	11.0	−1.4
	♀	4.03	13.4	−2.4	4.05	4.7	−1.9	4.02	10.9	− .9
5	♂	5.02	9.4	−2.9	5.04	6.5	−4.7	5.04	11.4	−2.2
	♀	5.02	14.9	−2.8	5.04	16.0	−3.6	5.02	14.7	−1.6
6	♂	6.01	10.5	−3.4	6.01	11.4	−3.9	6.04	15.6	−3.5
	♀	6.00	16.2	−3.6	5.98	15.3	−2.8	6.03	21.7	−2.8
7	♂	7.02	12.2	−3.7	7.00	14.2	−5.2	7.00	12.7	−4.7
	♀	6.98	18.1	−4.5	6.97	18.4	−4.4	7.02	22.8	−3.4
8	♂	8.00	14.4	−4.2	7.99	19.6	−4.4	8.01	17.3	−4.3
	♀	7.99	20.5	−5.0	8.02	17.6	−4.2	8.06	21.4	−3.0

REFERENCES

Lindquist, J., Sundberg, J., 1971. Perception of the Octave Interval. — Proceedings of the 7th International Congress on Acoustics, Vol. 3, pp. 617—620.

Remmel, M., Rohtla, M., 1977. Some Topics of Speech Recognition. Preprint KKI-7. Academy of Sciences of the Estonian SSR. Department of Social Sciences. Tallinn.

Rohtla, M., 1976. On the Invariance of F-pattern. — Estonian Papers in Phonetics 1976. Tallinn, pp. 75—78.

Terhardt, E., 1971. Die Tonhöhe harmonischer Klänge und das Oktavintervall. — Acustica, 24, 3, pp. 126—136.

Terhardt, E., 1974. Pitch, Consonance, and Harmony. — JASA, 55, 5, pp. 1061—1069.

Walliser, K., 1969. Über die Spreizung von empfundenen Intervallen gegenüber mathematisch harmonischen Intervallen bei Sinustönen. — Frequenz, 23, 5, pp. 139—143.

Алексеев Э., 1976. Проблемы формирования лада (на материале якутской народной песни). Москва.

Academy of Sciences of the Estonian SSR,
Institute of Language and Literature

Received
Jan. 22, 1977

U. LIPPUS, M. REMMEL, J. ROSS

INTERVALLIDE TAJU JA VIRTUAALSE HELIKÕRGUSE TEOORIA

Resümee

Käsitletakse põhjusi, miks inimene tajub laienenud muusikalisi intervalle puhastena. Katsetulemuste ning Terhardti virtuaalse helikõrguse mudeli kõrvutamisel on tuldud järeldusele, et selline taju võib inimesel (lapsel) tekkida kõnemehhanismide õppimise kõrvaltulemusena.

Eesti NSV Teaduste Akadeemia
Keele ja Kirjanduse Instituut

Toimetusse saabunud
22. I 1977

У. ЛИППУС, М. РЕММЕЛЬ, Я. РОСС

ВОСПРИЯТИЕ ИНТЕРВАЛОВ И ТЕОРИЯ ВИРТУАЛЬНОЙ ВЫСОТЫ

Резюме

Исследуются причины восприятия людьми расширенных музыкальных интервалов как «чистых». На основе сопоставления полученных результатов с моделью виртуальной высоты Терхардта выдвинута гипотеза, что такое восприятие является побочным результатом в процессе обучения человека (ребенка) механизмам речи.

Институт языка и литературы
Академии наук Эстонской ССР

Поступила в редакцию
22/I 1977