

PERCEPTION OF TONAL CONSONANCE

R. Plomp and W. J. M. Levelt*

Consonant chords, as used in music, are characterized by simple frequency ratios of the constituent tones. Although this relation between ratio simplicity and consonance has drawn considerable interest in the past, no unanimous opinion about its origin exists. Recent developments in hearing theory, equipment design, and measuring techniques justified a new study of the phenomenon. It appeared that the experimental results of this study were confirmed by statistical analysis of the chords of musical compositions.

* Present address: Center of Cognitive Studies, Harvard University, Cambridge (Mass), USA.

PERCEPTION OF TONAL CONSONANCE

R. Plomp and W. J. M. Levelt

One of the oldest discoveries in the field of tone perception, dating back to Pythagoras, concerns the singular character of chords produced by a string vibrating in two parts with length ratios of 1 : 1, 1 : 2, 2 : 3, and 3 : 4, respectively. These tone intervals were called *consonances* and on them the harmony of Western music has been developed, especially so since, in the Middle Ages, other intervals with ratios of 4 : 5, 3 : 5, 5 : 6, and 5 : 8 had been accepted as imperfect consonances. Nowadays we know that these consonant tone intervals are characterized by simple frequency ratios of the constituent tones. As such, however, this change in physical description does not throw more light on the phenomenon of consonance, and opinions are still divided as to its origin. Helmholtz's explanation of the difference between consonant and dissonant intervals on the basis of interference between adjacent harmonics (Helmholtz, 1863) has been opposed by many other investigators who tried to explain the phenomenon in terms of perception of the frequency ratio itself, coincidence of harmonics, difference tones or fusion (for references cf. Plomp and Levelt, 1965).

This divergence of opinions on the origin of consonance, combined with the fact that most research on it dates from before 1920, made it attractive to submit the phenomenon to a new investigation in which full use could be made of more recent developments in hearing theory, equipment design and measuring techniques. In this study, a review will be given of the most interesting results of these experiments, as far as they are available now. To avoid misunderstandings, it may be

useful to emphasize that the sole concern of our investigation was why consonance is related to simple frequency ratio. Although the concept of consonance is rather vague and may be different for musicians and laymen, this relationship is always involved. In our opinion it refers to the peculiar sensorial experience associated with isolated tone pairs with simple frequency ratios. We shall use the term *tonal consonance* to indicate this characteristic experience. The study of this phenomenon was the purpose of our investigations.

Dimensions in the perception of tone intervals

One possible approach towards the question why consonance is related to simple frequency ratios is to determine in which way tone intervals are categorized or, in other words, which criteria are used in discriminating them. This means, in fact, that we are looking for the psychological dimensions in our perception of tone intervals. If frequency ratio is one of the criteria we may expect to find a corresponding dimension in the experimental results. As, after Helmholtz's conception, dissonance is related to interference, it is of interest to examine intervals consisting of simple tones (sinusoids) as well as intervals consisting of complex tones (fundamental plus harmonics).

The best way to investigate the psychological dimensions in the perception of tone intervals is to use a non-verbal technique, in which a subject is asked to compare his impressions of different intervals. Since, in this respect, the method of triadic comparisons is very attractive, it was used in the present investigation. The subject, then, can operate three pushbuttons, each corresponding with a different tone interval. By pressing these pushbuttons successively the subject has to select the two intervals which sound most similar to him and the two which sound least similar. In this investigation (Levelt, van de Geer and Plomp, 1966) 15 different tone intervals were involved, given by the frequency ratios 1 : 2, 2 : 3, 3 : 4, 2 : 5, 3 : 5, 4 : 5, 5 : 6, 4 : 7, 5 : 7, 5 : 8, 4 : 9, 8 : 9, 11 : 12, 8 : 15, 15 : 16. To avoid the influence of pitch as much as possible, all stimuli had the same mean frequency viz. 500 cycles per second (cps). The component tones of the intervals were either simple tones, or complex tones (with harmonics up to 4000 cps, all with equal loudness). The stimuli were reproduced by a loudspeaker, placed in front of the subject at a sound-pressure level of 55-60 dB (normal listening level for speech and music). Since with a set of 15 stimuli $\binom{15}{3} = 455$ different triads correspond, it was impos-

sible to present any subject with the complete set of triads. We therefore developed an incomplete balanced design consisting of 4 blocks, each of 35 triads. Each block was presented to 4 subjects (non-musicians); they all judged one block of stimuli twice, once for simple-tone and once for complex-tone intervals.

As the blocks were so designed that all 105 possible pairs of stimuli were judged once by each subject, we obtained 16 judgements of each pair. On the basis of these responses, for each stimulus pair a 'similarity index' was computed, defined as the number of times that a specific pair was judged more similar than the other pairs. This resulted in two 15×15 matrices, one for simple-tone and one for complex-tone intervals. Any matrix element thus determines the 'psychological distance' between two intervals, a large similarity index corresponding with a small distance and a small similarity index with a large distance. The most suitable technique for further analysis of such similarity matrices is that developed recently by Kruskal (1964a, 1964b). For a better understanding of the problem, and of the way in which the data were handled, the following brief explication may be of value. Suppose that the perception of tone intervals is determined by three different psychological factors. Then, each of the 15 stimuli can be represented as a point in a three-dimensional space with orthogonal axes, the coordinates of the point corresponding to the contribution or loading of each factor in the perception of the stimulus concerned. The similarity matrix informs us about the relative distances between the 15 points. It will be clear that on the basis of this matrix the best-fitting configuration of the points can be computed. The only criterion of Kruskal's technique is that the rank order of the distances between interval points in the stimulus space must be the inverse of the rank order of the corresponding similarity indices. The degree to which this criterion is not fulfilled can be expressed in a percentage of 'stress'. In our case we do not know a priori how many independent factors are involved. Therefore, perhaps a space with more than three dimensions is required to nonviolate the data. On the other hand, 15 points can anyway be fitted in a 14-dimensional space, meeting Kruskal's criterion with zero stress. Of course, in the final solution that configuration will be taken which has the minimum number of dimensions still acceptable. The best-fitting configurations were computed in all spaces with 10 or less dimensions. This was done by means of Kruskal's MDSCAL computer programme. On the basis of what Kruskal considers as 'fair' stress (10%), it appeared that, both for

simple-tone and complex-tone intervals, a three-dimensional space sufficed to fit the data.

For the interpretation of this result it is of interest to look for 'common dimensions' in both spaces. We mean by this that, in the simple-tone interval space, a direction might be found for which the projections of the interval points on an axis in that direction closely correlate with the projections on an axis of the corresponding points (same frequency ratio) in the complex-tone interval space. Computation showed that both spaces had two dimensions in common, in other words in each space a plane was found for which the projections of the corresponding interval points are near to each other (correlation coefficients of 0.935 and 0.944 for the two dimensions, respectively)*. The projections of the stimulus points on this plane are not evenly distributed over the plane but appear to be ordered along a horseshoe-like pattern with interval width (frequency difference between the constituent tones) as a parameter. Without going into further details we may conclude that this relation shows that we have in fact only one underlying dimension, common for both simple-tone and complex-tone intervals, namely interval width.

Furthermore, the question is of interest whether there is a dimension in the perception of tone intervals related to frequency ratio, for we know that consonance depends on it. For that reason in both the simple-tone and complex-tone spaces, a direction has been determined for which the projections of the interval points on an axis in that direction has a maximum correlation with ratio simplicity. Although other measures are also acceptable, the frequency of the lowest common harmonic was taken as a criterion for ratio simplicity (for instance for the interval 1 : 2 the actual frequencies were 333 and 666 cps, so 666 cps was taken; for 2 : 3 the frequencies were 400 and 600 cps, giving 1200 cps, etc.). As the distribution of these frequencies is rather skew, the logarithm was used in the computations. It appeared that there existed an a priori correlation between interval width and ratio simplicity, owing to the intervals selected; therefore, measures had to be taken to eliminate this artefact from our results. After correction for it, we found that, in the complex-tone interval space, a direction could be determined giving a high correlation ($r = 0.914$) with ratio

* We are much indebted to J. P. van de Geer for providing the techniques to compute the common dimensions and the dimension for maximum correlation with ratio simplicity described below.

simplicity, whereas in the simple-tone interval space this was not possible ($r = 0.215$).

Summarizing the results of this experiment, we may conclude that tone intervals, whether consisting of simple or complex tones, are differentiated on the basis of interval width and that, in addition, only complex tones are also differentiated on the basis of simplicity of frequency ratio. This conclusion implies that, for simple-tone intervals, tonal consonance as a sensorial experience can only be related to interval width and not to ratio simplicity. This means that all explanations of consonance in which the harmonics do not play a rôle, have to be abandoned.

Relation between consonance and interval width for simple-tone intervals

Knowing that, for simple-tone intervals, consonance is related to interval width, it is of interest to focus attention on this relation. It was investigated through experiments (Plomp and Levelt, 1965) in which subjects judged simple-tone intervals with different interval width and mean frequency on a 7-point rating scale, 1 corresponding with most dissonant and 7 with most consonant. Some subjects asked for the meaning of 'consonant'. In that case the term was circumscribed as 'beautiful' and 'euphonious'. This is justified because, as had been found earlierly (Van de Geer, Levelt and Plomp, 1962), 'consonant', 'beautiful' and 'euphonious' are highly correlated for naive subjects. The tones were reproduced by a loudspeaker at a sound-pressure level of about 65 dB.

The experiments were carried out for mean frequencies of the intervals of 125, 250, 500, 1000, and 2000 cps. Each subject participated only in one test session in which he had to judge 12 to 14 different interval widths around one of these mean frequencies. The intervals were presented five times in a random order, and only the data of those subjects were maintained who gave sufficiently consistent responses ($r > 0.5$ between the scores of first and last series). In this way, results were obtained for about ten subjects at each mean frequency.

As an example, in Fig. 1 the results for intervals around 500 cps are reproduced. At other mean frequencies similar curves were obtained, thus indicating that there is a clear minimum in the consonance score followed by a broad maximum for wider intervals. This result confirms the validity of Helmholtz's assumption that consonance is related to interference. Two tones very near to each other, give slow beats which

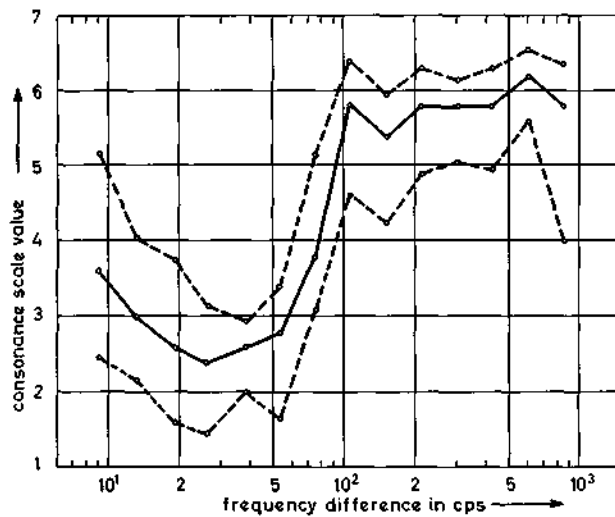


Fig. 1. Consonance rating scores of simple-tone intervals with a mean frequency of 500 cps as a function of frequency difference between the tones. The solid curve corresponds with the median, the dashed curves with the lower and upper quartiles of the scores (averaged over 11 subjects).

are evaluated as consonant. For larger interval widths, these beats are so rapid that the sound obtains a rough and dissonant character which disappears for still wider intervals. However, it was found that, contradictory to Helmholtz's view on the matter, the frequency difference for which the interval is most dissonant, depends on the mean frequency of the interval; this is also the case for the interval width for which consonance score increases no longer.

The question can now be asked whether these data can be related to other properties of hearing. They can, indeed. In recent years, many investigations have been published in which the concept of the 'critical band' plays an important rôle. This critical band, whose width is a function of frequency, can be considered as the resolving power of the hearing organ for sounds of different frequencies (for more details, see the second study). A comparison with the present results shows that tone intervals wider than critical bandwidth are judged as consonant, whereas maximum dissonance occurs for an interval width around a quarter of critical bandwidth. This supports the assumption that dissonance is due to interference.

Consonance for complex-tone intervals

In practice, tones as produced by musical instruments, usually consist of a fundamental and a number of harmonics. This implies that, for intervals composed of these tones harmonics of one of them may interfere with harmonics of the other. Accordingly, the degree to which this occurs will affect the consonance value of the interval. This influence can be illustrated in the following way. From the results of the preceding experiments a standard curve was derived which represents consonance – or dissonance – of simple-tone intervals as a function of interval width with critical bandwidth as a unit (Fig. 2). Assuming that the total dissonance value of a complex-tone interval is equal to the sum of the dissonance values of each pair of adjacent harmonics, this total can be computed by using the right-hand scale of Fig. 2. Although this assumption is rather speculative, it is not unreasonable as a first approximation. Its use may be justified in illustrating how it predicts, for complex-tone intervals, the dependence of consonance on interval width and frequency ratio. On this basis the curves of Figs. 3 and 4 were computed for complex tones consisting of six harmonics. The first figure shows in which way consonance varies as a function of interval width, whereas the other graph illustrates how the consonance of some intervals, given by simple frequency ratios, depends on frequency.

The curves illustrate: (1) peaks of the curve of Fig. 3 correspond with

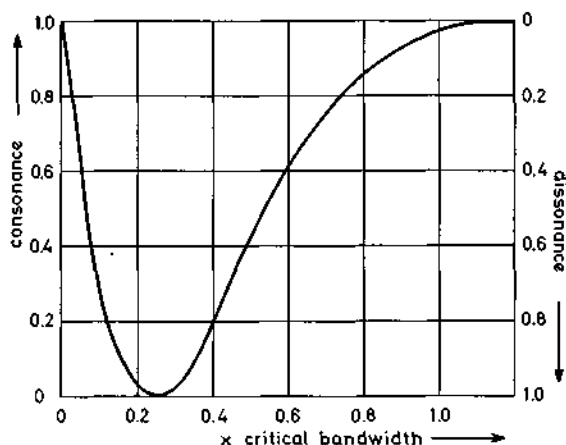


Fig. 2. Standard curve representing consonance of simple-tone intervals as a function of frequency difference with critical bandwidth as a unit. The consonance and dissonance scales are arbitrary.

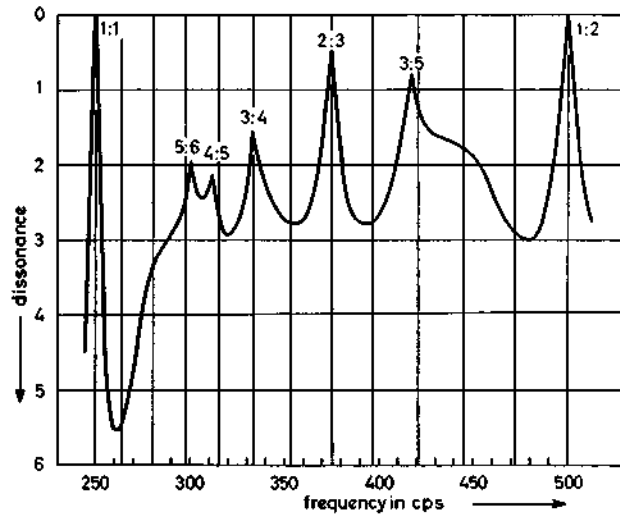


Fig. 3. Illustration of the way in which consonance of an interval with a lower complex tone of 250 cps and a variable higher one depends on the frequency of this tone. Both complex tones consist of six harmonics. The vertical lines represent interval widths after the equally-tempered scale.

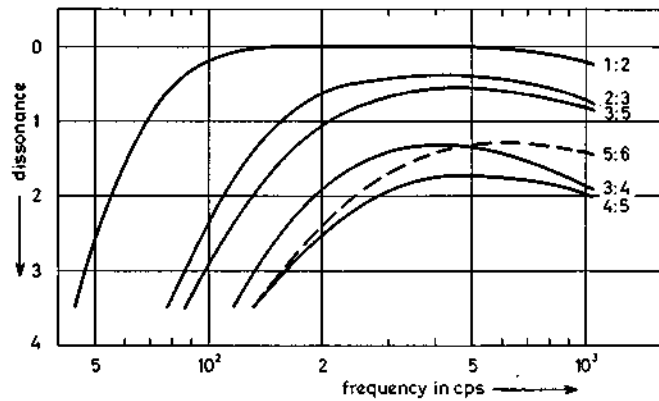


Fig. 4. Illustration of the way in which consonance of some intervals with simple frequency ratio depends on the frequency of the lower tone. Both complex tones consist of six harmonics.

simple frequency ratios of the component tones; this shows that, for complex tones, consonance is related to these simple ratios; (2) more simple ratios are represented by sharper peaks, so the octave (1 : 2) and fifth (2 : 3) are much more sensitive to a deviation from

their correct ratio than are other intervals; this explains why in the equally-tempered scale (vertical lines of Fig. 4) the impure thirds (4 : 5 and 5 : 6) are much more tolerable than impure octaves and fifths would have been; (3) the relative heights of the peaks of Fig. 3 and the curves of Fig. 4 fit rather well the rank order of consonant intervals as accepted by musicians; (4) with decreasing frequency consonance is better preserved for more simple than for more complex frequency ratios (Fig. 4), reflecting the musical practice to avoid thirds (4 : 5, 5 : 6) at low frequencies.

Statistical analysis of chords in music

The close relation between consonance and critical bandwidth raised the question whether in music, too, we may find this relation. As the experiments showed, maximum dissonance corresponds with an interval width of about a quarter of the critical bandwidth, whereas consonance increases no longer for interval widths exceeding this critical value. This suggests that especially this range will be of interest in music to differentiate between more consonant and more dissonant chords. To check this assumption statistical analyses of the chords in some musical compositions, the 'vertical' dimension of music, were carried out.

An illustration may serve to explain how the analyses were done. Suppose we are interested in the occurrence of intervals with $c^2 = 523.2$ cps as the lower tone. We then sort out all chords containing c^2 and a higher tone simultaneously and determine the fraction of time during which the nearest higher tone is separated from c^2 by a distance of 1 semitone, by 2 semitones, etc. It may appear that distances of 1 semitone as well as of 15 semitones are rare, whereas a distance of 4 semitones is rather common. Of this 'density distribution' of intervals we can determine the 50% (median)-, 25%- and 75%-points, respectively. These numbers represent a good measure of the widths of intervals with c^2 as the lower tone. By repeating the procedure for other tones over the relevant frequency range, we can find how the 50%-, 25%- and 75%-points depend on frequency.

As in practice musical tones nearly always consist of a number of harmonics with amplitudes comparable with the amplitude of the fundamental, we are also interested in density distributions in which harmonics have been taken into account. This can be done, popularly said, by plotting the notes of the first n harmonics in the score of the

music and treating these additional notes in the same way as the original ones. In this way density distributions can be determined for various values of n , premising that the additional notes fit our tone scale. This is rather well the case up to $n = 10$.

Density distributions as a function of both frequency and number of harmonics were computed, using especially developed equipment, for parts of musical compositions of J. S. Bach, A. Dvořák and A. Schoenberg. In Fig. 5, some results are reproduced graphically, based on the third movement of Dvořák's String Quartet Op. 51. The density distributions of the other compositions led to similar graphs. A comparison of the solid curves with the dashed curves corresponding with critical bandwidth and a quarter of this bandwidth, respectively, shows that, for a number of harmonics representative for musical instruments, all curves have about the same shape. Moreover, the situation of the solid curves demonstrates that most intervals have a width between critical bandwidth and a quarter of it.

These results suggest that critical bandwidth plays an important rôle in music. Apparently the region over which for simple tones the consonance impression strongly depends on interval width is used for 'modulation' between more consonant and more dissonant chords. We should realize that this equally deep 'penetration' in the borderland between pronouncedly consonant and dissonant sounds is a result of many such factors as the Western tone scale, the number of simultaneous tones and the primary intervals selected by the composer. Concerning the latter factor, a comparison of the density distributions of three compositions showed that in modern music the just-mentioned borderland is penetrated more deeply than in older music.

Conclusions

The investigations strongly suggest that the relation between tonal consonance and simple frequency ratio, as it has been found in practice, is a result of interference of adjacent harmonics. The fact that the

Fig. 5. Results of a statistical analysis of the chords of the third movement of A. Dvořák's String Quartet Op. 51 in e-flat major with n (= number of harmonics taken into account) as a parameter. The solid curves represent the 25%-, 50% and 75%-points, respectively, of the cumulative density distribution of intervals, plotted as a function of frequency. The dotted curves correspond with critical bandwidth and a quarter of this bandwidth.

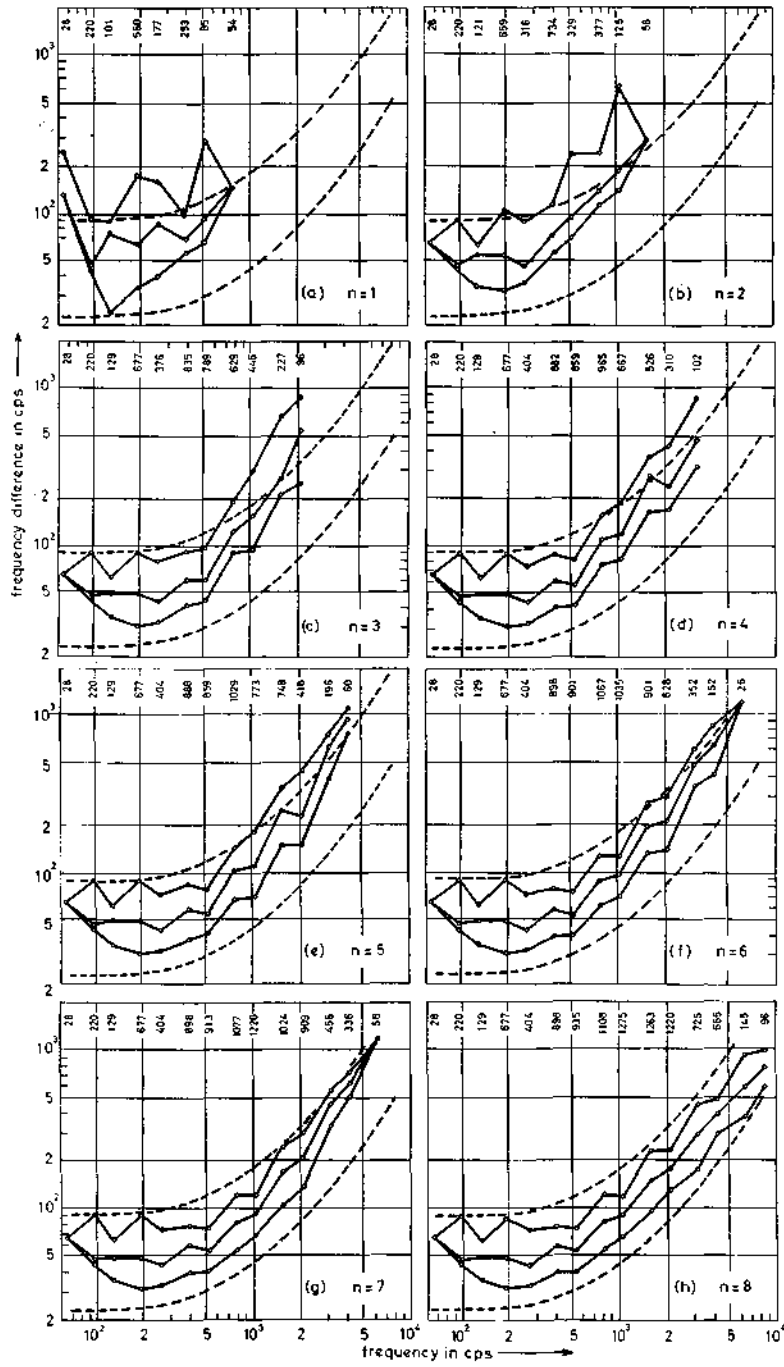


Fig. 5

perception of intervals consisting of simple tones is not governed by frequency ratio, but only by interval width, is confirmed by the finding that for naive subjects, the consonance value attached to this type of intervals is a continuous function of frequency difference. The experimental fact that frequency ratio is an important factor in discriminating complex-tone intervals can be explained by using the hypothesis that, in this case, not only the fundamentals interfere, but that adjacent harmonics do too. Moreover, the experiments showed that this interference only occurs for frequency distances within the 'critical bandwidth'; the most dissonant intervals correspond with a frequency distance of about a quarter of this bandwidth. This relevance of critical bandwidth for the perception of tonal consonance is supported by the results of statistical analyses of chords in musical compositions.

REFERENCES

- HELMHOLTZ, H. L. F. VON (1863), *Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik*. F. Vieweg & Sohn, Braunschweig.
- GEER, J. P. VAN DE, W. J. M. LEVELT and R. PLOMP (1962), The connotation of musical consonance. *Acta Psychol.* 20, 308-319.
- KRUSKAL, J. B. (1964a), Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika* 29, 1-27.
- KRUSKAL, J. B. (1964b), Nonmetric multidimensional scaling: a numerical method. *Psychometrika* 29, 115-129.
- LEVELT, W. J. M., J. P. VAN DE GEER and R. PLOMP, Triadic comparisons of musical intervals. *Brit. J. Mathemat. Statist. Psychol.*, in press.
- PLOMP, R. and W. J. M. LEVELT (1965), Tonal consonance and critical bandwidth. *J. Acoust. Soc. Am.* 38, 548-560.