Music Perception
Fall 1988, Vol. 6, No. 1, 65–94

Revision of Terhardt's Psychoacoustical Model of the Root(s) of a Musical Chord

RICHARD PARNCUTT
University of New England, Australia

The predictions of Terhardt's octave-generalized model of the root of a musical chord occasionally disagree with music theory (notably, in the case of the minor triad). The model is improved by assigning appropriate weights to the intervals used in the model's "subharmonic matching" routine. These intervals, called "root-supports," include the P8 (unison), P5, M3, m7, M9 (M2), and m3. The new model calculates the salience of each pitch class (C, C♯/D♭. . .B) as an absolute value. The most likely candidate for the root of a chord corresponds to the most salient pitch class in all cases where the root is unambiguously defined in music theory. The model also calculates a "root ambiguity" value for each chord, a measure of its dissonance. Effects of voicing (inversion, spacing, and doubling) and context on the root are considered.

Introduction

The root of a chord may be defined most simply as the note or pitch after which the chord is named in Western music terminology and theory. According to this definition, the root of C–E–G is C simply because of the name "C major." In a deeper sense, the root is that single note which most satisfactorily represents the function of a chord in a harmonic progression. If the chord C–E–G in a chord sequence were to be replaced by a single note, the note which would least disturb the harmonic progression would normally be C.

The concept of the root of a musical chord is essential for the understanding of Western harmony. Yet despite centuries of music-theoretical development, there still exists no widely accepted theory of the root's nature and origins.

Rameau (1750) was the first to make significant progress in this area. He compared the notes of a musical chord with the harmonics of a single complex tone (such as typical musical tone) and hypothesized a correspondence between the root of a chord and the fundamental of a harmonic series of frequencies. Rameau's work was revived during the nineteenth century, for

Requests for reprints may be sent to Richard Parncutt, Departments of Music and Psychology, University of New England, Armidale NSW 2351, Australia.

example, by Hauptmann in Germany and Sechter in Vienna (see Watson, 1982). It was regarded as the basis for a "vertical" view of harmony, complementing the "horizontal" view based on the diatonic scale and figured bass (e.g., as maintained by Sechter). Some later theorists (e.g., Bruckner) tended to place more emphasis on the "vertical" aspect of harmony, asserting that not only triads and sevenths but also ninths were explicable in terms of Rameau's theory. Others (e.g., Schenker) emphasized the "horizontal" aspect, claiming that only triads were derived from the harmonic series ("Nature") while sevenths and ninths were produced by melodic or horizontal elaboration of triads (by the "Artist").

Rameauian music theory has always been plagued by its failure to explain satisfactorily the nature and root of the second most common chord in mainstream Western music, the minor triad. The triad can be found among the harmonics of a complex tone: the sixth, seventh, and ninth harmonics, or the tenth, twelfth, fifteenth. However neither the sixth nor the tenth harmonic is octave-equivalent to the fundamental; so the harmonic series method fails to predict the triad's music-theoretical root. Riemann (1893) tried instead to explain the minor triad by a theory of "undertones," in contrast to the overtones by which Rameau had accounted for the major triad. This led, however, to the counterintuitive prediction that the root of the minor triad corresponds to its fifth.

Hindemith (1940/1942) preferred to explain the root in terms of combination tones—"distortion" tones produced in the inner ear by nonlinear interaction between tone components. However, his analysis made much the same predictions about the roots of chords as analyses based on Rameau's approach. In any case, combination tones were later shown to be inaudible in ordinary music (see Plomp, 1965) and therefore unlikely to influence the perception of harmony.

Along with a number of other music theorists, Hindemith tried to escape from the theoretical dilemma surrounding the minor triad by regarding it as a distorted form of the major triad. This explanation might be acceptable if the minor triad were a relatively unusual or unimportant chord. On the contrary, however, the minor triad is almost as fundamental as the major triad in determining the pitch structure of mainstream Western music. This suggests that the minor triad should be regarded as a chord in its own right, and that any explanation of its nature and origin, to be plausible, should be consistent with this view.

More recently, Terhardt (1974) explained the perceptual nature of the root of a chord in terms of the psychoacoustics of tone perception (see below) and, on the basis of this theory, devised a general model by which the roots of common musical chords may be predicted (Terhardt, 1982). While his concept of the root provides considerable insight into the nature of harmony, his chord-root model still fails to predict the music-theoretical root of the minor triad (see below). The model nevertheless provides the most satisfactory account to date of the nature of musical chords and the determination of their roots. This article aims to resolve problems with Terhardt's chord-root model—in particular, those related to the minor triad—and to present a revised model that is suitable for application in music theory, analysis and composition.

Conventional Music Theory

The conventional theory of the roots of musical chords is generally understood to be *octave-generalized*. In such theory, the root of a chord is regarded as a *pitch class*, that is, a pitch whose octave register is not specified.[1] Similarly, the root may be thought to be a function of a *chord class* (a simultaneity of pitch classes) and so to be independent of voicing (inversion, spacing, and doubling).

To define the root of a chord in this way may be convenient from the point of view of music theory and terminology, but it does not always accurately reflect musical reality. Sometimes, the root of a chord does depend on its voicing. For example, the chords A minor seventh (A–C–E–G) and C major sixth (C–E–G–A) are normally thought to have different roots (A and C, respectively), even though they belong to the same chord class (as defined above). The root can also depend on context: for example, the diminished seventh chord B–D–F–A♭ (in its various enharmonic spellings) may function as a dominant ninth chord with missing root in the minor keys of C, A, F♯, and E♭, that is, its root may be G, E, C♯ or B♭. Because effects of voicing and context on the root are considerably weaker than the effect of choice of pitch classes, their detailed discussion is postponed until later in this article.

According to conventional music theory, chord classes are constructed by superposing thirds. The lowest note of the resultant construct is the conventional root of the chord. Two superposed thirds produce a triad, the basic structural unit of diatonic music. Additional thirds produce seventh, ninth, eleventh,and thirteenth chords. Superposed thirds may be major or minor, so that various chromatic alterations are possible for each of the higher notes in the chord.

The conventional theory of stacked thirds provides a convenient system of chord terminology that is compatible with diatonically based, conventional music notation. However the theory has some serious shortcomings. To start with, it suggests that there is something quite fundamental about

1. The term "pitch class" was introduced in the context of the theory of atonal music by Babbitt (1955). The meaning of the term is modified somewhat in the present approach, which is mainly concerned with tonal music.

major and minor thirds for the construction of musical chords. On the contrary, the most important harmonic interval (apart from the octave) is clearly the fifth, not the third: the importance of the major and minor triads, for example, is due more to the harmonic properties of the perfect fifth interval than to the way in which major and minor thirds may be superposed to form triads. Another shortcoming of the model concerns the so-called eleventh and thirteenth chords. It is arguable whether the model reflects the musical properties of such chords in an appropriate way. For example, it is often necessary to omit the third from eleventh chords, and the fifth from thirteenths, suggesting that such chords might more appropriately be regarded as added (or suspended) fourths and sixths (respectively).

This article aims to replace (or, at least, to complement) the conventional theory of stacked thirds by a more general and less arbitrary theory of chords and their roots, based on the psychoacoustics of the perception of individual complex tones.

Terhardt's Theory of Pitch and Harmony

The theory of stacked thirds fails to address the basic issue of how chords are *perceived*. The issue of perception is acknowledged in the work of Rameau (who explained musical chords in terms of the physical structure of individual complex tones) and Schenker (who regarded chords as artistic limitations and extensions of Nature), but neither of these authors tackled the issue in a thorough or systematic way. In particular, they were not in a position to explain why the root of a chord is related to the harmonic structure of a single tone even though composers, performers, and listeners are seldom directly aware this structure.

For Terhardt, neither the (musical) structure of diatonic scales nor the (physical) structure of complex tones has any *direct* bearing on the nature of musical chords. He hypothesized an indirect connection, in which the "missing link" is an unconscious and universal process of *conditioning* of the auditory system by repeated exposure to complex tones (especially, speech vowels). By studying this conditioning effect, Terhardt was able to turn Rameau's intuitively based theory into science—into a theory based on comparisons between experimental data and the predictions of an objective model for the pitch of any steady-state complex sound (Terhardt, Stoll, & Seewann, 1982).

According to Terhardt's theory, the pitch of a complex tone is determined by a spontaneous process of *pattern recognition*. The audible harmonics of a complex-tone describe a specific pitch pattern: the interval between the first and second harmonics corresponds to the musical interval of an octave, between the second and third to a fifth, and so on. This "harmonic pitch pattern" is very common in ordinary sounds, and is especially common in speech. It is therefore highly familiar to the auditory system—so familiar that the listener is unaware of ever "recognizing" the pattern. Instead, the listener is (normally) aware of a single tone sensation whose pitch corresponds to that of the fundamental (the first harmonic).

A tone sensation is normally experienced at the fundamental if part of the harmonic pitch pattern is missing, even if this includes the fundamental itself. It is quite common for the harmonic pitch pattern to be incomplete due to masking: a harmonic or group of harmonics may be masked ("drowned out") by neighboring harmonics or by other, simultaneous sounds. Alternatively, certain harmonics may be absent from the physical stimulus in the first place. The fundamental, for example, is normally physically absent from the speech vowels reproduced by the earpiece of a telephone or the loudspeaker of a small radio. The pitch of such "residue tones" usually corresponds to the missing fundamental frequency.

Sounds in the environment and in music often include simultaneous complex tones. The resultant sound spectrum is made up of many pure tone components, each of which could be a harmonic of a number of different possible complex tones. In the process of identifying environmental sound sources, the auditory system needs to "decide" somehow which pure tone component belongs to which complex tone. This process is simulated mathematically in the pitch model of Terhardt et al. (1982). The first part of the model calculates the pitches and audibilities of the (pure) tone components of an input sound. On the basis of this information,the second part models the identification of harmonic sound sources (complex tone components). It does this by a procedure called "subharmonic matching." Each audible pure tone component is assigned a number of "subharmonic pitches": for example, a component at A440, or A4 (according to the system for numbering octave registers adopted by the American Standards Association, 1960) has subharmonics at A3, D3, A2, F2, D2, B1, A1, G1 etc. Matches between subharmonics of different pure tone components suggest that they are both harmonics of the same tone—the tone whose fundamental lies at the point of coincidence. For example, the third subharmonic of a pure tone component at A4 lies at D3. So does the second subharmonic of a component at D4. This suggests that the two components A4 and D4 are the third and second harmonics (respectively) of a complex tone whose fundamental pitch is D3. The more subharmonics coincide at a particular pitch, the more likely it is that the pitch represents the fundamental of a complex tone, and hence indicates the environmental presence of a harmonic (or periodic) sound source.

Terhardt's Chord-Root Model

A possible way of using the pitch model of Terhardt et al. (1982) to determine the root of a chord might be to input the chord to the model, complete with all the harmonics of each of its notes. But this would be inappropriate for ordinary music theory. First, it would be too complicated: a long computer program would be needed to make the necessary calculations. Second, results would depend to a considerable extent on the voicing (inversion, spacing, doubling) of the notes of the chord. As discussed above, this is not normally the case in music theory.

Terhardt's (1982) chord-root model avoids these problems by neglecting information on octave register altogether. The input to the model consists of the pitch classes of the notes of the chord. The harmonics of each tone (other than, effectively, those which are octave-equivalent to the fundamental) are not considered. The output of the model consists of the pitch classes of various candidates for the chord's root.

The model is basically an octave-generalized version of the subharmonic matching procedure of the pitch model of Terhardt et al. (1982). It combines the psychoacoustical principle of harmonic pitch pattern recognition with the music-theoretical principle of octave equivalence.[2]

The subharmonics considered in the model are limited to the first ten. The eleventh is hard to categorize in the chromatic scale (it lies roughly midway between a fourth and a tritone) and only rarely influences tone perception anyway, due to masking by neighboring harmonics. Within the confines of this model, then, it is inappropriate to speak of the minor triad as a chord whose tones lie in a frequency ratio of (approximately) 10:12:15, as the tenth, twelfth, and fifteenth harmonics are seldom audible in isolated complex tones. Instead, the minor triad is regarded simply as a particular configuration of the intervals P5, M3, and m3.

The seventh harmonic of a complex tone is noticeably mistuned relative to equal temperament: it is about a third of a semitone flatter than an equally tempered minor seventh interval above the fourth harmonic. This does not affect the model, as individual harmonics in a complex tone may be mistuned by as much as half a semitone (3% in frequency) and still be perceived as "belonging" to the tone (Moore, Peters, & Glasberg, 1985). Similarly, considerable variations in tuning may occur during performance of diatonic music (see, e.g., Fransson, Sundberg, & Tjernlund, 1974) with-

2. Regarding octave equivalence: Terhardt's theory implies that notes an octave apart are perceived as strongly related or similar, but not as "equivalent" in the music-theoretical sense. The psychoacoustical property of similarity appears to have been strengthened or exaggerated to the point of "equivalence" over a long period of music-historical development.

out affecting the music's perceived tonal structure. Such variations are easier to explain if musical intervals are regarded as no more than *pitch distances* to which the auditory system is sensitive (as in Terhardt's theory) rather than as frequency ratios (e.g., as in the theory of Boomsliter & Creel, 1961).

Octave equivalence is accounted for in the model by transposing the first ten elements of the harmonic series into a single octave. This reduces the number of different intervals above the fundamental to five: the octave or unison (from harmonics 1, 2, 4, and 8); the fifth (from 3 and 6); the major third (from 5 and 10); the minor seventh (from 7); and the major ninth or second (from 9). These "root-support" intervals (as they were called by Stuckey, 1983) may be written in abbreviated form as P1, P5, M3, m7 and M2, where P stands for "perfect," M for "major," and m for "minor."

The model predicts the root(s) of a specific chord in the following way. First, the subharmonic pitch classes of each note are listed, by transposing each note down through each of the root-support intervals in turn. For example, the subharmonic pitch classes of C are listed as C, F, A♭, D, and B♭. Next, matches are sought among the subharmonics. In the case of the major triad C–E–G, C is found to be a subharmonic of all three notes. In other words, C is "supported" by all three notes of the chord. It is therefore a more likely candidate for the root of the chord than F, D, and A, each of which is supported by only two of the three notes of the chord.

The predictions of the model in the case of the major triad are in accord with music theory and practice. C is almost always the functional root of the chord. The notes F, D, and A may each be added to the bass of the triad to produce the (root-position) chords F major ninth (with missing third), D (dominant) eleventh, and A minor seventh. The notes F, D, and A are also the most commonly occurring melodic notes associated with the C major triad in music. In the key of C major, the most stable pitches (besides the notes C, E, and G of the tonic triad) are the notes F, D, and A. The leading note B—like the more common chromatic notes in C major (e.g., F♯ and A♭)—is weaker and less stable than the other diatonic notes: in conventional theory, it cannot normally be doubled, and it normally resolves through a semitone step to the nearest diatonic note.

That Terhardt's (1982) chord-root model makes musically feasible predictions for most other important chord classes may readily be verified. An important chord whose root is *not* accounted for satisfactorily by the model—as mentioned in the introduction—is the minor triad. According to the music-theoretical method of stacking thirds, the root of the triad C–E♭–G is C. The note C is also the most common root of this chord in music. The second most common root of C–E♭–G in music would appear to be E♭: the first inversion, E♭–G–C, can function as an E♭ major sixth

chord with a missing fifth. Yet Terhardt's chord-root model predicts that F is the most likely candidate for the root of the C minor triad (Table 1a), as the tones C, E♭, and G all correspond to harmonics of F. This prediction is not entirely in error: F *can* act as the root of this triad in certain contexts. But F could not by any stretch of the imagination be regarded as "the" root of the C minor triad.

Another chord whose root is not well predicted by Terhardt's (1982) chord-root model is the half-diminished seventh, C–E♭–G♭–B♭ (Table 1b). This chord has at least three commonly occurring roots, depending on the context in which it appears. The root may be C, for example when the chord is included in a diatonic sequence of sevenths in D♭ major or B♭ minor. The root may be E♭, for example when the first inversion of the chord (an E♭ minor sixth chord) precedes perfect cadences in B♭ minor, or when the chord acts as a tonic in the key of E♭ (Dorian) minor (e.g., in jazz). It may be an A♭, when the chord (in any inversion) appears in the key of D♭ major as a "dominant ninth with missing root."

TABLE 1
Terhardt's (1982) Chord-Root Model

a.	Minor Triad C–E♭–G											
Root Candidate	C	D♭	D	E♭	E	F	G♭	G	A♭	A	B♭	C♭
P1	C	—	—	E♭	—	—	—	G	—	—	—	—
P5	G	—	—	—	—	C	—	—	E♭	—	—	—
M3	—	—	—	G	—	—	—	—	C	—	—	E♭
m7	—	—	C	—	—	E♭	—	—	—	G	—	—
M2	—	E♭	—	—	—	G	—	—	—	—	C	—
Number of Supports	2	1	1	2	0	3	0	1	2	1	1	1

b.	Half-Diminished Seventh Chord C–E♭–G♭–B♭											
Root Candidate	C	D♭	D	E♭	F♭	F	G♭	G	A♭	A	B♭	C♭
P1	C	—	—	E♭	—	—	G♭	—	—	—	B♭	—
P5	—	—	—	B♭	—	C	—	—	E♭	—	—	G♭
M3	—	—	G♭	—	—	—	B♭	—	C	—	—	E♭
m7	B♭	—	C	—	—	E♭	—	—	G♭	—	—	—
M2	—	E♭	—	—	G♭	—	—	—	B♭	—	C	—
Number of Supports	2	1	2	2	1	2	2	0	4	0	2	2

NOTE: Abbreviations for root-support intervals: P1 = perfect unison (or octave); P5 = perfect fifth (or twelfth); M3 = major third (or tenth); m7 = minor seventh (or fourteenth); M2 = major second (or ninth).

The ambiguity of the root of the half-diminished seventh chord is reflected neither by conventional musical theory nor by Terhardt's chord-root model (Table 1b). According to the stacked thirds method, the root of C–E♭–G♭–B♭ is C; the method, as it is normally interpreted, does not implicate A♭ and E♭ as further possible roots. Terhardt's model, on the other hand, clearly predicts that the root of the chord is A♭: in emphasizing the similarity between the chord and a "residue tone", it forces the analogy between the chord's root and the missing fundamamental of such a tone. E♭ and C are listed by Terhardt's model at the same level as musically relatively unlikely candidates such as D, F, G♭, B♭, and C♭ (or B).

Weighting of Subharmonics

Terhardt's (1982) chord-root model correctly predicts which pitches in a chord *could* act as roots, but fails to weight these root candidates appropriately relative to each other. For example, the pitch class F *can* act as the root of the chord C–E♭–G, and A♭ *can* be the root of C–E♭–G♭–B♭, but the importance of these possibilities in music theory and practice is not as great as implied by the model. This appears to be a consequence of Terhardt's oversimplification of his pitch model to make it easier to apply in music theory. The model could be improved by reintroducing some of the complexity of the pitch model, in the form of a more sophisticated method of weighting predicted roots relative to each other.[3]

In Terhardt's chord-root model, each subharmonic effectively receives equal weight (Table 2, row 1). In the subharmonic matching procedure of the pitch model of Terhardt et al. (1982), however, subharmonics are weighted such that those with higher harmonic numbers carry less weight: the *n*th subharmonic is assigned the weight $1/n$. This is because higher harmonics tend to be less audible, and therefore less likely to influence complex tone perception, than lower harmonics. The large difference between the weights assigned to the first and second subharmonics (1 and 1/2 respectively) reflects the large difference observed between the pitch salience (and pitch ambiguity) of full complex tones and residue tones, that is, between complex tones with and without tone components at their fundamentals (Stoll, 1983).

3. Robert Stuckey suggested the possibility of weighting the root support intervals in this way in a letter to Ernst Terhardt dated 22 May, 1984. Specifically, he recommended assigning the weights 35, 17.5, 12, 7, and 5 to the intervals P1, P5, M3, m7, and M9 (respectively), on the basis of an analysis of the average audibilities of the harmonics in speech vowels.

TABLE 2
Derivation of Weights of Subharmonics

	Subharmonic Interval Class				
Weight	P1	P5	M3	m7	M2
1. After Terhardt (1982)	1	1	1	1	1
2. After Terhardt et al. (1982)	15/8	1/2	3/10	1/7	1/9
3. Row 2 divided by 15/8	1.00	0.27	0.16	0.08	0.06
4. Row 3 raised to the power 0.55	1.00	0.48	0.36	0.24	0.21
5. Easy approximation to Row 4.	1	1/2	1/3	1/4	1/5

The formulation for the weight of a subharmonic in the pitch model of Terhardt et al. (1982) may be used to estimate appropriate weighting factors for the "octave-generalized subharmonics" (root supports) of his chord-root model, simply by adding contributions from actual subharmonics which are octave-equivalent (Table 2, row 2). The first, second, fourth, and eighth subharmonics are octave-equivalent, and have weights of 1, 1/2, 1/4, and 1/8 in the pitch model. So the weight of the octave/unison in the octave-generalized chord-root model may be estimated at $1 + 1/2 + 1/4 + 1/8 = 15/8$. Similarly, the weight of the P5 may be estimated at $1/3 + 1/6 = 1/2$; of the M3, at $1/5 + 1/10 = 3/10$. Only one audible harmonic contributes to each of the weights assigned to the m7 and M2.

For ease of comparison, all five values are normalized by dividing through by the value for the octave/unison (row 3 of the table). The general trend of the resultant weights agrees with musical experience. For example, the root-supporting effect of the fifth (P5) is greater than that of the major third (M3). However the trend appears to be too steep: the value assigned to M2 is very small compared to that assigned to P1. An intuitively more reasonable set of values is obtained by raising the numbers in row three to a power less than one. The arbitrary value of 0.55 is chosen for the exponent, as it leads to a simple and easily memorable approximation (row 5). These final values are used throughout the rest of this article. When used to predict the roots of musical chords (below), they are found to produce results that are consistently in accord with musical experience.[4]

The five listed intervals support the root directly: their fundamentals are octave-equivalent to harmonics of the root. It is also possible for a note to support the root indirectly, via selected overtones (rather than its fundamental). Consider, for example, a note a minor third (m3) above the root.

4. Future experimental work (e.g., along the lines of Thompson & Cuddy, 1988) may yield more precise vahies for the weights of root support intervals; however, such an improvement is unlikely to affect the *music-theoretical* applications of the model, as the model already operates at a level of precision greater than that normally needed for music theory.

According to Terhardt's theory, the fundamental of this note lends no support to the root, as no minor third occurs between the fundamental and the first 10 octave-generalized harmonics of a complex tone. However its third and fifth harmonics (P5, M3) *can* support the root: they are octave-equivalent to the root's seventh and third harmonics (m7 and P5 respectively). These relationships may be summarized by the expressions m3 + P5 = m7 and m3 + M3 = P5.

The indirect root-supporting properties of the intervals P4, TT, m6, and M6 may be seen similarly, from the (octave-generalized) expressions P4 + P5 = P1, TT + M3 = m7, m6 + M3 = P1, and M6 + P5 = M3. However the support provided to the root by these intervals is negligible, because their root-supporting harmonics normally coincide with harmonics of the root that are already physically present and audible. In the case of the P4, for example, the third harmonic of the upper note coincides with the fourth harmonic of the lower (or, in the case of the P11, with the eighth harmonic of the lower). New pitch information leading to reinforcement of the root can only be introduced by intervals such as the P4, TT, m6, or M6 if one of the harmonics of the lower tone happens to be missing or inaudible. This occurs only rarely (e.g., when the lower tone is played by an instrument of the clarinet family). Even in such cases, the missing harmonics may be assumed already to have been "filled in" by the auditory system, as part of the process recognizing incomplete patterns (see Wertheimer, 1923).

In the case of the m3, the third harmonic of the higher note falls between the third and fourth harmonics of the lower. For example, the third harmonic of E♭4 (B♭5) falls between the third and fourth harmonics of C4 (G5 and C6). In the case of the octave-equivalent interval m10, the root-supporting third harmonic of the upper note coincides with a harmonic of the lower note and so does not lend significant additional root support to the lower note. So the interval class m3 only provides root support to the pitch class of its lower note when it appears in close position, or in inversion (e.g., in the case of a minor triad in first inversion). The root-support weight chosen for the interval m3 in the model is correspondingly small. The value 1/10, or half the weight assigned to the interval M2, appears to be an appropriate value.

Stuckey included the m3 (along with the P1, P5, M3, m7, and M2) among the root-support intervals. He called the other six interval classes "detractors". The main harmonic functions of detractors, according to Stuckey, are (1) to suggest other possible roots of a chord (i.e., to increase the ambiguity of the root) and (2) to give different chords characteristic "colors," in much the same way that nonharmonic tone components add "color" to the sound of a bell. The concept of root support and detractor intervals may be regarded as an alternative to the conventional concept of stacked thirds in music theory.

Weighting of Pitch Classes

The model described in this article may be divided into three stages. First, each pitch class is assigned a "weight": a relative measure of its perceptual importance of salience. Second, the calculated pitch weights of a chord are used to estimate its "root ambiguity." Finally, the calculated root ambiguity value is used to convert pitch weights into absolute estimates of "salience." Calculated weights and saliences in the model are octave-generalized: they may be regarded as averaged over different voicings (inversions, spacings, doublings) of a chord class, within which the relative perceptual importances of different tones belonging to the same pitch class can actually vary quite considerably (see Parncutt, 1987).

This section is concerned with the assignment of weights to pitch classes. The primary aim is to establish the root(s) of a chord class by systematic application of psychoacoustical theory. The root of a chord class is assumed to correspond to the pitch class with the highest calculated weight.

Examples of how pitch class weights are calculated in the model are provided in Table 3. The method is the same as that used in Table 1, except that (1) the subharmonic interval classes P1, P5, M3, m7, and M2 are assigned different weights, as described in the previous section; and (2) the weak root-supporting property of the interval m3 is taken into account. The sum of each column of the table is the "weight" assigned by the model to the pitch class represented by that column.

TABLE 3
Revised Chord-Root Model

a.	Minor Triad C–E♭–G											
Pitch Class	C	D♭	D	E♭	E	F	G♭	G	A♭	A	B♭	C♭
P1	1.00	—	—	1.00	—	—	—	1.00	—	—	—	—
P5	0.50	—	—	—	—	0.50	—	—	0.50	—	—	—
M3	—	—	—	0.33	—	—	—	—	0.33	—	—	0.33
m7	—	—	0.25	—	—	0.25	—	—	—	0.25	—	—
M2	—	0.20	—	—	—	0.20	—	—	—	—	0.20	—
m3	0.10	—	—	—	0.10	—	—	—	—	0.10	—	—
Weight (sum)	1.60	0.20	0.25	1.33	0.10	0.95	0.00	1.00	0.83	0.35	0.20	0.33

b.	Half-Diminished Seventh Chord C–E♭–G♭–B♭											
Pitch Class	C	D♭	D	E♭	F♭	F	G♭	G	A♭	A	B♭	C♭
P1	1.00	—	—	1.00	—	—	1.00	—	—	—	1.00	—
P5	—	—	—	0.50	—	0.50	—	—	0.50	—	—	0.50
M3	—	—	0.33	—	—	—	0.33	—	0.33	—	—	0.33
m7	0.25	—	0.25	—	—	0.25	—	—	0.25	—	—	—
M2	—	0.20	—	—	0.20	—	—	—	0.20	—	0.20	—
m3	0.10	—	—	0.10	—	—	—	0.10	—	0.10	—	—
Weight (sum)	1.35	0.20	0.58	1.60	0.20	0.75	1.33	0.10	1.28	0.10	1.20	0.83

The results of these calculations agree more closely with music theory than the results of the simpler procedure of Table 1. Parts (a) of Tables 1 and 3 both deal with the minor triad C–E♭–G. According to Table 1a, F is the root of this triad. In Table 3a, the root of this triad is predicted to be C, with E♭ in second place; other possible roots are predicted to be G, F, and A♭. In part (b) of the table, the half-diminished seventh chord C–E♭–G♭–B♭ is analyzed. The predicted root here is E♭, with C, G♭, and A♭ competing for second place. These results are more sensible that those of Table 1b, in which A♭ is predicted to be easily the strongest candidate for the root of the half-diminished seventh chord on C.

The calculations of Table 3 take longer to perform than those of Terhardt's (1982) chord-root model in Table 1. The easiest way to realize the new version of Terhardt's model is by computer. The following mathematical formulation of the model provides the basis for a suitable computer program.[5]

At the input to the model, the pitch classes of the notes of a chord class are stored in an array $N(p)$, $p = 0...11$, where N stands for "note" and p stands for "pitch". The value $p = 0$ corresponds to the pitch class C; $p = 1$ corresponds to C♯/D♭; and so on up to $p = 11$, which corresponds to B. N can only take two values: 1 and 0. If $N(p) = 1$, then the pitch class p is included among the notes of the chord; if $N(p) = 0$, pitch class p is not included. According to this system, a D major triad would be represented by setting $N(p) = 1$ for $p = 2$, 6 and 9 (D, F♯, and A), and $N(p) = 0$ for other values of p; 0, 1, 3, 4, 5, 7, 8, 10, and 11.

The subharmonic pitch weights derived in the last section are stored in an array $w(i)$, $i = 0...11$, where w stands for "weight" and i stands for "interval"—more precisely, "octave-generalized interval" or "interval class," in semitones. Referring to Table 2, $w(0) = 1.00$, $w(7) = 0.50$, $w(4) = 0.33$, $w(10) = 0.25$, and $w(2) = 0.20$. The weak root-supporting role of the minor third is accounted for by setting $w(3) = 0.10$. Other elements of the array (corresponding to the "root detractors") are set to zero: $w(i) = 0$ for $i = 1, 5, 6, 8, 9$, and 11.

Contributions to the weight W of the pitch class p are made when notes N lie at root-support intervals i above p, that is, at pitches $p+i$:

$$W(p) = \mathrm{sum}_i\{N(\mathrm{mod}_{12}[p+i]) \times w(i)\}. \qquad (1)$$

In this equation, the operator sum_i denotes summation of the contents of the brackets { } over all values $i = 0...11$. The operator mod_{12} indicates that the expression in brackets [ ] is in "modulo 12".

5. A FORTRAN implementation of the program is available from the author.

Modulo 12 arithmetic may be performed by counting around the twelve points of a clock. In the present application of modulo 12 arithmetic, each point of the clock may be thought of as a pitch class: 12 o'clock is C, 7 o'clock is G, etc. Mod_{12} $[p+i]$ is the same as $p+i$ *if* $p+i$ lies in the range 0...11; otherwise, a whole multiple of 12 is added or subtracted such that the result lies in this range.

Root Ambiguity

In tonal music, the root of a major triad or major–minor (e.g., dominant) seventh chord is quite unambiguous, while that of a half-diminished or diminished seventh chord can be quite ambiguous. There is a general tendency for more dissonant chords to have more ambiguous roots. The present section describes how the root ambiguity of a chord may be estimated quantitatively. By means of a model, root ambiguity values are calculated that are relatively low for consonant chords like the major triad and relatively high for dissonant chords like the diminished seventh chord.

Root ambiguity may be expressed in terms of the number of different possible roots of a chord, where different possibilities are weighted according to their perceptual importance or probability of occurrence in typical musical contexts. As an example, consider a (hypothetical) chord which has just two possible roots, of which one is relatively strong and the other relatively weak. According to the above definition, the root ambiguity of such a chord would be obtained by adding a number close to one for the first root to a number considerably less than one for the second. The root ambiguity of the chord would be greater if this second root were stronger, or if the chord had a third possible root.

The weights calculated in the previous section for the pitch classes of a chord provide an appropriate basis for calculating a chord's root ambiguity. Root ambiguity is low for chords whose distribution of pitch-class-weights is quite "peaked," namely, for chords whose maximum weight is considerably greater than other weights in the distribution. The following equation expresses this idea as simply as possible (i.e., by means of the simplest arithmetical operations) but without sacrificing generality (i.e., taking into account all pitch classes):

$$A' = \mathrm{sum}_p\{W(p)/W_{\max}\}, \qquad (2)$$

where A' denotes a preliminary estimate of root ambiguity, and $W_{\max}$ denotes the maximum weight in the distribution—the weight assigned to the predicted root.

When this formulation is applied to calculations made using equation (1) above, the root ambiguity A' of a major triad is predicted to be around 4, while that of a diminished seventh chord is around 8. According to the concept of root ambiguity described here, these figures are too high. The major triad has one main root, and a number of relatively weak, subsidiary roots: one would therefore expect its root ambiguity to lie somewhere in the vicinity of 2. Judging from the various ways of notating and theorizing about the diminished seventh chord, it has possible roots not only at, but also a semitone below, each of its four pitch classes.[6] As all eight possible roots of the diminished seventh chord are quite weak, the root ambiguity of this chord should be much less than 8.

The simplest way of reducing calculated root ambiguity values to intuitively more reasonable values, but without allowing them to fall below a minimum possible value of 1, is to take the square root of the above expression (eq. 2), that is, to raise it to the (arbitrary) power 0.5 (as in the model of "sensory multiplicity" developed by Parncutt, 1987):

$$A = [\mathrm{sum}_p\{W(p)/W_{\max}\}]^{0.5}. \qquad (3)$$

This formulation produces reasonable root ambiguity values, as set out for common dyads and chords in Table 4. The root ambiguity of each of the dyads is near two. It exceeds two for most triads and tetrads (seventh chords), but is less than the number of tones, reflecting the perceptual "fusion" of tones in musical chords (see Stumpf, 1898).

The root ambiguity values presented in Table 4 may be interpreted as estimates of dissonance. According to this interpretation, the calculated rank order of dyads, from most consonant to most dissonant, is P4/P5, M3/m6, M2/m7, m3/M6, and TT/m2/M7. The rank order of triads is major, minor, augmented, and diminished. The rank order of seventh chords is major–minor (dominant), minor, major, half-diminished, and diminished.

TABLE 4

Calculated Root Ambiguity Values A

(a) Dyads	m2 2.2	M2 2.0	m3 2.1	M3 1.9	P4 1.8	TT 2.2
(b) Triads	Major 2.0	Minor 2.1	Augmented 2.3	Diminished 2.5		
(c) Seventh Chords	Maj/min(dom) 2.1	Minor 2.3	Major 2.3	Half-dim 2.4	Dim 2.9	

6. Those root candidates not corresponding to actual notes, if added to the diminished seventh chord, would turn it into a dominant minor ninth chord.

These ranks agree with expectations based on music theory and practice (e.g., as developed by Malmberg, 1918) with the exception of the ordering of the dyads M2/m7 and m3/M6. In music, the dyads M2 and m7 are normally considered to be more dissonant that the m3 and M6, because more beating (interference between nearby tone components) is associated with the M2 and m7 than with the m3 and M6 (Helmholtz, 1863/1954). The present model predicts that the dyads M2 and m7 are more consonant than m3 and M6, because the interval class M2/m7 more strongly suggests a root than m3/M6. Beating and root ambiguity may be regarded as separate psychoacoustical contributors to the dissonance of music chords (Terhardt, 1977). In this case, the dissonant effect of beating appears to be greater (more salient) than the dissonant effect of root ambiguity.

The values in the table mirror to some extent the historical introduction of new chords into the vocabulary of Western music. For example, in the Middle Ages, fifths were generally regarded as consonances, and thirds as dissonances. In the Renaissance, major and minor triads were regarded sonorities in their own right, while diminished and augmented triads were not. In the early Baroque, major–minor (or dominant) seventh chords were introduced as independent sonorities earlier other tetrads (such as major and minor seventh chords).[7]

Salience of Pitch Classes

The salience of a pitch class may be quantitatively defined as its probability of being noticed or experienced (see Parncutt, 1987). In this way, salience may be expressed as an absolute value (by contrast to the weights calculated above, which are relative). For the purposes of the present, simple psychoacoustical model, individual differences in pitch perception are neglected. Probabilities of noticing are assumed to apply to an idealized, average, Western nonmusician: a listener who is familiar with the sound of diatonic music, but is incapable of recognizing and naming musical elements such as intervals, chords, and chord progressions.

The salience of a pitch class in a chord may be formulated mathematically on the basis of the following two assumptions. First, the definitions of the weight W of a pitch class p (above) and its salience S imply that S is proportional to W:

7. The psychoacoustical model developed in this article says nothing about the *process* by which new, more dissonant sounds were introduced. In general, this involved contrapuntal experimentation: for example, minor seventh intervals heard above major triads in contrapuntal textures (e.g., in Monteverdi) gradually took on the character of independent sonorities (major-minor seventh chords). Another factor influencing the historical development of harmonic vocabulary (not considered here) is the effect of roughness (see the section entitled "Construction of Chords").

$$S(p) = k \times W(p), \tag{4}$$

where k is some positive constant.

Second, the sum of the saliences of all 12 pitch classes in a chord may be assumed to equal the average number of pitch classes simultaneously noticed in the chord (see the experiment on "sensory multiplicity" in Parncutt, 1987). This, in turn, may be assumed to be the same as the sound's root ambiguity A. In this interpretation of the parameter A, the model predicts that roughly two pitch classes are normally noticed simultaneously in major triads, and roughly three in diminished seventh chords (see Table 4). In general, the model predicts that fewer tones are noticed simultaneously in consonant than dissonant chords, as consonant chords blend more easily.[8] This second assumption about the salience of a pitch class may be expressed mathematically as follows:

$$\text{sum}_p\{S(p)\} = A. \tag{5}$$

The following simple formulation of the salience of a pitch class is consistent with the above two assumptions:

$$S(p) = W(p) \times A/\text{sum}_p\{W(p)\}. \tag{6}$$

This equation may be simplified by substituting an expression derived from equation (3) above:

$$\text{sum}_p\{W(p)\} = A^2 \times W_{\max}. \tag{7}$$

The expression now reads:

$$S(p) = [1/A] \times [W(p)/W_{\max}]. \tag{8}$$

Since it is always true that $1 \leq A$ and $W(p) \leq W_{\max}$, salience values $S(p)$ according to this equation always lie in the range 0 to 1. This is consistent with their interpretation as approximate probabilities.

The salience of the predicted root of a chord according to the model may be found by replacing $S(p)$ by $S_{\max}$ and $W(p)$ by $W_{\max}$ in equation (8). This yields simply

$$S_{\max} = 1/A. \tag{9}$$

8. Note that the model does not specify exactly which tones are noticed in a particular case; instead, it estimates the *probabilities* of particular tones being noticed.

According to the model, then, the salience $S_{\max}$ of the root of a chord is inversely proportional to the chord's root ambiguity. $S_{\max}$ may be regarded as a measure of the (psychoacoustical) "tonalness" of the chord (Aures, 1984), or the extent to which the chord sounds like a single tone (see Parncutt, 1987).

Examples

Examples of calculated salience distributions of dyads, triads, and tetrads (seventh chords) are presented in Figures 1 to 3 in the form of histograms. The maximum of each histogram corresponds to "the" (or "the most likely") root of each simultaneity, as predicted by the model. The height of the maximum ($S_{\max}$) is a measure of the "tonalness" (or consonance) of the sound. The figures show that, on average, the calculated tonalness of a simultaneity decreases as the number of simultaneous tones increases. In other words, chords containing more tones sound less like a single tone.

Figure 1 provides information on the roots of dyads (simultaneities of two musical tones). Only dyads spanning intervals up to and including the tritone TT are included, as larger intervals (up to the octave) are equivalent to their inversions within the octave in an octave-generalized analysis. The dyad with the most clearly defined root is the perfect fourth P4, the inversion of P5: its root is the upper note of P4 or the lower note of P5. The root of M3 (m6) is also quite clear: it is the lower note of M3. The roots of M2 (m7) and m3 (M6) are more ambiguous. However the predominant root candidates as predicted by the model (upper note of M2, lower note of m3) are in accordance with musical experience: presented in isolation, the M2 dyad tends to be interpreted as part of a dominant seventh chord, and the m3 is normally heard as part of a minor triad. The predicted roots of the m2 (M7) and TT (A4/d5) are quite ambiguous, as expected from musical experience: played out of context, these dyads do not suggest a tonal center.

In Figure 2, the major and minor triads are predicted to have clearly defined roots, corresponding to their roots in music theory. The predicted roots of the augmented and diminished triads are more ambiguous. The salience distribution of the augmented triad suggests that the triad is compatible with whole-tone scale patterns. The main root candidate of the diminished triad (apart from the notes of the triad themselves) lies a major third below the conventional root. This explains why the diminished triad on the leading note in diatonic music normally functions as an incomplete dominant seventh chord.

The predicted roots of the tetrads in Figure 3 do not always correspond to their conventional roots as determined by the theory of super-

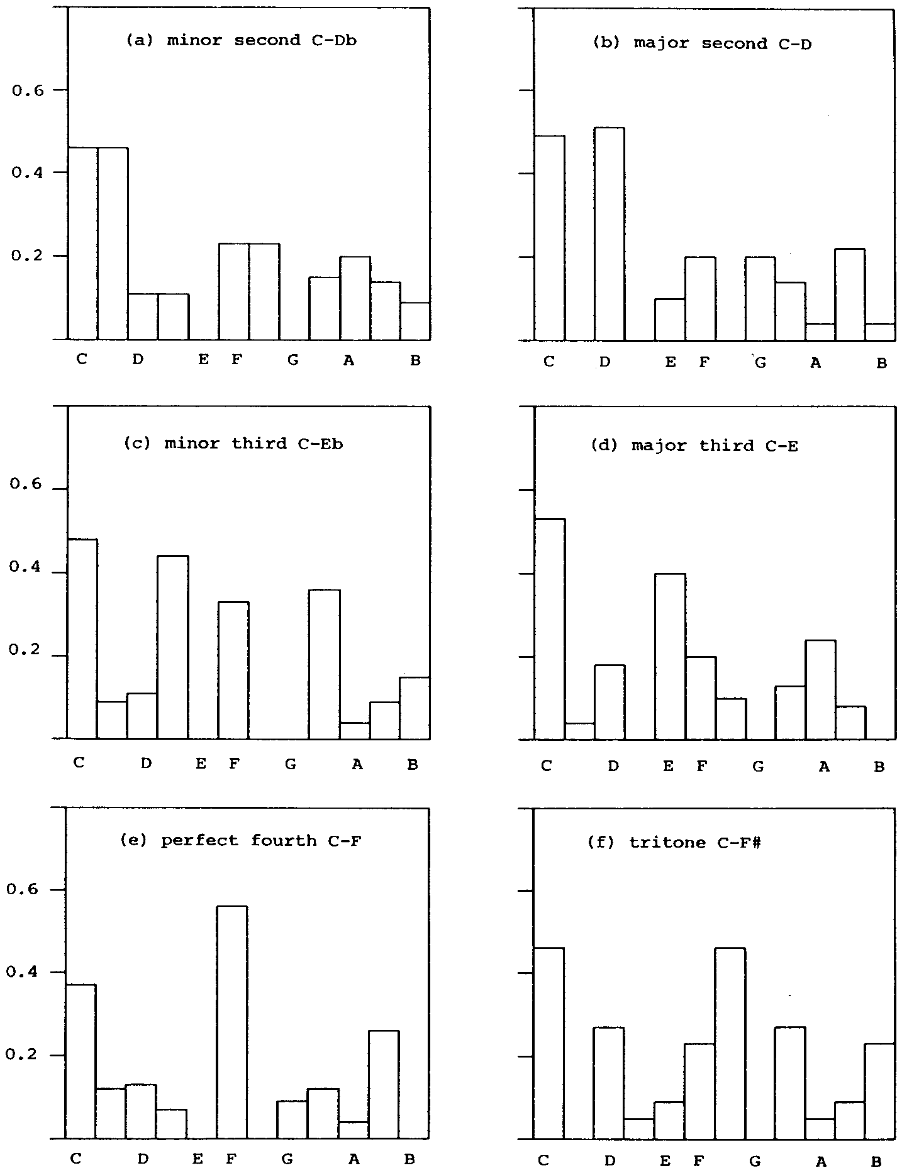

Fig. 1. Distributions of calculated salience *S* against pitch class *p* for six dyad classes. Each applies to inversions as well as compound intervals (e.g., part (a) applies to the intervals m2, M7, m9, M14 et.).

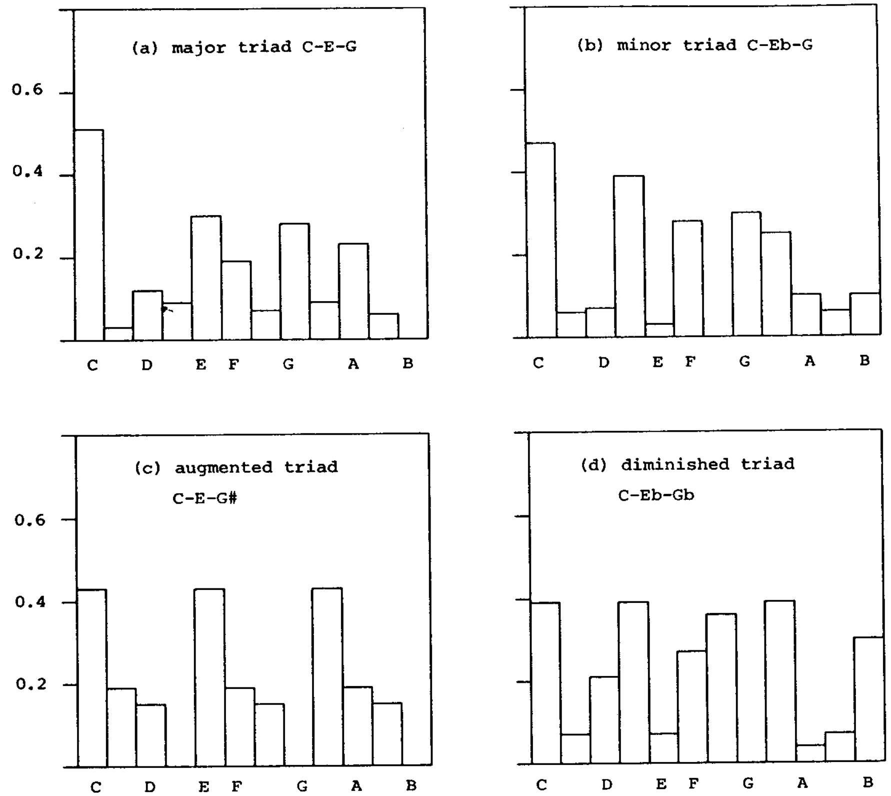

Fig. 2. Distributions of calculated salience *S* against pitch class *p* for four triad classes. Each applies to all difference voicings (inversions, spacings) of a chord.

thirds. In the case of the minor seventh chord, the third of the chord is predicted to be just as salient as the conventional root, implying that the chord's first inversion—the major sixth chord—should be no more or less consonant (or frequently used) than its root position. This implication is borne out, for example, in the case of the diatonic II^7 chord in a major key, which often appears in first inversion (as a IV^6) in preparation for a perfect cadence. The third of the half-diminished seventh chord is predicted to be considerably more salient than its conventional root, which explains why this chord is found less often in root position than in first inversion (e.g., as a IV^6 chord preceding a perfect cadence in a minor key), and why only

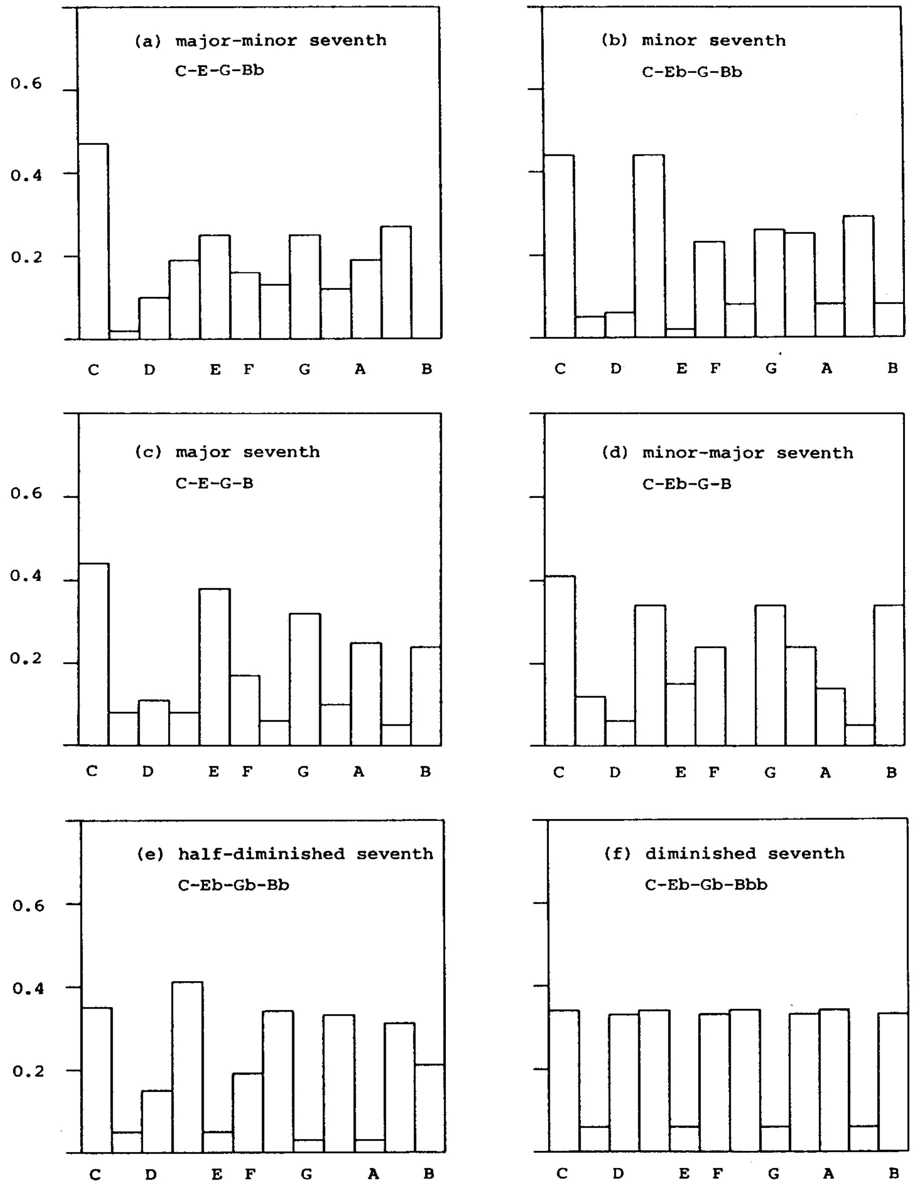

Fig. 3. Distributions of calculated salience S against pitch class p for six tetrad (seventh chord) classes.

the first inversion of the chord can function as a tonic chord (e.g., in Gershwin's *Summertime*). The calculated salience distribution of the diminished seventh chord suggests no less than eight possible roots, all of which are possible in music theory. It also shows how the "diminished" or "octatonic" scale C D E♭ F F♯ G♯ A B (as found in the works of Debussy, some jazz styles, etc.) may be derived from the chord. This is an example of how the model may be used to explain the general concept of chord-scale compatibility, not only in advanced harmony and jazz theory but also in conventional harmony.

The calculated salience distributions of the M3 and P5 dyads and the major triad are compared with experimental data of Thompson and Cuddy (1988) in Table 5. Thompson and Cuddy constructed these target sounds

TABLE 5

Comparison with Experimental Data

a.	Major Third Dyad C–E											
Pitch Class	C		D		E	F		G		A		B
1. Experimental Data	6.3	3.3	3.6	3.5	5.3	3.8	3.1	5.0	3.3	4.9	3.1	3.6
2. Calc. Tone Saliences	.53	.04	.18	.00	.40	.20	.10	.00	.13	.24	.08	.00
3. Calc. Pitch Commonalities	.77	.08	.27	.08	.77	.21	.18	.28	.28	.30	.17	.20
Correlation coefficients: $r_{1,2}$ = 0.75; $r_{1,3}$ = 0.85												

b.	Perfect Fifth Dyad C–G											
Pitch Class	C		D		E	F		G		A		B
4. Experimental Data	6.3	2.8	4.7	3.9	4.1	3.9	2.3	6.1	3.9	4.1	3.9	3.5
5. Calc. Tone Saliences	.56	.00	.09	.12	.04	.26	.00	.37	.12	.13	.07	.00
6. Calc. Pitch Commonalities	.81	.04	.31	.22	.20	.33	.04	.80	.13	.21	.21	.13
Correlation coefficients: $r_{4,5}$ = 0.86; $r_{4,6}$ = 0.95												

c.	Major Triad C–E–G											
Pitch Class	C		D		E	F		G		A		B
7. Experimental Data	6.3	2.5	3.5	2.9	4.7	4.7	2.9	5.8	3.5	4.1	4.3	3.3
8. Calc. Tone Saliences	.51	.03	.12	.09	.30	.19	.07	.28	.09	.23	.06	.00
9. Calc. Pitch Commonalities	.74	.09	.34	.18	.61	.26	.13	.67	.20	.32	.20	.23
Correlation coefficients: $r_{7,8}$ = 0.87; $r_{7,9}$ = 0.89												

Note. Experimental data" are those of Thompson and Cuddy (1988). "Calc. Tone Saliences" are tone saliences calculated according to the model developed in this article. "Calc. Pitch Commonalities" are pitch commonality values between target sounds and probe tones calculated according to the model of Parncutt (1987).

from octave-spaced tones (Shepard, 1964)[9] and asked listeners to rate how well, in a musical sense, they went with single octave-spaced probe tones presented just after the targets. Twelve different probe tones were used, corresponding to the 12 steps of the chromatic scale. Average ratings from a seven-point scale (1 = "goes badly," 7 = "goes well") are presented in rows 1, 4, and 7 of the table. Rows 2, 5, and 8 consist of tone saliences calculated according to the present model (see Figures 1d, 1e, and 2a). The results of a more sophisticated calculation, taking into account masking between tone components over the entire audible pitch range and based on the concept of "pitch commonality" between target sounds and probe tones (see Parncutt, 1987), are presented in rows 3, 6, and 9 of the table. The correlation coefficients show that the more sophisticated calculation is more successful in predicting the results of the experiments of Thompson and Cuddy. But the present model provides a reasonable approximation to the experimental data, especially considering its relative simplicity. Its predictions could be improved slightly by adjusting the weights assigned to the root support intervals; however, this would make the model more complex without tangibly improving its usefulness in music theory. In any case, some deviations between theory and experiment are to be expected due to cultural conditioning effects.

Voicing and Context

As mentioned above, the root of a chord depends mainly on the pitch classes that make it up. It depends to a smaller extent on the chord's voicing (inversion, spacing, doubling) and on the context in which the chord appears.

As a general rule, the consonance ("root disambiguity") of a chord is enhanced if root supports (P1, P5, M3, m7, M9, m3) are voiced below detractors (m2, P4, TT, m6, M6, M7), and if strong supports (P1, P5, M3) are voiced below weak supports (M9, m3). For example, in romantic and jazz harmony, m7s are usually voiced below thirteenths (M6s and m6s), and M3s are usually voiced below m3s (m10s or A9s). This may be understood in terms of the arrangement of support and detractor intervals among the audible harmonics of complex tones: tone components at strong support intervals above the fundamental (P1, P5, M3, and their octave equivalents) are mostly lower in pitch than those at weak support intervals (m7, M2). This state of affairs is familiar to the auditory system through regular exposure to complex tones, especially in speech.

9. The octave-spaced components of the tones used by Thompson and Cuddy actually had harmonic overtones up to the fifth harmonic, with amplitudes inversely proportional to harmonic number. Neither of the two models presented here takes these harmonics into account (although they could quite easily be taken into account in the second model). Accounting for third harmonics (in particular) could noticeably improve predictions.

Harmonics above the tenth are occasionally audible in ordinary complex tones. Most of these correspond to "detractors" in musical chords. They fuse into the overall sound of the complex tone only if they are appropriately high-pitched (compare Schouten's 1940 concept of the "residue"). When such "detractors" occur at lower pitches, they tend to be heard as separate from the rest of the complex tone (e.g., when individual tone components are heard out of bell sounds). Similarly, chords in which weak supports or detractors are voiced among or below strong supports tend to sound bitonal—they seem to have more than one root. For example, the chord C–B♭–E–A has only one clearly defined root (C), while the root of the chord C–A–E–B♭ may be either C or A. This example is also a good illustration of the dissonant effect of voicing a detractor below a support (in this case, M6 below m7).

The bass note of an ordinary, "monotonal" chord is always a root support; it cannot be a detractor, as this would change the root. So the bass note of an inverted chord must lie at an interval of P5, M3, m7, m3, or M2 above the "missing root" in the bass. Most inverted chords have a P5, M3, m7, or m3 in the bass. An example of a chord whose bass lies an M2 above the root is a dominant eleventh chord with missing third and fifth: a IV triad on a V bass. Because the M2 is such a weak root support, however, the root of this chord is ambiguous; it may be regarded as either IV or V. An apparent exception to the proposed rule involves the third inversion of the major seventh chord (e.g., the chord B–C–E–G). In the proposed theoretical framework, this chord may be regarded either as a C major triad with a passing note (M7) in the bass, or as a second inversion E minor triad with an m6 added above the root.

The strongest root support that can appear in the bass is the root itself. Consequently, any note in a chord is more likely to be perceived as the root if it is voiced in the bass. For example, the root of a minor seventh chord (Figure 3b) corresponds to the bass note, not only when the chord appears in root position, but also when it is in first inversion (major sixth chord). This effect is strengthened by increasing the pitch distance between the bass and the upper notes of the chord, that is, by widening the chord's spacing.

The root of a chord also depends to a small extent on the *context* in which it appears. This effect is usually only noticeable for chords whose roots in isolation are already quite ambiguous. In a tonal context, important scale degrees (such as I, V, and IV) are more likely to act as roots than other pitches (such as VII). For example, although the Neapolitan sixth chord is "actually" just the first inversion of a major triad on ♭II, it may be argued that the "functional" root of this chord corresponds to its bass note, IV.

The chances of a note being heard as a root are increased if that note is *doubled*, that is, if it appears simultaneously in more than one octave. Doubling the root of a chord reduces the chord's root ambiguity, and thereby increases the consonance of the chord. Other notes may also be doubled without affecting consonance very much, providing they are already likely candidates for the root, or tonally important in context (e.g., the fifths of major and minor triads: see Figure 2). The doubling of notes that are neither likely root candidates nor tonally particularly strong—for example, the third of the major triad or the conventional root of the diminished triad, especially when these doubled notes coincide with the tonally weak leading note—can noticeably increase root ambiguity and thereby reduce consonance. These observations readily explain the "doubling rules" of conventional harmony theory.

The combined effects of voicing and context can even affect chords whose roots are normally considered to be quite clear and unambiguous. Consider the chord G–C–E (where G is in the bass). According to conventional music theory (as well as the model developed above), the root of this chord is clearly C. This is the case, for example, in the progression A–C–D–F G–C–E–G F–C–F–A, where the middle chord clearly functions as a second inversion. In the common cadential progression G–C–E to G–B–D, however, the first chord is often regarded as a doubly-suspended G chord: as a "cadential 6/4" in G major. This is because the pitch G is harmonically very important in the context of the chord. The "G-ness" of the first chord is further enhanced by the doubling of the G in conventional four-part harmony. It is apparent from this example that *all* chord classes have some degree of root ambiguity.

The Root Position Half-Diminished Seventh Chord

The root of the chord C–E♭–G♭–B♭ (E♭ minor with C in the bass) is highly ambiguous, and deserves special consideration in light of the principles outlined above. According to the model developed in this article, the root of the chord *class* C–E♭–G♭–B♭ (i.e., the root of the chord when its voicing—including its inversion—is not specified) is E♭ (see Figure 3e). But when C appears in the bass of the chord, E♭ cannot be the root, as C—which lies an M6 above E♭—detracts from the (potential) root E♭. G♭ cannot be the root either, as the bass, C, is a tritone away from G♭ and the TT is also a detractor interval. B♭ could be the root, as the bass, C, supports it (via the root-support interval M2); however, the M2 is a relatively weak

root support, and B♭ is already fifth in the predicted rank order of possible roots of the chord class C–E♭–G♭–B♭.

By a process of elimination, then, the only possible roots of the chord C–E♭–G♭–B♭ (with C in the bass) are C and A♭. Their saliences, according to the model, are about equal. They are also about the same as the relatively low saliences of the many possible roots of the diminished seventh chord (cf. parts e and f of Figure 3). The root of the chord is therefore not only quite ambiguous but also quite weak. Given that none of the notes of the chord are doubled, the "actual" root of the chord can only be determined by context. If the context is one of B♭ minor, the root is C (II), as A♭ (♭VII) does not figure prominently in this key. If the context is one of D♭ major, the root is A♭(V), as this scale degree is tonally stronger and more stable than C (VII).

The first chord of Wagner's opera *Tristan und Isolde*, the chord F–B–D♯–G♯, is enharmonically equivalent to a half-diminished seventh chord in conventional root position. Some music theorists (e.g., Kurth, 1920/1923) have interpreted the G♯ in this chord as an accented appoggiatura to the A which follows it; however, this interpretation is not very feasible, considering that the G♯ is held for five slow quaver beats, while its "resolution", the A, is held for only one. According to the present analysis, the only two possible roots of the *Tristan* chord are F (as might be suggested by the conventional theory of superposed thirds) and C♯ (as predicted by Terhardt, 1982). The melodic fragment A F E, which precedes the chord, weakly suggests that F is a more likely root for the chord than C♯. The chord that immediately follows the *Tristan* chord seems to confirm this guess, but—again—only weakly. It is a French augmented sixth chord in the key of A minor (F–B–D♯–A), whose root, according to the model developed in this article, is F. (B cannot act as the root of this chord, as the bass note F, at the interval of a tritone, would detract from it.) The root of the *Tristan* chord is so weakly suggested by the chord in isolation and so weakly confirmed by its context that it is reasonable to assert that the chord has no root at all.

The striking effect of the *Tristan* chord is largely due to the relationship between the chord (regardless of any root that may be assigned to it) and the tonal context in which it appears. The chord is enharmonically equivalent to the diatonic chord II^7 in D♯ minor (with root E♯) or V^9 in F♯ major (with "missing root" C♯) (see, e.g., Arend, 1901). In the *Tristan* prelude, the chord appears in a context of A minor, a key that is diametrically opposite to D♯ minor on the cycle of fifths. The chord itself is familiar, yet its context is unfamiliar; hence the mysteriously delicious feeling of *déja entendu*. The feeling of mystery is enhanced by the fact that the key of A minor is itself only weakly defined.

Construction of Chords

Rameau's approach to harmony suggests that chords are built by combining notes in a way that imitates the overtone structure of a single complex tone. However, successful composers (including Rameau himself) do not seem to have invented new harmonies and chord progressions in this way at all. Instead, they experimented with sound and followed their intuition. The experience of "hearing out" the harmonics of a complex tone is hardly likely to inspire the composition of new harmonies and chord progressions.

A more appropriate basis for an acoustical theory of the construction (and historical development) of musical chords is provided by psychoacoustical concepts of musical consonance. According to Terhardt (1977), consonance in Western music has two main components. The first involves beating between simultaneous tone components. Beating at frequencies greater than 20 Hz is perceived as "roughness." The roughest musical intervals (dyads), for tones with typical spectral envelopes, are the seconds (m2 and M2) and their inversions (sevenths) and compounds (ninths). The second component of consonance, according to Terhardt, involves the root of a chord. The more ambiguous the root, the more dissonant the chord.

The harmonic vocabulary of Western music may be regarded as the result of a cumulative historical process of trial and error, by which a wide range of different combinations of tones and intervals were tested in isolation and in context by generations of composers, performers, and audiences. Gradually, more and more dissonant tone combinations came to be tolerated and, eventually, enjoyed.

The prominence of the major and minor triads (as opposed to any other triads) in Western tonal music may be understood within this paradigm. According to the two-component theory of consonance, the most consonant chords have unambiguous roots and low roughness. The strongest root-support interval class (besides the octave/unison) is the perfect fifth; and the roughest dyad classes are the major and minor seconds. The major and minor triads are the only possible simultaneities of three difference pitch classes that contain a perfect fifth and no seconds.

Conclusions and Applications

The conventional theory of chord classes and their roots provides information on the roots of certain commonly found chords, but does not give an adequate explanation of the perceptual nature of the root, nor does it

explain *why* a particular chord has a particular root or roots. The theory described in this article provides insight into the roots and consonance of chords by relating these musical phenomena to the psychoacoustics of tone perception. In addition, a model is described that is based on this theory and allows the root(s) of any chord to be established in a systematic way. Unlike both the conventional theory of stacked thirds and Terhardt's (1982) chord-root model, the new model makes predictions that are in harmony with musical experience for all important chords in mainstream Western music. These advantages of the model could make it a useful addition to harmony courses and texts.

This article has been primarily concerned with chords as isolated vertical phenomena. From a musical standpoint, this is unsatisfactory: the most interesting aspect of the harmony of a piece of music involves harmonic *progression*. The output of the present model provides the basis for a psychoacoustical analysis of chord progressions. As explained elsewhere (Parncutt, 1987), the strength of the harmonic relationship between two chords may be estimated from their "pitch commonality," calculated from the pitch salience distributions of the chords; and experimentally measured distributions of pitch salience following short chord progressions (Krumhansl & Kessler, 1982) may be accurately predicted by summing calculated pitch salience distributions of individual chords over time. In this way, the model may contribute to an understanding of various aspects of tonality.

Quantitative methods of harmonic analysis based on the model (such as those described above) could be implemented on computer and used in composition. A composition program of this kind could include parameters to control the tonal style of chord progressions at the output. "Composing" with such a program could yield further insight into the nature and origins of tonality.[10]

References

American Standards Association. *Acoustical terminology SI.* New York: American Standards Association, 1960.

Arend, M. Harmonische Analyse des Tristan Vorspiels. *Bayreuther Blätter*, 1901, *24*, 160–169.

Aures, W. *Berechnungsverfahren für den Wohlklang beliebiger Schallsignale, ein Beitrag zur gehörbezogenen Schallanalyse.* Doctoral thesis, Technical University of Munich, 1984.

Babbitt, M. Some aspects of twelve-tone composition. *The Score and I.M.A. Magazine*, 1955, *12*.

10. For helpful comments and discussions, I am grateful to Ernst Terhardt, Bob Stuckey, Peter Maddox, Rex Eakins, Margaret Gummow, and Neil Buckland. This study grew out of my Ph.D. research at the University of New England, which was supervised by Neville Fletcher (physics), Catherine Ellis (music), and William Noble (psychology).

Boomsliter, P., & Creel, W. The long pattern hypothesis in harmony and hearing. *Journal of Music Theory*, 1961, *5*, 2–31.
Fransson, F., Sundberg, J., & Tjernlund, P. The scale in played music. *Swedish Journal of Musicology*, 1974, *56*, 49–54.
Helmholtz, H. L. F. von. *On the sensations of tone as a physiological basis for the theory of music* (A. J. Ellis, Trans.). New York: Dover, 1954. (Original work published 1863)
Hindemith, P. *The craft of musical composition* (A. Mendel, Trans.). Mainz: Schott, 1942. (Original work published 1940)
Krumhansl, C. L., & Kessler, E. J. Tracing the dynamic changes in perceived tonal organisation in a spatial representation of musical keys. *Psychological Review*, 1982, *89*, 334–368.
Kurth, E. Romantische Harmonik und ihre Krise in Wagners "Tristan". Berlin: Hess, 1923 (2nd Ed.) (Original work published 1920).
Malmberg, E. F. The perception of consonance and dissonance. *Psychological Monographs*, 1918, *25*, 93–133.
Moore, B. C. J., Peters, R. W., & Glasberg, R. W. Thresholds for the detection of inharmonicity in complex tones. *Journal of the Acoustical Society of America*, 1985, *77*, 1861–1867.
Parncutt, R. *Sensory bases of harmony in Western music.* Ph.D. thesis, University of New England, 1987.[11]
Plomp, R. Detectability threshold for combination tones. *Journal of the Acoustical Society of America*, 1965, *37*, 1110–1123.
Rameau, J.-P. *Demonstration du principe de l'harmonie.* Paris: Durant, 1750.
Riemann, H. *Vereinfachte Harmonielehre.* London: Augener, 1893.
Schouten, J. F. The residue, a new component in subjective sound analysis. *Proc. Konikl. Ned. Akad. Wetenschap.*, 1940, *43*, 356–365.
Shepard, R. N. Circularity in judgments of relative pitch. *Journal of the Acoustical Society of America*, 1964, *36*, 2346–2353.
Stoll, G. Ambiguity of pitch of complex tones. Contributed paper at the *Fourth Workshop on Physical and Neuropsychological Foundations of Music*, Ossiach, Austria, 1983.
Stuckey, R. *Music in cyclic twelfths.* Unpublished manuscript, London, 1983.
Stumpf, C. Konsonanz und Dissonanz. *Beiträge zur akustischen Musikwissenschaft*, 1898, *1*, 1–108.
Terhardt, E. Pitch, consonance and harmony. *Journal of the Acoustical Society of America*, 1974, *55*, 1061–1069.
Terhardt, E. The two-component theory of musical consonance. In E. F. Evans & J. P. Wilson (Eds.), *Psychophysics and physiology of hearing*. London: Academic Press, 1977.
Terhardt, E. Die psychoakustischen Grundlagen der musikalischen Akkordgrundtöne und deren algorithmische Bestimmung. In C. Dahlhaus & M. Krause (Eds.), *Tiefenstruktur der Musik*. Berlin: Technical University of Berlin, 1982.
Terhardt, E., Stoll, G., & Seewann, M. Algorithm for extraction of pitch and pitch salience from complex tonal signals. *Journal of the Acoustical Society of America*, 1982, *71*, 679–688.
Thompson, E. F., & Cuddy, L. L. Acoustic and sensory factors in the perception of chords. Submitted to *Canadian Acoustics*, 1988.
Watson, R. W. *Viennese harmonic theory from Albrechtsberger to Schenker and Schoenberg*. Ann Arbor: UMI Research Press, 1982.
Wertheimer, M. Untersuchungen zur Lehre der Gestalt II. *Psychologische Forschung*, 1923, *4*, 301–350.

11. To be published with the title *Harmony: A Psychoacoustical Approach* by Springer Verlag, Heidelberg.