ARCHIVES OF ACOUSTICS
13, 1-2, 147-156 (1988)

# INTONATION OF TONE SCALES: PSYCHOACOUSTIC CONSIDERATIONS

ERNST TERHARDT

Institute for Electroacoustics, Technical University, Munich (FRG)

Conventional music theory ordinarily is based more or less explicitly on the concept that intonation of musical tones can with sufficient precision be described by one physical parameter, i.e. "tone frequency". Ratios of "tone frequencies" play a predominant role in theories of consonance and tone scales. Closer inspection of the physical nature of musical tones, and particularly of pitch-dependent auditory effects, reveals, however, that the aforementioned classical concept is insufficient. A psychoacoustically-oriented dualistic approach to intonation is suggested maintaining that "correct" intonation of a musical tone interval basically depends on (1) harmonic purity, and (2) sensory purity. Harmonic purity depends on memorized pitch-interval templates which in turn are partly of natural, partly of cultural origin. Sensory purity largely depends on perception of fluctuations, and its basic aspects are independent of cultural effects. It is concluded that optimal intonation is a compromise which at every moment of a musical performance must be achieved by active evaluation of the two aforementioned criteria. Theories of perception of pitch and fluctuations readily explain why this is so, and provide promising tools for achieving that compromise.

## 1. Introduction

The problem of tone scale intonation has been discussed and attacked for some thousand years with remarkably little success. This paper attempts to provide that problem with some new conceptual aspects which may be helpful when modern psychoacoustic methods are considered along the way to a final solution. Actually, scale intonation can hardly be properly discussed without taking into consideration how tone scales as such may have been developed. Fortunately, it is not essential to know every detail of that development; rather it is regarded sufficient to have a concept about what the route of development could have been, on the basis of a number of universal and consistent auditory criteria. In that sense, the introduction to the intonation problem following in the next paragraph itself is part of a proposed concept.

The pitch dimension, which basically is continuous, was dissected into discrete pitch categories quite early in history. As is well known, musical tones bear certain relationships to each other not only in terms of their pitch differences but also in terms of certain additional qualities. Those qualities appear to have provided the criteria for tone categorization from the very beginning of music. As was made quite obvious by Pythagoras' principle of tone scale generation, i.e. concatenation of fifths and octaves, the most pronounced tonal affinities, i.e. octave- and fifth-affinity, provide the criteria necessary and sufficient to explain why tone scales could hardly have developed in a different manner than they actually did. In this view, the development from the pentatonic through the diatonic to today's chromatic scales appears straightforward and cogent. This is so, at least, if one for a moment ignores the intricacies of intonation and just considers tone categories. Probably, the main mistake inherent in most of the classical tone scale theories is an intermingling of the aspects of tone category and tone intonation. As will be further pointed out below, tone intonation is much more complex than ordinarily has been assumed; in particular, it is not just a matter of small-integer "tone-frequency" ratios.

Thus when we first ignore intonation problems, the development and typical features of tone scales may be seen as follows. Each of the three scales shown in Fig. 1 are an ordered collection of tone categories, i.e. notes, arranged according to their height; and the notes of each scale have been chosen according to the criteria of octave- and fifth-affinity, respectively. However, neither the pentatonic nor the diatonic scales are conclusive in a sense. While in these scales for any arbitrarily chosen tone another can be found which is in an octave relationship, this is not true in each case for fifth-third-, etc. relationships. This can generally be expressed by saying that in these two scales *interval width* (i.e. the number of steps encompassed by two tones) is not an unequivocal indicator of *interval quality* (where quality means the type of tone affinity pertinent to a particular interval category; e.g. the qualities of

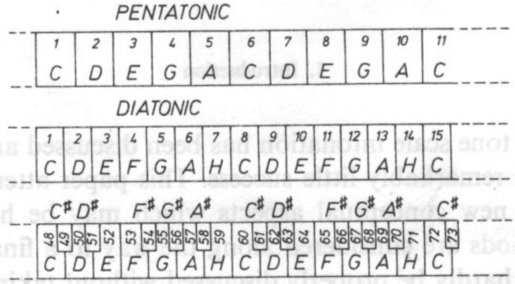


FIG. 1. Schematic representation of the pentatonic, diatonic, and chromatic tone scales. They can be regarded as representing different states of one and the same line of development, i.e. selection of notes in terms of octave- and fifth-affinity. While octave periodicity is implemented in all of them, the first two scales are not conclusive with respect to providing fifths and other intervals. The chromatic scale represents the final state of development, as interval width i.e. number of steps between two tones is an unequivocal indicator of interval quality (i.e. "octaveness", "fifthness", etc.), no matter which note is chosen as an interval's basis

"octaveness", "fifthness", etc). In this sense it can be said that today's 12-note chromatic scale in fact is conclusive, as interval width unequivocally indicates interval quality, no matter which tone is chosen as an interval's basis.

Actually it is that characteristic of the chromatic scale that determines its advantage of being "well-tempered". The aforementioned way of expressing the chromatic scale's unique features is different from, yet musically more relevant than saying that the chromatic scale is obtained by subdividing the octave into twelve "equal steps". The insufficiency of the latter statement is that it does not say in what respect the steps are equal.

Conceptually, a musically relevant theory of tone scales and intonation can hardly be achieved unless the basic difference between "scale" and "intonation" is understood and strictly observed. Here the "Three Worlds Concept" put forward by POPPER [6] and ECCLES [3] is extremely helpful, as it provides to the problem a perfectly fitting frame. That concept maintains that there are three basically different areas ("worlds"), into which all human experiences can be assigned. These are: 1) the world of physical/chemical processes and states, i.e. the "real world", it is called "World 1"; 2) the world of sensory experiences in the widest sense ("World 2"), and 3) the world of information, in particular of symbolically represented products of the human brain ("World 3").

In fact, music exists in three fundamentally different representations, each of which pertains to one of the three worlds (Fig. 2). In World 1, music exists as sound; in World 2, as auditory sensation, in World 3, as a score. In that conceptual frame, a musical tone scale as such is a symbolic representation of tone categories, i.e. notes;

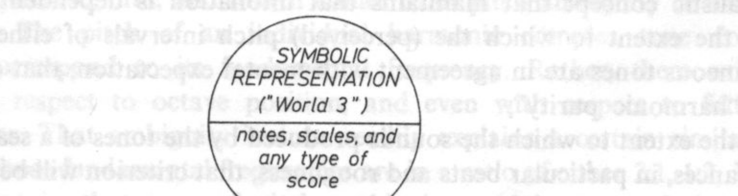


FIG. 2. The three "worlds" of musical reality. In this concept, a musical tone scale is a collection of notes, i.e. symbolic representation of tones; it is pertinent to "World 3". Intonation of a scale is a collection of corresponding physical parameters such as part-tone frequencies and amplitudes ("World 1"). Auditory sensation such as pitch, and roughness provide the decisive criteria of intonation ("World 2")

it is pertinent to World 3. The scale's intonation implies all the physical parameters (mainly, but not only, frequencies) which are correlated with the auditory sensations produced by the sound these are mainly the pitch and beat sensations, as will be pointed out below; those physical parameters are obviously pertinent to World 1. Finally, the auditory sensations mentioned are pertinent to World 2. The theory of tone scales must be concerned with the relationships and interactions between those three manifestations of musical tones. Those relationships can generally be termed *realisation*, *perception*, and *abstraction* (Fig. 2).

In the frame of that concept we can now start the discussion of tone scale intonation with the help of the following two statements:

a) Intonation is dependent on auditory perception and abstraction. It is neither purely by mathematical (numerical) nor purely by phisical arguments that one can decide about intonation. Intonation of a tone scale is optimal if it satisfies the complex mechanisms of auditory perception and abstraction at any instant of a musical performance.

b) Intonation as such means specification of those physical sound parameters which affect the perceptual and abstract criteria mentioned. Besides the "fundamental frequency" of tones, virtually every other parameter such as spectral envelope and sound pressure level must be taken into consideration.

## 2. Criteria of intonation

The most promising, yet little recognized, concept of intonation appears to be the dualistic concept that maintains that intonation is dependent on:

1) the extent to which the (perceived) pitch intervals of either successive or simultaneous tones are in agreement with mental expectation; that criterion will be called "harmonic purity";

2) the extent to which the sounds produced by the tones of a scale are free from disturbances, in particular beats and roughness; that criterion will be called "sensory purity".

These two criteria have been described earlier as establishing a useful concept of musical consonance (TERHARDT [10], [11]. In particular, the term "harmony" was given a special meaning, implying the most basic musical phenomena: *tonal affinity*, *compatibility*, and *fundamental-note-relationship*. Of these phenomena it is in particular tonal affinity (i.e. octave-, fifth-, and fourth-affinity) that is important for scale intonation.

The basis of harmonic purity is provided by pitch. Each musical interval is represented by a certain distance on the pitch continuum. There is no justification for presuming that a particular interval quality corresponds to one and the same distance on every level of the pitch continuum. Rather, the pitch distance corresponding to a particular music interval in general will be dependent on absolute pitch height. The actual pitch distance by which a particular musical interval is represented on the continuum depends on the metrics by which pitch height is measured. Fortunately, for the theory of musical intervals and intonation it is not necessary to make a decision on the metrics. It is sufficient to be aware that to each musical interval quality corresponds a certain pitch distance that even may be a function of pitch height, and that can be compared with corresponding memorized distances.

The decision of whether or not a given tone interval is "harmonically pure" will depend, among other effects, on the precision by which the corresponding pitch distances are represented in memory. What can be said about the origin and the precision of that representation?

On the basis of psychoacoustic evidence, there is no indication that *every* interval of today's chromatic scale is of natural origin, i.e. either acquired in basic perceptual processes or "hard-wired" in the auditory nervous system. Research on children's intonation in singing appears to indicate a considerable amount of learning in early life. While young children may show a good sense of melodic contour, they usually care little about tone- and interval-categories, let alone precise intonation (DOWLING [2]). Appreciation of harmony appears to develop relatively late (SHUTER-DYSON [8]). It will thus appear reasonable to assume that for instance second- and perhaps even third-intervals, and in particular their intonation, are to a considerable extent culturally acquired.

On the other hand, there are a number of solid arguments in favour of the presumption that the most basic intervals, i.e. octave, fifth, and fourth, are acquired, or predominantly determined, by auditory spectrum analysis in the perception of natural speech (cf. TERHARDT [11]). Additional support to the basic and natural character of the octave, fifth, and fourth comes from pitch ambiguity of harmonic complex tones. The pitch of an individual harmonic complex tone does *not unequivocally* correspond to its fundamental frequency. Rather, there exists an ambiguity with respect to octave position, and even with respect to fifth- and fourth-confusions. That ambiguity in turn readily explains a certain similarity of musical tones whose fundamental frequencies are in a ratio of either 2:1, 1:2, 3:2, 2:3, 4:3, or 3:4. That is, that type of pitch ambiguity explains one of the crucial phenomena involved in tonal music: tonal affinity of octaves, fifths, and fourths. Pitch confusions in terms of third- and other intervals have practically never been observed. It is in this sense that octaves, fifths and fourths can be regarded as more "natural" than the rest of intervals.

If *anything* can be concluded from experimental data and observations, it appears reasonable to assume that the pitch distances stored in memory as representations of the harmonic aspects of musical intervals are of natural origin in the cases of the octave, fifth, and fourth; and for the rest of intervals are culturally dependent. If this is so, then it obviously does not make sense to raise the question of what the "natural" intonation is of thirds, seconds, etc. Experiments designed to determine the optimal intonation of those culturally dependent intervals will necessarily reflect only the listeners' mean previous intonation experiences. Results of that type of experiment in fact confirm that view (for an extensive review see BURNS and WARD [1]).

With respect to the octave, fifth and fourth, the memorized pitch distances corresponding to them may be affected by actual intonation of real musical instruments as well. Therefore, the harmonic purity of those intervals should be regarded as possibly dependent on both a natural and a cultural component. A sort of competition between these two components may be observed in auditory tuning of octaves of successive tones by a trained musician: while his/her "natural" (and thus "naive") evaluation will ordinarily produce a considerable stretch (i.e. a fundamental-frequency ratio larger than 2:1; cf. WARD [14], WALLISER [13], TERHARDT [11]), cultural experience in ensemble playing may prevent him from stretching the octave by too great an amount. Evidence for that conclusion may be seen for example in recent results by MAKEIG [5].

Of the two intonation criteria, i.e. harmonic purity and sensory purity, the first can be regarded as functionally most relevant; in particular, it applies to both successive and simultaneous tones. Sensory purity, which virtually is dependent on fluctuation effects, i.e. beats, will in many cases provide another strong criterion of intonation which is significant only with simultaneous tones. As was extensively pointed out by HELMHOLTZ [4], musical chords composed of harmonic complex tones attain the highest sensory purity if the fundamental frequencies of the complex tones are in ratios of small integer numbers, e.g. 4:5:6 for the major triad. If the fundamental frequencies depart just a little therefrom, beats can be heard which ordinarily are disliked as being indicative of "mistuning". Harmonic purity is hardly affected by such a small amount of mistuning. As is common experience, slight mistuning does not disturb or even destroy the essential "musical message" but can considerably reduce the sensory pleasantness of a sound. In that sense harmonic purity may be regarded as "functional", while sensory purity is "cosmetical".

In general, the intonation which provides maximal harmonic purity is not necessarily identical with the intonation producing maximal sensory purity. For example, when two successive tones are tuned to give an optimal octave, their fundamental frequencies will turn out to be in a ratio slightly greater than 2:1. As a consequence, the same tones, when sounding simultaneously, may produce audible beats, i.e. less than optimal sensory purity. The latter would be achieved by tuning the tones exactly in the ratio 2:1, which however would render harmonic purity less than optimal. Obviously, intonation in general is a compromise.

So far we have been concerned only with the intonation criteria which depend on tone intervals, as opposed to individual tones. Naturally, intonation of individual, isolated tones depends on absolute-pitch recognition and thus is confined to listeners having absolute pitch. Since only a small percentage of musical listeners possess that ability, and since the development of musical scales is essentially dependent on tone intervals rather than on recognition of individual tones, the latter case is disregarded in the present study.

The criteria of intonation and their role in musical scales can thus be summarized as follows:

Intonation of successive tone intervals is exclusively governed by harmonic purity, i.e. matching of pitch distances to corresponding "templates" stored in memory.

The "templates" which correspond to octaves, fifths, and fourths, probably are of natural origin, i.e. independent of previous musical experiences. The "templates" corresponding to the rest of intervals probably are essentially dependent on previous experience and learning; i.e. they can be developed only if musical scales already exist.

Intonation of simultaneous tone intervals is dependent both on harmonic and sensory purity. In many cases, the latter will provide the most sensitive criterion, and in principle it will apply to all intervals, be they of natural or cultural origin.

Since, with a given musical sound and intonation, it is not always possible to fully satisfy both the criteria of harmonic and sensory purity, intonation generally is a compromise. This implies that there does not exist such a thing as a fixed ideal intonation.

To further understand the advantage provided by the present approach it will be helpful to critically discuss the classical concept of describing intonation merely in terms of "tone frequency" alone.

### 3. Criticism of the "tone-frequency concept of intonation"

In musical acoustics and music theory, the intonation of tones is ordinarily described by "the" frequency, in the sense of "oscillations per second". However, to make that concept fully valid and significant, two preconditions must be fulfilled. The first is a physical one: the oscillation frequency as such must be defined with sufficient precision; i.e. the tone's oscillations must be strictly periodic. The second condition is a psychoacoustical one: the pitch sensation must be *solely* dependent on the oscillation frequency, i.e. pitch must not depend, for instance, on the tone's spectral composition and sound pressure level. If either one or both of these conditions are violated, any discussion of intonation based on "tone frequency" becomes more or less inadequate. In fact it turns out that for certain types of tones (in particular, percussive tones such as of the piano) the first condition is violated; and the second condition is violated for practically every type of tone.

Concerning piano tones, it is well known that their spectra are slightly inharmonic (Schuck and Young [7]). This means that the period of the tone's entire oscillation is not identical with that of the first partial alone: rather, the former is much longer. As a piano tone's pitch corresponds approximately to the period of its first partial, one can also say that its entire physical period is by far longer than that corresponding to its pitch. In other words, in this case the concept of associating the tone's pitch with its physical period fails profoundly. It is thus apparent that neither harmonic nor sensory purity can be properly accounted for by the concept of "tone frequency". In particular, sensory purity of the piano tone scale cannot consistently be evaluated by the lowest partial frequencies alone, since it depends on beats between higher partials as well; therefore also precise specification of the frequencies of higher partials is required.

Musical tones which are with sufficient precision periodic are produced by instruments with a steady energy supplement, i.e. strings, horns, woodwinds, and organ pipes. The higher partials of those tones are harmonics, i.e. their frequencies are with sufficient precision determined by that of the fundamental, i.e. they are just integer multiples of the latter. Therefore assessment of sensory purity (beats of simultaneous part tones) can be accomplished largely by knowledge of the fundamental frequencies alone (that in turn are identical with oscillation frequencies in that case). However, beats are further dependent on the amplitudes of partials and,

**Table 1.** Typical magnitudes of effects relevant to intonation. Upper part: "numerical effects", i.e. phenomena which have been considered in terms of the "tone frequency approach". Middle part: piano string inharmonicity as an example of a physical effect. Lower part: aural effects which are pitch dependent and thus relevant to "harmonic purity" of intonation. Further descriptions and explanations of the latter effects can be found, e.g. in TERHARDT et al. [12].

| | equivalent frequency cents | difference percent |
|---|---|---|
| **Numerical effects** | | |
| Pythagorean comma, $(3/2)^{12}:2^7$ | 23.4 | 1.36 |
| synthonic comma (departure of II-VI from pure fifth in just intonation), 81/80 | 21.5 | 1.25 |
| difference between Pythagorean and natural major third, $(3/2)^4:5$ | 21.5 | 1.25 |
| difference between natural and tempered fifth $(3/2):2^{7/12}$ | 2.0 | 0.11 |
| difference between natural and tempered major third | −13.7 | −0.79 |
| **Physical effect** | | |
| inharmonicity of aurally relevant partials of piano-bass-strings (typical example) | 30 | 1.8 |
| **Pitch-dependent, i.e. aural effects** | | |
| just noticeable pitch difference of musical tones (typical example) | 5 | 0.3 |
| aural pitch shifts of true harmonics of complex tones: 1st harmonic | −20 … 0 | −1 … 0 |
| higher harmonics | 0 … 160 | 0 … 10 |
| difference between true and nominal pitch of musical tones: low tones | −50 … 0 | −3 … 0 |
| high tones | 0 … 50 | 0 … 3 |
| octave enlargement (octave matching of successive tones) | 9 … 50 | 0.5 … 3 |

moreover, the perceived pitch of any complex tone depends in a complex way on the entire part-tone spectrum. In general, the pitch of a musical tone cannot be precisely determined by the fundamental frequency (or oscillation frequency) alone; rather, the frequencies and amplitudes of many partials play a significant role as well (cf. TERHARDT [12]. Therefore the second of the two aforementioned preconditions is not fulfilled with either type of musical tone.

Naturally the relevance and consequences of these arguments depend on the magnitude of the respective effects. Table 1 presents typical magnitudes of some numerical, physical, and auditory effects pertinent to intonation. The figures indicate that in fact the "mistuning" introduced by inharmonicity of piano strings (i.e. a physical effect) and by pitch shifts (i.e. departures of pitch from supposed nominal values) are at least of the same order of magnitude as classical numerical intonation effects such as for instance the pythagorean comma. It is thus apparent that the classical method of describing intonation just by "tone frequency" can only roughly account for perceptually relevant intonation criteria.

## 4. Conclusions

Whatever objections may seem justified against the details of the present approach to tone scale development and intonation, one conclusion appears to be quite safe: There is not even theoretically such a thing as an ideal fixed intonation which can be described by "tone frequencies" without making further specifications. Optimal intonation in every case and instant is a compromise dependent on partly contradictory criteria. Optimal intonation can neither be regarded as fixed nor can it be sufficiently specified by just one frequency per tone. Optimal intonation of tonal music must be flexible, i.e. adapting to momentary requirements. Any rigid assignment of frequencies to tones can thus serve only as an abstract reference pattern from which the optimal intonation in every instant will depart more or less distinctly. What type of intonation (i.e. just, pythagorean, or equally tempered) is used as a reference pattern, is of secondary importance, though equally tempered intonation appears to be most convenient for that purpose.

The dualistic concept of harmonic and sensory purity sketched in the present study may provide a systematic solution to the intonation problem. To take full advantage of that approach a theory of pitch perception to evaluate harmonic purity, and of beat- and roughness-perception to evaluate sensory purity are required. The pitch theory must in particular account for pitch shift effects such as those mentioned in Table 1. The virtual-pitch theory (TERHARDT [9], TERHARDT [12]) meets these criteria to a considerable extent. Since perception of beats and roughness is also well understood, it will appear that, in spite of the complexity of effects involved, we are beginning to understand musical tone scale intonation and likewise we are beginning to be able to predict quantitatively the physical sound parameters yielding optimal intonation.

**Acknowledgments**

The author gratefully acknowledges many helpful comments on the manuscript given by Dr A. JAROSZEWSKI and Dr I. LINDEVALD. This study was worked out in the Sonderforschungsbereich 204 Gehör, München, supported by the Deutsche Forschungsgemeinschaft.

**References**

[1] E. M. BURNS and W. D. WARD, *Intervals, scales, and tuning*, In: D. Deutsch [ed], The Psychology of Music, Academic Press, New York 1982.

[2] W. J. DOWLING, *Melodic information processing and its development*, In: D. Deutsch [ed], The Psychology of Music, Academic Press, New York 1982.

[3] J. C. ECCLES, *The understanding of the brain*, New York, McGraw-Hill, 1973.

[4] H. HELMHOLTZ, *Die Lehre von den Tonempfindungen*, Braunschweig, Vieweg 1963.

[5] S. MAKEIG, *Octave tuning — two modes of perception*, Res. Sympos. on the Psychol. and Acoustic of Music, Lawrence, Kansas 1982.

[6] K. POPPER, *Objective knowledge*, University Press, Oxford 1972.

[7] O. H. SCHUCK, and R. W. YOUNG, *Observations on the vibrations of piano strings*, J. Acoust. Soc. Am., **15**, 1-11 (1943).

[8] R. SHUTER-DYSON, *Musical ability*, In: D. Deutsch [ed], The Psychology of Music, Academic Press, New York 1982.

[9] E. TERHARDT, *Zur Tönhöhenwahrnehmung von Klängen*, Acustica, **26**, 173-199 (1972).

[10] E. TERHARDT, *Ein psychoakustisch begündetes Konzept der Musikalischen Konsonanz*, Acustica, **36**, 121–137 (1976).

[11] E. TERHARDT, *The concept of musical consonance, a link between music and psychoccoustics*, Music Perception, **1**, 276–295 (1984).

[12] E. TERHARDT, G. STOLL and M. SEEWANN, *Algorithm for extraction of pitch and pitch salience from complex tonal signals*, J. Acoust. Soc. Am., **71**, 679-688 (1982).

[13] K. WALLISER, *Über die Spreizung von empfundenen Intervallen gegenüber mathematisch harmonischen Intervallen bei Sinustönen*, Frequenz, **23**, 139–143 (1969).

[14] W. D. WARD, *Subjective musical pitch*, J. Acoust. Soc. Am., **26**, 369-380 (1954).

*Received August 21, 1987.*