Psycho-Acoustic Experiments on the Sensory Consonance of Musical Two-Tones

Reinhart Frosch
Sommerhaldenstrasse 5B, CH-5200 Brugg, Switzerland
PSI (Paul Scherrer Institute), Villigen and ETH (Eidgenoessische Technische Hochschule), Zurich (retired)

1. INTRODUCTION

The topic of the present study is the sensory consonance of two-tones (i.e., pairs of simultaneous tones) composed of two harmonic complex tones similar to vibrato-less bowed-string tones. Per definition, a complex tone contains two or more simultaneous "simple" (i.e., sinusoidal) partial tones; in a *harmonic* complex tone, all partial-tone frequencies are integral multiples of the fundamental frequency f, i.e., of the frequency of the first partial.

According, e.g., to Plomp and Levelt (1965), *sensory consonance* is approximately equivalent to pleasantness, euphony, and beauty.

In previous sensory-consonance experiments with bowed-string-like tones, it was found that the *consonance curve* (i.e., the sensory consonance of the studied two-tone plotted versus its fundamental-frequency ratio $R = f_2 / f_1$) forms narrow peaks at small-integer ratios R such as 3/2, 5/3, etc.

The abscissa of the experimental consonance curve shown in Fig. 1 [taken from Section 3.2 of Frosch (2002)] is the *interval size* x of the two-tone:

$$x = (1200 \text{ cents}) \cdot \frac{\log(R)}{\log(2)}; \qquad (1)$$

e.g., an *octave* ($R = 2$) has an interval size $x = 1200$ cents, and a *perfect fifth* ($R = 3/2$) has $x \approx 702$ cents.

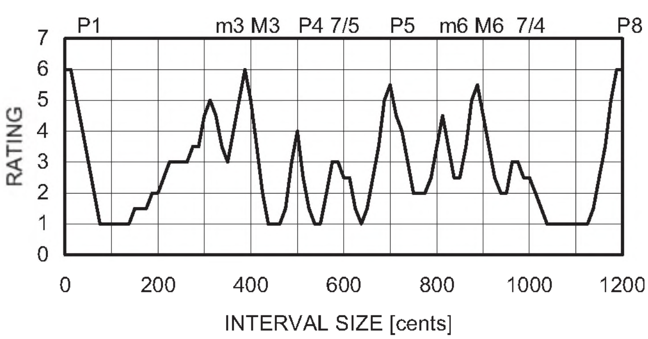

Fig. 1. Results of my "one-man" experiment on the sensory consonance of bowed-string-like two-tones; $f_1 = 264$ Hz ; rating 6 means "very consonant", and rating 1 means "very dissonant".

The symbols P1, m3, ... above Fig. 1 are the standard names of the diatonic intervals corresponding to the consonance peaks. Seven of these symbols are explained in Table 1. The names "perfect prime", "minor third", etc. can be understood by counting notes on a diatonic scale; e.g., the minor third *E*4-*G*4 involves three notes (namely *E*4, *F*4, *G*4); one of the two corresponding steps, *E*4-*F*4, is a diatonic semitone, and the other step *F*4-*G*4, is a diatonic whole tone; see Section 3.6 of Frosch (2002).

Interval	Symbol	*R*	*x* [cents]
Perfect prime	P1	1 / 1	0
Minor third	m3	6 / 5	316
Major third	M3	5 / 4	386
Perfect fourth	P4	4 / 3	498
Perfect fifth	P5	3 / 2	702
Major sixth	M6	5 / 3	884
Perfect octave	P8	2 / 1	1200

Table 1. Interval name, symbol, frequency ratio *R*, and interval size *x* of seven two-tones yielding consonance peaks in Fig. 1.

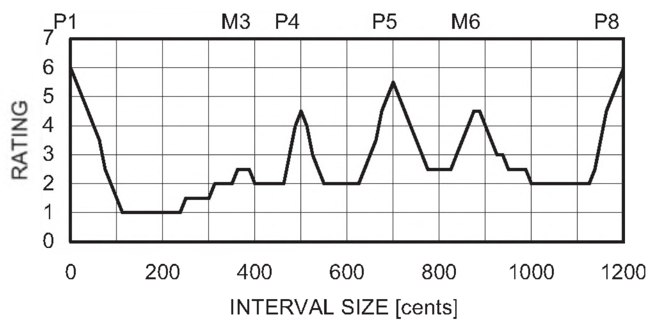

Fig. 2. Same as Fig.1; $f_1 = 99$ Hz .

In the case of Fig. 1, the deeper of the two simultaneous bowed-string-like tones had a fundamental frequency f_1 of 264 Hz; i.e., its note was *C*4. The fundamental frequency f_2 of the higher of the two simultaneous tones was chosen to be higher than f_1 by 0, 12.5, 25, 37.5, ... , 1200 cents; i.e., as many as 97 different two-tones were judged. The two-tones were generated at a sound-pressure level of about 70 dB on a Yamaha DX11 music synthesizer. The "voice" (i.e., the timbre) of the synthesizer was chosen to be "harmonica". The corresponding partial-tone spectrum differs little from that of vibrato-less bowed-string tones.

The experimental details of Fig. 2 were equal to those of Fig.1, with the exception of the fundamental frequencies. For Fig. 2, the frequency f_1 was 99 Hz; i.e., the note of the deeper tone was *G*2. Thus the two-tones of Fig. 1 were in the female-singing pitch region, whereas those of Fig. 2 were in the bass region. In comparison with Fig. 1, Fig. 2 contains fewer consonance peaks, its peaks are wider, and

:he peak heights differ; e.g., the major-third peak (M3) of Fig. 2 is lower than that of Fig. 1.

Tests similar to that just described have been done by other authors; e.g., the experiments of Hall and Hess (1984) and of Vos (1986) both yielded solid evidence for narrow sensory-consonance peaks at small-integer fundamental-frequency ratios. These experiments, however, did not yield complete sets of consonance-peak heights in the interval range 0–1200 cents.

The just mentioned interval range was covered fairly completely in the old experiment of Kaestner (1909); see Figs. 3 and 4. The method of complete paired comparisons used by Kaestner will be briefly described in Section 2 below.

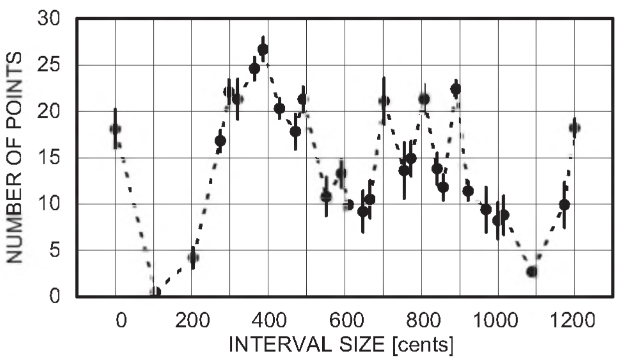

Fig. 3. Numbers of points obtained by 30 different two-tones in the range 0–1200 cents in the experiment of Kaestner (1909); deeper- tone frequency $f_1 = 256\ \mathrm{Hz}$.

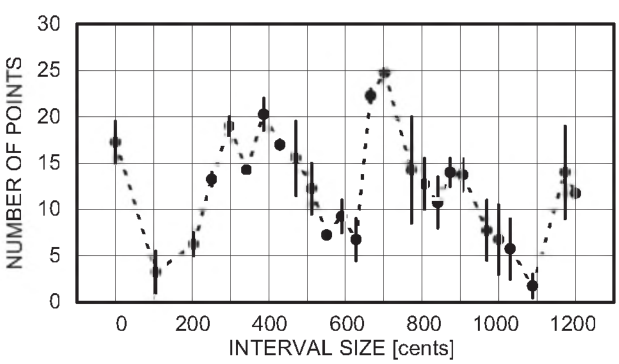

Fig.4. Same as Fig. 3; $f_1 = 128\ \mathrm{Hz}$.

The points with error bars in Fig. 3 represent the means of the point numbers obtained by 5 observers and the standard errors of those means. In the case of Fig. 4 there were 26 different two-tones and 2 observers.

My one-man experiment (Figs. 1 and 2) and the experiment of Kaestner (Figs. 3 and 4) both yielded consonance peaks at small-integer frequency ratios. The ranking order of the peak-heights, however, disagrees in several cases; the octave (1200 cents), e.g., in Figs. 1 and 2 obtains the first place (jointly with other two-tones); in Figs. 3 and 4, however, the octave is relegated to 7th place. In the deep-tone case (Figs. 2 and 4), disagreeing peak heights may be due to pitch differences (deeper-tone notes *G*2 and *C*3, respectively); in the higher-tone case (Figs. 1 and 3), however, the deeper-tone frequencies differ little (note *C*4 in both cases). In order to determine more reliably the relative consonance-peak heights, I conducted two "peak-height experiments" featuring complete paired comparisons similar to those of Kaestner (1909). These newer experiments, first reported in Frosch (2003) and Frosch (2005), will be briefly described in the following two sections.

2. PEAK-HEIGHT EXPERIMENT AT FEMALE-SINGING PITCH

In this experiment, the seven two-tones specified in Table 1 were judged by 18 subjects (no professional musicians). Table 1 contains all fundamental-frequency ratios *R* that have values ranging from 1.0 to 2.0 and are equal to ratios of integers m and n ranging from 1 to 6. The consonance curve in Fig. 1 has peaks at the interval sizes *x* corresponding to these seven *R*-values. Thus the purpose of the experiment was to determine the relative heights of the seven consonance peaks. The frequency of the deeper of the two simultaneous bowed-string-like tones was equal to that for Fig. 1 (i.e., 264 Hz, note *C*4). The properties of the tones were as described in the text below Fig. 2.

The 18 subjects were asked to listen to an audio-tape recording lasting about 15 minutes. The seven stimuli S_i (S_1 = prime, S_2 = minor third, etc.) were made to take part in a championship consisting of 7·6 = 42 matches. In contrast to Kaestner (1909), return matches were included, because they allowed the reliability of the subjects to be verified (see below). The order of the 42 matches was chosen at random; e.g., game 1 came out to be S_1 versus S_6, game 2 was S_3 versus S_1, etc.; the return match of game 1, i.e., S_6 versus S_1, turned out to be game 22.

Sequence of events on audio tape:
1) spoken words "game one" (duration: 2 seconds);
2) break (2 seconds);
3) stimulus S_1 is played (1 second);
4) break (0.5 second);
5) stimulus S_6 is played (1 second);
6)–9) repeat 2)–4) (4.5 seconds);
10)–13) repeat 2)–4) (4.5 seconds);
14) break (4 seconds; subjects write down their judgment);
15) spoken words "game two" (2 seconds); etc.

Possible judgments:
same as those in experiment of Kaestner (1909);
"*A*": the first of the two stimuli is more consonant;
"*B*": the second of the two stimuli is more consonant;
"*X*": the two stimuli are equally consonant.

Data taking:
Before judging the above-mentioned audio-tape recording, each of the 18 subjects was given a one-page form on which they were to write down their judgments on the 42 games, and they were instructed to attempt to decide which of the

:wo stimuli taking part in a game was more pleasant, eu- ɔhonious, and beautiful.

Total number of judgments:
Among the 18·42 = 756 judgments, there were 274 judg- nents *A*, 302 judgments *B*, and 180 judgments *X*. In a Monte-Carlo calculation based on the assumption that the ɔrobabilities of *A* and *B* were both equal to the arithmetic nean of 274/756 and 302/756, i.e., equal to 0.381, it was found that the judgment-number ratio *A*/*B* was smaller than ts experimental value of 274/302 in about 12200 of 100000 Monte-Carlo experiments; i.e., the (unexpected) difference ɔetween the just mentioned experimental judgment numbers 274 and 302 was found to be statistically not very signifi- cant.

Reproducibility of judgments:
The 42 judgments per subject form 21 pairs; each pair con- sists of a home match (e.g., stimulus S_6 versus stimulus S_2) and its return match (S_2 versus S_6). The judgment ɔairs *AB*, *BA*, and *XX* are consistent, whereas the remain- ng six combinations (*AA*, *AX*, *BB*, *BX*, *XA*, *XB*) signify that he subject has not judged consistently. A second Monte- Carlo program based on the judgment probabilities 274/756, 302/756, and 180/756 (see above) and on the assumption of random judgments yielded the bad-pair number distribution given by black discs in Fig. 5. The experimental histogram

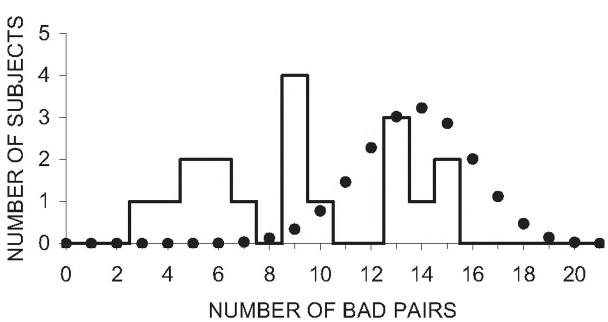

Fig. 5. Solid line: histogram of the numbers of subjects who totalled the number of bad judgment pairs specified on the abscissa. Black filled circles: numbers of subjects expected if the judgments are assumed to have been made at random (see text).

n Fig. 5 indicates that indeed 6 of the 18 subjects must be suspected to have judged randomly. The remainder of the analysis was restricted to the 12 subjects which had ≤ 10 ɔad pairs.

The twelve "reliable" subjects made a total of 12·42 = 504 udgments. If a judgment was either *A* or *B*, the winning stimulus received one point. In case of a draw (judgment *X*) each of the two participating stimuli was given 0.5 point. Each of the 7 stimuli took part in 12 games. The summed numbers of points are listed in Table 2. It is seen that sub- ect 1 attributed a total of 5.0 points to the perfect prime (P1, he single complex tone *C*4), whereas subject 8 gave the naximum of 12.0 points to that stimulus, and subject 10 attributed the minimum of 0.0 point. In the last column of Table 2, the sum of the 7 numbers of points from the considered subject is listed, and is seen to be equal to the number of judgments, i.e., to 42.0, as expected.

Results:

Subject	P1	m3	M3	P4	P5	M6	P8	Sum
1	5.0	9.0	11.0	2.0	5.0	6.5	3.5	42.0
2	6.0	9.5	12.0	0.0	2.5	7.0	5.0	42.0
3	10.5	7.0	6.5	7.0	6.0	5.0	0.0	42.0
4	4.0	9.0	12.0	6.0	6.0	5.0	0.0	42.0
5	5.5	10.0	9.0	6.5	2.5	7.5	1.0	42.0
6	7.5	8.0	8.5	5.5	5.5	5.5	1.5	42.0
7	1.0	11.0	7.0	6.0	6.0	8.0	3.0	42.0
8	12.0	3.5	7.0	6.5	6.5	3.5	3.0	42.0
9	6.5	9.5	10.0	4.0	2.5	6.0	3.5	42.0
10	0.0	5.0	8.5	6.0	7.5	8.0	7.0	42.0
11	6.0	10.5	10.5	7.5	4.0	3.5	0.0	42.0
12	5.5	3.0	10.0	8.0	7.5	7.0	1.0	42.0
Mean	5.79	7.92	9.33	5.42	5.13	6.04	2.38	42.0
Standard error	0.97	0.78	0.55	0.67	0.53	0.45	0.63	
x [cents]	0	316	386	498	702	884	1200	

Table 2. Results of the peak-height experiment at female-singing pitch.

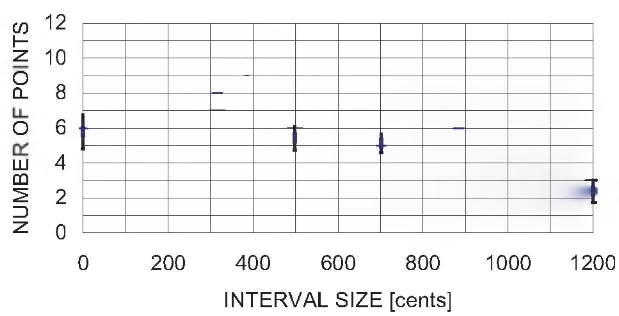

Fig. 6. Mean numbers of points obtained by the seven two-tones in Table 2; deeper-tone frequency $f_1 = 264$ Hz.

In the 14th line of Table 2, and also in Fig. 6, the arithmetic mean of the 12 numbers of points for the considered stimulus is presented. The 15th line of Table 2 and the error bars in Fig. 6 give the standard errors of those means.
The ranking order of the seven two-tones obtained by an *attitude scale construction* according to Edwards (1957) was equal to that of the seven mean point numbers in Table 2 and Fig. 6. That ranking order is similar to that obtained by Kaestner (1909); see Fig. 3 and Table 6 below.

3. PEAK-HEIGHT EXPERIMENT AT BASS PITCH

The details of this second experiment were similar to those of the female-singing pitch experiment described in the previous section; the fundamental frequency of the deeper of the two simultaneous bowed-string-like tones, however, was 99 Hz; i.e., the note of that tone was *G*2. Twenty-nine subjects (no professional musicians) took part. The histogram of the number of bad judgment pairs was similar to that shown in Fig. 5. Eighteen subjects had ≤ 10 bad pairs and were thus considered as reliable. In this second experiment, too, the total number of judgments *B* was greater than that of judgments *A*. Among the $18 \cdot 42 = 756$ judgments of the 18 reliable subjects, there were 291 judgments A, 316 judgments B, and 149 judgments *X*. The deviation of that ratio 316/291 from 1.0 was again found not to be statistically significant; the corresponding number (~15700) of Monte-Carlo experiments was even greater than that for the first experiment (~12200; see previous section).

Results:

Subject	P1	m3	M3	P4	P5	M6	P8	Sum
1	10.0	0.0	3.0	3.0	6.5	8.5	11.0	42.0
2	7.5	0.0	3.0	9.0	11.0	5.5	6.0	42.0
3	12.0	2.5	7.0	5.0	6.5	7.0	2.0	42.0
4	11.0	0.0	7.0	5.0	4.0	5.0	10.0	42.0
5	7.0	3.0	2.5	3.5	4.0	12.0	10.0	42.0
6	4.0	7.5	11.0	6.5	9.5	2.5	1.0	42.0
7	12.0	0.0	6.5	7.5	5.5	8.0	2.5	42.0
8	7.5	9.0	3.5	5.5	5.5	4.0	7.0	42.0
9	12.0	2.5	3.5	3.5	4.5	6.0	10.0	42.0
10	2.5	6.5	11.5	3.5	9.0	5.5	3.5	42.0
11	3.0	6.0	5.0	10.0	11.0	5.5	1.5	42.0
12	2.0	10.5	10.5	5.0	2.5	8.5	3.0	42.0
13	10.5	4.0	10.5	4.5	2.0	3.0	7.5	42.0
14	6.0	6.0	6.5	6.5	6.5	4.0	6.5	42.0
15	4.0	2.5	4.0	6.0	6.5	7.0	12.0	42.0
16	11.0	2.0	6.0	3.5	6.5	6.5	6.5	42.0
17	11.0	6.0	8.0	5.5	7.0	3.0	1.5	42.0
18	10.0	4.0	9.5	3.0	3.0	4.0	8.5	42.0
Mean	7.94	4.00	6.58	5.33	6.17	5.86	6.11	42.0
Standard error	0.84	0.75	0.71	0.47	0.63	0.57	0.86	
x [cents]	0	316	386	498	702	884	1200	

Table 3. Results of the peak-height experiment at bass pitch.

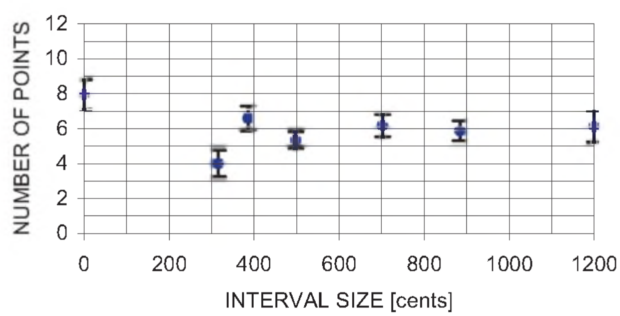

Fig. 7. Mean numbers of points obtained by the seven two-tones in Table 3; deeper-tone frequency $f_1 = 99\,\text{Hz}$.

The resulting mean numbers of points obtained by the seven bass-pitch two-tones are presented in Table 3 and Fig. 7, and are seen to differ strongly from those obtained by the corresponding female-singing-pitch two-tones (Table 2 and Fig. 6).

4. HELMHOLTZ CONSONANCE THEORY

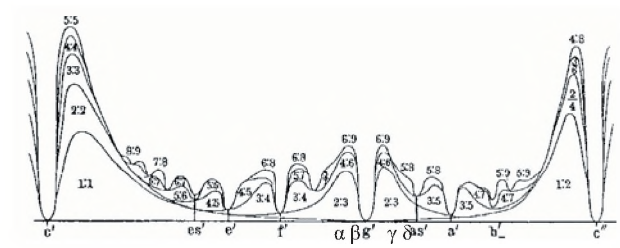

Fig. 8. Reproduction of Fig. 60a of von Helmholtz (1913); deeper-tone frequency $f_1 = 264\,\text{Hz}$; abscissa: interval size (c': 0 cent; c": 1200 cents; g': 702 cents, etc.); ordinate: roughness (≈dissonance). The added symbols α, β, γ, δ are defined in Table 4 below.

In Fig. 8, the fundamental frequency of the deeper of the two simultaneous bowed-string-like tones is assumed to be $f_1 = 264\,\text{Hz}$, i.e., to have the note *C*4. The ordinate of Fig. 8 is the roughness caused by beats generated by pairs of partial tones having nearly equal frequencies.

point	higher tone [Hz]	2nd harmonic [Hz]	beats per second	relative roughness
α	376	752	40	0.5
β	386	772	20	1
g'	396	792	0	0
γ	406	812	20	1
δ	416	832	40	0.5

Table 4. Five points of the curve "2:3" in Fig. 8.

In Table 4, five points of the curve "2:3" in Fig. 8 are listed. The number "3" signifies that one of the two beating partials is the third harmonic of the deeper of the two complex tones and thus has a frequency of 3·264 = 792 Hz. The number "2" means that the other beating partial is the second harmonic of the higher of the two complex tones. At point α, e.g., that second harmonic has a frequency of 2·376 = 752 Hz, so that 792–752 = 40 beats per second are generated. Von Helmholtz judged that these 40 beats per second cause a relative roughness of about 0.5; see Table 4 above and Appendix XV of von Helmholtz (1957). [There is a misprint in the last equation of that appendix: the second of the two brackets in the denominator has to be squared.]

At point β in Fig. 8, the two just defined partials generate 792–772 = 20 beats per second and are judged to cause a relative roughness of 1.0; thus point β is one of the two peaks of the partial-roughness curve "2:3".

My own synthesizer experiments on the roughness caused by two sine-tones of frequencies f_{A} and f_{B} (done at deeper-sine-tone frequencies f_{A} = 132, 264, 528, and 1056 Hz) have yielded roughness maxima at beat rates $b = |f_{\mathrm{B}} - f_{\mathrm{A}}|$ such that

$$b(\text{maximal roughness}) \approx k \cdot \sqrt{(f_{\mathrm{A}} + f_{\mathrm{B}})/2}\ . \qquad (2)$$

My experimental results for the constant k in Eq. (2) were

$$k = (1.068 \pm 0.052)\ \mathrm{s}^{-1/2} \text{ and } k = (1.075 \pm 0.099)\ \mathrm{s}^{-1/2} \qquad (3)$$

at sound pressure levels of 50 and 70 dB, respectively. The maximal-roughness beat-rates given by Eqs. (2) and (3) tend to be lower than those found in the published studies on this topic; see, e.g., Fig. 1 of Terhardt (1974) and Fig. 7 of Kameoka and Kuriyagawa (1969a). [The maximal-dissonance beat rates experimentally determined by these latter authors themselves are particularly high. A different feature of that study, namely the method of adding partial dissonances described in Kameoka and Kuriyagawa (1969b), has been criticized by Vos (1986).]

According to Eqs. (2) and (3), the peak of the curve "2:3" in Fig. 8 (point β; higher-complex-tone frequency 386 Hz) should be shifted to a beat rate of about $30\,\mathrm{s}^{-1}$, i.e., to the left, to a higher-complex-tone frequency of about 381 Hz.

In Fig. 8, the dissonances defined by the various curves such as "2:3", "4:6", etc. are added linearly. Aures (1985) has found that linear roughness addition is valid in a case which shares some features with the present one (namely in the case of two sine tones which are both amplitude-modulated at modulation frequencies of 40 Hz), if the two carrier frequencies differ from each other by several critical band widths or more.

The relative heights of the various curves in Fig. 8 are based on the partial-tone amplitudes of bowed-string tones. According, e.g., to Section 10.2.4 of Terhardt (1998), however, the partial dissonances in Fig. 8 decrease too strongly with rising partial-tone number; e.g., the curve "4:6" is too shallow. The correction of this inaccuracy shifts the total-roughness peak to the right, i.e., back towards its original position in Fig. 8.

It is seen in Fig. 8 that the Helmholtz theory predicts dissonance minima at small-integer fundamental-frequency ratios (e.g., e' = 5/4, g'= 3/2, etc.)

Theoretical consonance curves based on roughness due to beating partial-tone pairs have been presented also in recent years, e.g. in Fig. 12.15 of Terhardt (1998) and in Fig. 11 of Rasch and Plomp (1999). As specified in Table 5, the consonance rankings of the seven dyads treated in Section 2 above defined by these two more recent curves differ little from those defined by the Helmholtz curve (i.e., by our Fig. 8).

frequency ratio of dyad	ranking according to ...		
	Helmholtz	Terhardt	Rasch and Plomp
1/1 (P1)	1(a)	1	1(a)
6/5 (m3)	7	7	6
5/4 (M3)	6	6	7
4/3 (P4)	5	5	5
3/2 (P5)	3	4	3
5/3 (M6)	4	3	4
2/1 (P8)	1(b)	2	1(b)

Table 5. Consonance ranking of the seven dyads treated in Section 2 above, according to the Helmholtz consonance theory and to two of its modern versions (see text).

In Table 6 below, the theoretical consonance ranking order given in the Helmholtz column of Table 5 is compared with the experiment described in Section 2 and with the experiment of Kaestner (1909) [see our Fig. 3].

frequency ratio of dyad	ranking according to ...		
	Helmholtz theory	Experiment Kaestner	Experiment R. F.
1/1 (P1)	1(a)(!)	6(a)	4
6/5 (m3)	7(!)	3(a)	2
5/4 (M3)	6(!)	1	1
4/3 (P4)	5	3(b)	5
3/2 (P5)	3	5	6
5/3 (M6)	4	2	3
2/1 (P8)	1(b)(!)	6(b)	7

Table 6. Consonance ranking of the seven dyads treated in Section 2 above (deeper-complex-tone frequency 264 Hz), according to the Helmholtz consonance theory (Table 5 and Fig. 8), to the experiment of Kaestner (1909; our Fig. 3), and to the experiment described in Section 2 (Table 2 and Fig. 6).

The symbol (!) in Table 6 indicates that this theoretical ranking differs from at least one of the two experimental rankings by 5 or more.

Perfect prime (P1) and perfect octave (P8): Kaestner (1909) noted that his observers found these two dyads "dull", and thus not pleasant. The twelve "reliable" subjects of the experiment described in Section 2 were not asked to comment on their judgments; they, however, attributed few points to these two dyads, too.

Major third (M3) and minor third (m3): The experimental rankings of these two dyads are better than the theoretical ones. Most composers of the last few centuries appear to have shared this preference of thirds at female-singing pitch. For the last 200 years, children have heard equal-temperament thirds (300 cents, 400 cents) more often than just thirds (316 cents, 386 cents). Nevertheless, the consonance peaks in our Fig. 1, and also those in Fig. 3 of Hall and Hess (1984) and in Fig. 2a of Vos (1986) occur at the just-third interval sizes, indicating that such preferences are not due to education, but rather to the physiological properties of our hearing system.

5. A MODIFIED CONSONANCE THEORY

It is hereby postulated that the sensory consonance of a dyad (i.e., of a two-tone) formed of two simultaneous harmonic complex tones is high if the dyad fulfils both of the following conditions.

Condition a:
The dyad contains few or no pairs of partial tones which generate disagreeable beats.

Condition b:
In the excitation pattern generated by the dyad on the basilar membrane of the inner ear there are few or no wide gaps.

The just specified condition a is the "Helmholtz condition"; see, e.g., Chapter X and Appendix XV of von Helmholtz (1954). Condition b (which constitutes our proposed modification of the Helmholtz theory) has been presented first in Section 3.3 of Frosch (2001), and then (in English) in Section 3.4 of Frosch (2002).

The structure of the mammalian cochlea is described in many textbooks, e.g., in Slepecki(1996) and in Zwicker and Fastl (1999). The basilar membrane (BM) in the human cochlea is about 35 mm long. The BM end near the base of the cochlea, i.e., near the oval window, is usually given the coordinate $x_{\mathrm{BM}} = 0$, and the other end, near the apex of the cochlea and near the helicotrema, is said to be at $x_{\mathrm{BM}} \approx$ 35 mm. The BM supports the organ of Corti, which contains inner and outer hair cells (OHC's).

Because of strong amplification in the OHC's of a healthy inner ear, a soft or medium-level sine-tone causes a strong vibration of an about 1 mm long piece of the BM. At high frequency (typically 10 kHz), the vibration peak is near $x_{\mathrm{BM}} = 0$, at low frequency (~0.1 kHz) the peak is near $x_{\mathrm{BM}} = 35$ mm, and at medium frequency (~2 kHz) it is near $x_{\mathrm{BM}} = 17.5$ mm.

In Fig. 6.11 of Zwicker and Fastl (1999), e.g., a linear relationship between the vibration-peak coordinate x_{BM} and the *critical-band rate z* is presented. In Chapter 10 of Hartmann (1998) it is pointed out that the quantity *z* is not a rate, and the better designation *critical-band number* is introduced. Chapter 10 of Hartmann (1998) also contains the definitions of the Bark (or "Munich") scale and the ERB (or "Cambridge") scale; "ERB" stands for "equivalent rectangular bandwidth". In the mentioned first presentation of condition b, Frosch (2001), Munich critical bands were used; in the later publication, Frosch (2002), I switched to Cambridge critical bands. Both these treatments yielded that condition b leads to an improvement of the bad agreement between theory and experiment specified in Table 6 above.

Equation (10.29) of Hartmann (1998) gives the following relation between the sine-tone frequency f and the Cambridge critical-band number z :

$$z = 9.26 \cdot \ln\left(\frac{f}{229\ \mathrm{Hz}} + 1\right); \qquad (4)$$

Hartmann (1998) commented that this equation "possibly ... correlates with the mapping of frequency in the cochlea". As mentioned above, such a correlation appears to be more positively inferred in Zwicker and Fastl (1999); in the remainder of the present study, the following relation between the vibration-peak coordinate x_{BM} and the (Cambridge) critical-band number z is assumed to hold approximately:

$$z \approx 35 - x_{\mathrm{BM}}[\mathrm{mm}]. \qquad (5)$$

In Fig. 9 below, the partial-tone patterns of the seven two-tones treated in Section 2 above are presented.

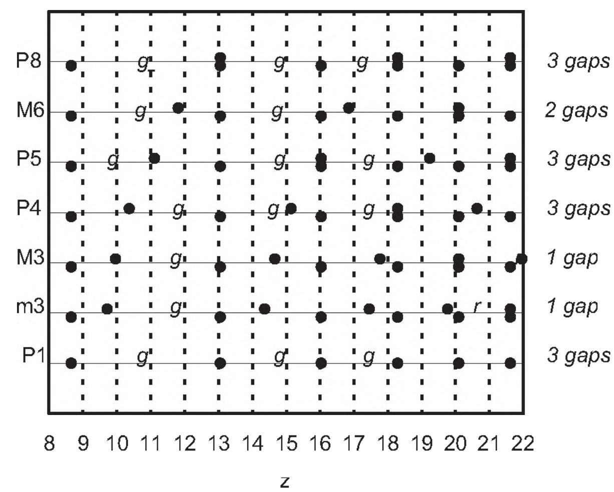

Fig. 9. Partial-tone patterns of the seven two-tones defined in Table 1, at female-singing pitch ($f_1 = 352$ Hz ; note *F*4); details see text.

In Figs. 9 and 10, the critical-band numbers z corresponding, according to Eq. (4) above, to the partial-tone frequencies f occurring in each of the seven dyads are indicated by filled circles. Pairs of adjacent partials separated by a critical-band-number difference $\Delta z > 2$ (and thus yielding, on the BM, excitation peaks separated approximately by a distance $\Delta x_{\mathrm{BM}} > 2$ mm) are conjectured to cause a disagreeable gap, and are indicated, in Figs. 9 and 10, by the symbol *g*. If Figs. 9 and 10 are extended to higher critical-band numbers z, no additional gaps with $\Delta z > 2$ are found in any of the seven dyads. Pairs of partials yielding, via Eq. (2), a constant k ranging from 0.5 to 2.0 are thought to cause strongly disagreeable roughness, and are indicated in Figs. 9 and 10 by the symbol *r*.

In Fig. 9, the major third (M3) and the minor third (m3) are seen to fulfil condition b especially well (only one gap), so that their good ranking in the two experimental columns of Table 6 is understandable.

The minor sixth (m6, R = 8/5, not shown) yields two gaps in Fig. 9, as does the major sixth (M6, 5/3).

In Fig. 9, the deeper-tone complex-tone frequency f_1 has been chosen to be 352 Hz (note *F*4), near the centre of the range of the corresponding frequencies chosen in our Section 2 above (264 Hz), in our Fig. 3 (Experiment of Kaestner, 256 Hz), in our Fig. 8 (Theory of von Helmholtz, 264 Hz), in Fig. 12.15 of Terhardt (1998, theoretical curve, 440 Hz), and in Fig. 11 of Rasch and Plomp (1999, theoretical curve, 250 Hz). If in our Fig. 9 the deeper complex tone *F*4 is replaced by *C*4 (264 Hz), then the only different number of gaps is that of the perfect fourth (P4; 4/3; 2 gaps). If, on the other hand, *F*4 is replaced by *C*5 (528 Hz), then the only different number of gaps is that of the perfect fifth (P5; 3/2; 4 gaps).

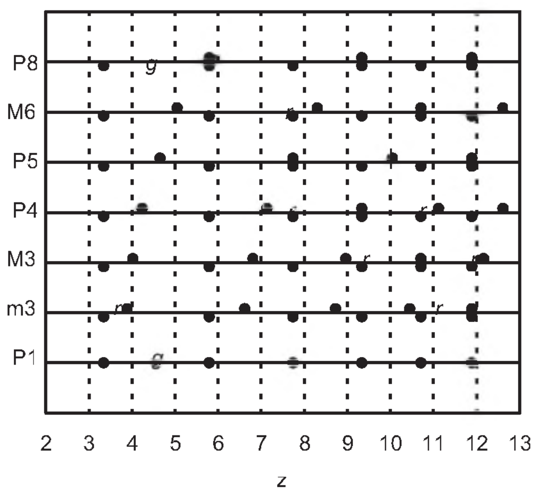

Fig. 10. Same as Fig. 9; bass pitch ($f_1 = 99$ Hz , note *G*2).

According to this modified theory, sensory consonance is *subtle*: An increase of dissonance is avoided if the frequency difference of two neighbouring partial tones is *large enough* (no disagreeable beats) but also *small enough* (no disagreeable gap in the excitation pattern on the BM).

Condition b (i.e., consonance is high if there are no or few wide gaps in excitation pattern on BM) becomes fairly plausible if one considers *dyads of deep tones* (i.e., of tones at bass pitch; deeper of the two simultaneous complex tones at ~100 Hz; see Fig. 10).

In the case of Fig. 10, perfect primes (P1), perfect fifths (P5), and octaves (P8) have no or one gap and no strongly disagreeable beats. In this case, condition b is about equally well fulfilled for all seven stimuli. The thirds (M3 and m3) are predicted to be dissonant because they violate condition a (the Helmholtz condition): each of them includes two pairs of partial tones (marked by the symbol r in Fig. 10) which generate strongly disagreeable beats [constant k in Eq. (2) between 0.5 and 2.0]. Thus the modified consonance theory explains, e.g., that the numbers of points for the thirds (M3 and m3) are high in Fig. 6 (female-singing pitch), but are distinctly lower in Fig. 7 (bass pitch).

6. CONCLUSIONS

The consonance curves for two simultaneous bowed-string-like harmonic complex tones (sensory consonance versus fundamental-frequency ratio $R = f_2 / f_1$) form narrow peaks.

The *R-values* of the peaks ($R = m / n$; m and n are small integers) agree with the Helmholtz consonance theory (rates of beats of partial tones in a certain range cause dissonance).

The relative *heights* of the consonance peaks become more understandable if one considers, in addition to the Helmholtz theory, also “condition b” (sensory consonance is high if there are no or few wide gaps in the excitation pattern caused by the partial tones on the basilar membrane in the inner ear).

REFERENCES

Aures, W. (1985). Ein Berechnungsverfahren der Rauhigkeit. Acustica, 58, 268-281.

Edwards, A. L. (1957). Techniques of Attitude Scale Construction. Irvington Publishers, New York. Chapters 1-3.

Frosch, R. (2001). Mitteltönig ist schöner! Studien über Stimmungen von Musikinstrumenten. Peter Lang, Bern. Sections 3.2 and 3.3.

Frosch, R. (2002). Meantone Is Beautiful! Studies on Tunings of Musical Instruments. Peter Lang, Berne. Sections 3.2-3.6.

Frosch, R. (2003). Psycho-Acoustic Experiment on the Sensory Consonance of Musical Two-Tones. Contribution to Swiss Physical Society meeting in Basel, March 21.

Frosch, R. (2005). Psycho-Acoustic Experiment on the Sensory Consonance of Deep Musical Two-Tones. Contribution to Swiss Physical Society meeting in Berne, July 14.

Hall, D. E. and Hess, J. T. (1984). Perception of Musical Interval Tuning. Music Perception, 2, 166-195.

Hartmann, W. M. (1998). Signals, Sound, and Sensation. Springer, New York. Chapter 10.

von Helmholtz, H. (1913). Die Lehre von den Tonempfindungen. Vieweg, Braunschweig. 6th ed. Chapter 10.

Helmholtz, H. (1954). On the Sensations of Tone. Dover, New York. Chapter X and Appendix XV.

Kaestner, G. (1909). Untersuchungen über den Gefühlseindruck unanalysierter Zweiklänge. Psychologische Studien, 4, 473-504.

Kameoka, A. and Kuriyagawa, M. (1969a). Consonance Theory Part I: Consonance of Dyads. J. Acoust. Soc. Am., 45, 1451-1459.

Kameoka, A. and Kuriyagawa, M. (1969b). Consonance Theory Part II: Consonance of Complex Tones and Its Calculation Method. J. Acoust. Soc. Am., 45, 1460-1469.

Plomp, R. and Levelt, W. J. M. (1965). Tonal Consonance and Critical Bandwidth. J. Acoust. Soc. Am., 38, 548-560.

Rasch, R. and Plomp, R. (1999). The Perception of Musical Tones, in: The Psychology of Music, D. Deutsch, ed., 2nd ed., Academic Press, San Diego, 89-112. Section III C.

Slepecki, N. B. (1996). Structure of the Mammalian Cochlea, in: The Cochlea, P. Dallos, A. N. Popper, and R. R. Fey, eds., Springer, New York, 44-129.

Terhardt, E. (1974). Pitch, consonance, and harmony. J. Acoust. Soc. Am., 55, 1061-1069.

Terhardt, E. (1998). Akustische Kommunikation. Springer, Berlin. Section 12.1.4.

Vos, J. (1986). Purity Ratings of Tempered Fifths and Major Thirds. Music Perception, 3, 221-258.

Zwicker, E. and Fastl, H. (1999). Psycho-Acoustics, Facts and Models. Springer, Berlin. Chapters 3 and 6.