Marc Leman (Ed.)

Music, Gestalt, and Computing

Studies in Cognitive and Systematic Musicology

With support from:
European Union (Caleidoscope 96/411063)
Fonds voor Wetenschappelijk Onderzoek - Vlaanderen (FWO-V)
International Society for Systematic and Comparative Musicology

Springer

Parncutt, R. (1997). A model of the perceptual root(s) of a chord accounting for voicing and prevailing tonality. In M. Leman (Ed.). *Music, gestalt, and computing - Studies in cognitive and systematic musicology* (pp. 181-199). Berlin, Germany: Springer-Verlag.

A Model of the Perceptual Root(s) of a Chord Accounting for Voicing and Prevailing Tonality

Richard Parncutt

Department of Psychology, Keele University, Staffordshire ST5 5BG, U.K.

Abstract. A theory of harmony is presented that predicts the perceptual root of any simultaneity within the chromatic scale, embedded in the context of any chord progression. The theory takes into account the pitch classes (pcs) of the chord, its voicing, the prevailing tonality, and local voice-leading. Pcs: the perceptual root of an isolated, octave-generalized chord is determined by the root-support intervals P1, P5, M3, m7 and M2. Voicing: other things being equal, the bass note is more likely to function as the root than the other notes. Tonality: the harmonic profile of a chord in a tonal context is given by the sum of the chord's profile in isolation and the pc-stability profile of the prevailing tonality. Voice leading: parallel movement of parts increases the likelihood that the roots of two sonorities will be heard to move in parallel with the voices. Predictions of the theory for the perceptual roots of diatonic triads and 7ths and of common chromatic chords in major and minor keys are intuitively reasonable, but will need to be tested against the results of perceptual-cognitive experiments and systematic studies of scores and improvised performances.

1 Introduction

Existing theories of the root of a chord can be inconsistent and contradictory. Consider, for example, the sonorities D-F-A-C (D in the bass) and F-A-C-D (F in the bass) in the key of C major. Rameau's theory of stacked thirds would have it that both chords are versions of II^7, and so have the same root – the second degree of the scale, $\hat{2}$ (here, D). (When taking voice leading into account, Rameau held that the root of the chord is F if the chord is followed by a C chord, and D if it is followed by a G chord.) According to Riemann's theory of function, both chords may be regarded as subdominants, so both have the root F. In jazz, it is standard practice to notate D-F-A-C as Dm7 (not F/D) and F-A-C-D as F6 (not Dm7/F), implying that the root depends upon which note is voiced in the bass.

The present article develops a harmonic theory that attempts to account for the root of any chord in any tonal context (by *context*, I am thinking primarily of the chords that immediately precede a chord in a progression). The theory addresses primarily how a chord sounds, not how it is notated; to emphasize this difference, the predictions of the theory may be referred to as *perceptual roots*. Incorporating both music-theoretic and perceptual aspects, the theory demonstrates the considerable and still barely tapped potential of perceptual theory as a basis for the theory of tonal music. The juxtaposition of *artistic* (music theory) and *scientific* (music perception) approaches opens up a number of issues that are not yet satisfactorily resolved, but await further research.

Both scale-step and functional theory can feasibly accommodate most diatonic chords. But it is not a straightforward matter to apply these theories to chords such as the cadential $^6_4$ (or second-inversion tonic triad Ic, resolving to a root-position dominant Va) and the augmented 6ths (built by superposing major 3rd and augmented 6th intervals, plus an optional perfect 5th or augmented 4th, on a bass that is usually a minor 6th interval above the tonic). According to the theory of stacked thirds, the root of the cadential $^6_4$ is the tonic ($\hat{1}$); but in Schenker-inspired harmonic theory (e.g. Forte, 1974, it is the dominant ($\hat{5}$)). Within the functional paradigm, arguments may be conceived for classifying the cadential $^6_4$ as a species of sub- or pre-dominant (if it is thought of as leading to or preparing for V^5_3), as a kind of dominant (as a retrospective prolongation of V), or as a tonic (after all, it contains the same three pitch classes, or pcs, as I). Applying the stacked-thirds approach to the augmented 6th chords (in their correct enharmonic spellings) yields the scale-degrees $\sharp\hat{4}$ (for the *Italian* and *German* versions of the chord) and $\hat{2}$ (for the *French*) as root candidates. But these chords usually appear with $\hat{6}$ in the bass, suggesting that $\hat{6}$ is a root candidate; and the most commonly doubled note – at least in the *Italian version* of the augmented 6th chord – is $\hat{1}$, suggesting that $\hat{1}$ is another root candidate.

The aim of the present paper is to develop a theory of the perceptual root that deals more plausibly and consistently with these issues than previous theories. The new theory will be:

- octave generalized (primarily for reasons of convenience, and ease of music-theoretical application);
- implemented in the form of an algorithm whose input can be any chord or chords in the chromatic scale;
- entirely *bottom up*, so that the root-determination procedure is applied in exactly the same way to any chord (giving the theory a clear advantage over the theory of harmonic dualism (von Oettingen, 1866), and later refinements thereof); and
- systematically testable, and hence falsifiable.

The theory will strive to satisfy the following additional criteria:

- to reliably predict the root whenever it is clearly and unambiguously defined in existing music theory;
- to achieve a suitable compromise where existing music theories (stacked thirds, harmonic function) disagree; and
- to plausibly and parsimoniously account for effects of voicing (registral positions of tones in the chord) and tonal context on the perceptual root of a sonority.

2 Enharmonic Spellings and Intonation

Before proceeding with an account of the theory, I should briefly address two important aspects of tonal harmony that the theory does not consider in any way.

First, the theory ignores enharmonic spellings. This somewhat radical omission may be justified as follows. Enharmonic spellings are primarily determined by the context of preceding and following notes (for a concise account of the underlying rules, see Parncutt & Stuckey, 1992). But in the present theory the effect of context on the root is accounted for independently of enharmonic spellings. Consider, for example, the functional and perceptual difference between the dominant 7th and the German augmented 6th chords. Relative to the chromatic scale and taken out of context, these two chords are identical: both have intervals of 4, 7 and 10 semitones above the root (and/or octave equivalents). The augmented 6th chord sounds differently because it is embedded in a different tonal context (represented quantitatively in the present model by the appropriate key profile, see Krumhansl & Kessler, 1982), and because it resolves differently (again, a matter that can be addressed separately in the model). Clearly, there is not enough information present in enharmonic spellings alone to capture the richness of the difference between these two chords, or between any other chords which are identical except for the context in which they are embedded (e.g., diatonic II^7 in a minor key, and the Tristan chord). In sum, enharmonic spellings are primarily notational conventions. They are not directly relevant to the present model, which is primarily about perception.

The theory also ignores intonation. The results of several published studies on intonation suggest that, contrary to musical intuition, sharps are not consistently intoned sharper or flatter than enharmonically equivalent flats (Parncutt & Stuckey, 1992). Assuming that sharps relative to a key signature or scale lie at major 3rd intervals above other scale degrees or salient pitches, and that flats lie at major 3rd intervals below other scale degrees, sharps are *sharper* than enharmonically equivalent flats in *Pythagorean* tuning (where major thirds have frequency ratios of 81:64, or 408 cents), and *flatter* in *just* tuning (where major thirds have frequency ratios of 5:4, or 386 cents). Numerous experimental studies of intonation have shown that good performers tend to tune between these two extremes, tending toward Pythagorean tuning in some musical situations (e.g., fast melodic passages) and toward just tuning in others (e.g., slow harmonic progressions). Thus, intonation is not related in any simple way to enharmonic spelling, but it is instead determined by musical context (for experimental data on this point see Fyk, 1995). The present model consequently makes no reference to intonation; but if a model of intonation were to be built – possibly on the basis of this model, or along the lines of (Agmon, 1993) – available experimental data on intonation suggest that the model should attempt to account for effects of context, but ignore enharmonic spellings.

3 Definitions

Although the concept of root is of central importance to harmonic theory, it is surprisingly poorly defined (Thomson, 1993). At the most simplistic level, the root is the name most commonly given to a chord (e.g., we call C-E-G a *C chord*, not an *E chord* or a *G chord*). But this definition gives no insight into the root's nature and origins. According to scale-step theory (Rameau), the root is the lowest note in a stack of thirds, while functional theory (Riemann) suggests that the root of just about any chord in tonal music corresponds to $\hat{1}$, $\hat{4}$, or $\hat{5}$, depending on whether a chord functions as a tonic, sub- or pre-dominant, or dominant, respectively.

The *perceptual root* may be defined in a number of other ways. One is based on the well-known study of Krumhansl and Kessler (1982). Although the authors did not say so directly, this study clearly demonstrated the psychological reality of chord roots, an achievement that has escaped others with that aim (e.g. Pritschet, 1992). Krumhansl and Kessler presented musically trained listeners with chords constructed from Shepard (octave-complex) tones, both in isolation and in progressions. They then asked the listeners to rate how well a probe tone followed the chord. For isolated major and minor triads, the peak of the resultant profile corresponded to the stacked-thirds root. In Parncutt (1993), I determined the perceptual root of a wider range of chords, by a similar experimental method. Results again agreed broadly with music theory.

Further possible definitions of the root of a chord are based on the statistical analysis of musical scores, which nowadays is best carried out by computer (Huron, 1994). The root may be defined either as the pc most commonly voiced in the bass, or as the pc which is most often doubled in a chord. Similarly, the root may be defined as the pc which is most often played against a given chord during melodic improvisation although in jazz, performers may deliberately avoid playing the root (Järvinen, 1995). I favour these notation and performance based definitions, because they are the most straightforward to implement, apply and test.

As we have seen for the augmented 6th chord, music theories based on notation may yield different root predictions from theories based on perception. The present theory is of the perceptual kind, as are Krumhansl's experiments and Huron's analyses. Predictions of the present theory are intended primarily to be compared with experimental data. If predictions of the theory agree with experimental data but disagree with notation-based music theory, then we may reasonably suspect that notation-based theory is incorrect.

4 Theory

In Parncutt (1988), I developed a model of the perceptual root that navigated a middle way between Terhardt's simple (over-simplified?) model of the root

(Terhardt, 1982) and his much more general pitch algorithm (Terhardt, Stoll, & Seewann, 1982). My aim was to retain only enough detail to predict the perceptual roots of commonly used chords. The model incorporated five intervals called *root-supports* (a term originally suggested by R. Stuckey): perfect unison (P1), perfect fifth (P5), major third (M3), minor seventh (m7) and major second (M2), plus enharmonic and octave equivalents. These intervals, listed here in decreasing order of importance for root determination, correspond to the intervals between the first ten harmonics and the fundamental of a harmonic complex tone, when they are collapsed into a single octave. The algorithm successfully predicts the perceptual roots of commonly used chords presented in isolation (e.g., major, minor, diminished, augmented and suspended triads, and Mm7, mm7, MM7, hd7 and d7 chords). Note that voicing is ignored; predictions may be regarded as averaged over various different voicings.

In the present article, the theory of root-supports will be regarded as only one (i) of four separate factors influencing the root of a chord. The three additional perceptual effects are (ii) voicing (addressed qualitatively in Parncutt, 1988), (iii) prevailing tonality (involving key profiles Krumhansl & Kessler, 1982), and (iv) melodic streaming (based on experiments performed by Bregman and colleagues (Bregman, 1990)). Effects (i) and (ii) apply to a chord presented in isolation, while effects (iii) and (iv) depend upon the contrapuntal-harmonic progression in which a chord is embedded.

(ii) *Voicing.* In Parncutt (1988), I assumed that, other things being equal, the bass note is more likely to function as a chord's perceptual root than the other notes. Consider, for example, the minor-7th chord D-F-A-C, or II^7 in the key of C major. Its first inversion, IV6, can also function as a root-position chord (Rameau's "sixte ajoutée"). The theory of root-support intervals predicts that both D and F may act as perceptual roots; the actual perceptual root is determined by which of these two is in the bass. Another familiar example of the voicing principle is the second inversion of the tonic triad (cadential $^6_4$). The theory of root-support intervals predicts that the 4th above the bass (the tonic) is the perceptual root, but according to the voicing principle the bass is the root. The tension between these two root candidates may explain the chord's dissonance, the chord's relative rarity by comparison to $^5_3$ and $^6_3$ chords, and the special procedures by which $^6_4$ chords are typically prepared and quitted.

(iii) *Prevailing tonality.* A computer-implemented analysis of a representative sample of tonal scores (D.Huron, personal communication) has revealed that the distribution of doublings of chromatic scale-steps within a given major/minor tonality corresponds remarkably closely with the corresponding Krumhansl key profile. In other words, a note's suitability for doubling corresponds to the perceived goodness of fit between it and the passage in which it is embedded. Huron's finding is consistent with the following two hypotheses: (i) the perceptual root of a chord is generally the best note in a chord to double; and (ii) a note is more likely to be the root if it corresponds to a strong scale-degree such as $\hat{1}$ or $\hat{5}$, whereas it is unlikely to act as a perceptual root if it is tonally unstable (leading note or non-diatonic note). Thus, for example, the root of a Neapolitan 6th chord is normally regarded as $\hat{4}$ (consistent with its interpretation as a sub- or pre-dominant harmony) rather than $b\hat{2}$, even though according to (i) above the latter is the only pitch whose root function is *supported* by the other notes in the chord. Another factor pointing to $\hat{4}$ as the perceptual root of the Neapolitan is voicing – the root is in the bass.

To recapitulate briefly, the theory of stacked thirds is incapable of explaining the root of the Neapolitan 6th, or of the cadential $^6_4$. Instead, these examples demonstrate that any note of a major triad and, by extension, any note of any chord can function as the root, provided the chord appears in a suitable context. Only two conditions would appear to need be satisfied: first, put that note in the bass; and second, line it up with a stable pitch a major/minor key.

(iv) *Local voice leading.* To understand how voice leading can affect the perceptual root, it is instructive to re-examine relevant experiments on melodic streaming from a music-theoretic viewpoint. Bregman (1990) has enumerated several factors that cause the pure-tone components of time-varying sounds to fuse into single streams. One of these is coherent frequency variation: If two pure tones move in parallel motion, maintaining a constant frequency ratio, they are likely to be perceived as a single *event*, even if the ratio between them is not harmonic (although harmonicity enhances the effect). The effect is strongest when the frequency ratio is held strictly constant. Musically, this means maintaining a constant number of semitones (a well-known example is the parallel motion that occurs in the melody line of Ravel's *Bolero*). Fusion is also observed experimentally, although less strongly, when the frequency ratio is almost constant. Musically, this may occur in diatonic stepwise progressions of triads (either in root position, as in Debussy's *La cathédrale engloutie*, or in first inversion, as in Renaissance Fauxbordon).

Parallel movement has at least two separate perceptual effects. First, the voices involved tend to lose their perceptual independence; so if perceptual independence is regarded as desirable, parallel 5ths and 8ves should be avoided. Second, and most relevant to the present argument, the parallel movement tends to increase the likelihood that roots of the two sonorities will be heard to move in parallel with the voices. For example, the chord A-C-F tends to be heard as an A chord if it is followed by A-C-E, but as an F chord if followed by G♯-B-E, producing the sliding root progression F-E. In the case of parallel 5ths, the lower tones of the fifths tend to be heard as roots.

5 Model

A model based on the first principle (root-support intervals) has been presented and tested elsewhere (Parncutt, 1988, 1993). Here, that model is retained with minor amendments, and extended to encapsulate, as parsimoniously as possible, for the second (voicing) and third (key) principles described above. The fourth principle (local voice leading) is not yet implemented.

Input. The notes of a chord are input to the model as a vector *N(p)*. The integer variable p denotes pitch class (pc, or chroma), and varies between 0 (corresponding to the note C) and 11 (B). If $N(p) = 1$, pitch p is included in the chord; if $N(p) = 0$, pitch p is not included. For example, a C-major triad is given by $N(0) = N(4) = N(7) = 1$, and $N(p) = 0$ for all other p. Equivalently, we may write that $N = \{1,0,0,0,1,0,0,1,0,0,0,0\}$.

Root-support intervals. The theory presented in Parncutt (1988) is implemented by means of a vector of root-support weights *w(i)*, where i is interval in semitones and varies from 0 to 11. Previously, I estimated the relative importance of the root-support intervals by assigning the following weights to them: P1=1, P5=1/2, M3=1/3, m7=1/4 and M2=1/5. That is, I set $w = \{1,0,1/5,1/10,1/3,0,0,1/2,0,0,1/4,0\}$. These weights have since been adjusted in three ways. (i) My argument supporting the assignment of a non-zero weight to the m3 interval (Parncutt, 1988, p.75) later turned out to be flawed, so I reduced $w(3)$ to zero. (ii) While comparing model predictions with experimental data (Parncutt, 1993), I found that my original values for the less important intervals (m7, M2) were somewhat too high by comparison to the values for the more important intervals (P1, P5). A similar conclusion followed from comparisons of calculated profiles for major and minor triads with data of Krumhansl and Kessler (Parncutt, 1994). The relative size of the corresponding values of $w(i)$ were adjusted accordingly. (iii) It is easier to remember and work with the weights if they are whole numbers, so I multiplied them by 10 and rounded. After these three adjustments, $w = \{10,0,1,0,3,0,0,5,0,0,2,0\}$. I should stress that the values in this vector are still only estimates – it has so far proven impossible to measure them directly – but their inherent imprecision is not a significant problem for the theory, because small variations in the values have relatively little effect on the model's output.

The two vectors are now combined to produce a vector of pc-weights $W(p), p = 0,11$. The first element of W is produced by vector multiplication: $W(0) = Nw$. In the case of a C-major triad, $W(0)$ = 1x10 + 0x0 + 0x1 + 0x0 + 1x3 + ... 0x0 = 10 + 3 + 5 = 18. Other elements of W are produced by performing cyclic permutations of w before multiplying. So $W(1) = Nw'$ where $w' = \{0,10,0,1,0,3,0,0,5,0,0,2\}$; $W(2) = Nw''$ where $w'' = \{2,0,10,0,1,0,3,0,0,5,0,0\}$; and so on. Another way of performing this procedure is to create a 12x12 *circulant matrix* (Ferrar, 1941) as follows:

$$w_c = \begin{array}{cccccccccccc}
10 & 0 & 2 & 0 & 0 & 5 & 0 & 0 & 3 & 0 & 1 & 0 \\
0 & 10 & 0 & 2 & 0 & 0 & 5 & 0 & 0 & 3 & 0 & 1 \\
1 & 0 & 10 & 0 & 2 & 0 & 0 & 5 & 0 & 0 & 3 & 0 \\
0 & 1 & 0 & 10 & 0 & 2 & 0 & 0 & 5 & 0 & 0 & 3 \\
3 & 0 & 1 & 0 & 10 & 0 & 2 & 0 & 0 & 5 & 0 & 0 \\
0 & 3 & 0 & 1 & 0 & 10 & 0 & 2 & 0 & 0 & 5 & 0 \\
0 & 0 & 3 & 0 & 1 & 0 & 10 & 0 & 2 & 0 & 0 & 5 \\
5 & 0 & 0 & 3 & 0 & 1 & 0 & 10 & 0 & 2 & 0 & 0 \\
0 & 5 & 0 & 0 & 3 & 0 & 1 & 0 & 10 & 0 & 2 & 0 \\
0 & 0 & 5 & 0 & 0 & 3 & 0 & 1 & 0 & 10 & 0 & 2 \\
2 & 0 & 0 & 5 & 0 & 0 & 3 & 0 & 1 & 0 & 10 & 0 \\
0 & 2 & 0 & 0 & 5 & 0 & 0 & 3 & 0 & 1 & 0 & 10
\end{array}$$

and to set $W = Nw_c$.

The vector W is a series of pc-weights whose values are only meaningful relative to each other. For a C-major triad, for example, $W = \{18, 0, 3, 3, 10, 6, 2, 10, 3, 7, 1, 0\}$; this means that, when a C-major triad is sounded, C (W=18) is typically more salient than E and G (W=10), and F (W=6) is more salient than F♯ (W=2).

The absolute values of the weights become meaningful in the present model by normalization, achieved by multiplying each pc-weight by a constant factor such that the mean of all 12 weights is 10 (an arbitrary value chosen for convenience). Some examples of normalized pc-salience profiles are presented below. Each element is rounded to the nearest whole number. The top line of the table shows the interval in semitones above the conventional (stacked-thirds) root. For clarity, salience values corresponding to actual notes are printed in bold.

semitones	0	1	2	3	4	5	6	7	8	9	10	11
maj. triad	**34**	0	6	6	**19**	11	4	**19**	6	13	2	0
min. triad	**29**	2	4	**25**	0	15	0	**19**	15	4	2	6
dim. triad	**19**	2	10	**19**	2	13	**19**	0	19	0	2	15

Voicing. The effect of voicing on the perceptual root is accounted for in the model by adding a constant to the normalized weight at the pc of the bass note. I have arbitrarily chosen the value 20, as it consistently gives results that agree with musical theory and practice and with my musical intuition. As an example, consider the major triad. If the triad is in root position, the weight assigned to the stacked-thirds root is boosted by from 34 to 54 in the above table. If it is in first inversion, the weight assigned to its 3rd is increased from 19 to 39.

It would of course be preferable to establish the value of the added constant more systematically, but the necessary experimental work has not yet been done. Profiles produced by the model would need to be compared with experimentally determined profiles of chords in the context of a chord progression in which each tone was harmonic-complex – not octave-complex, as in

(Krumhansl & Kessler, 1982) – so that chords could be presented in different inversions. Probe tones, on the other hand, would need to be octave-complex, because the model is octave-generalized.

Prevailing tonality. To account, in a simple but appropriate way, for the effect of the prevailing tonality on the perceptual root, it is assumed that the chord is embedded in an unambiguous tonal context, a major or minor key. A suitably normalized version of the pc-stability profile of that key (after Krumhansl & Kessler, 1982) is simply added to the chord's profile in isolation.

The normalization of the key profiles is carried out in two stages. First, a constant is subtracted from all 12 values such that the minimum value of the profile becomes zero. Second, all 12 values are multiplied by a constant value such that the mean of all profile elements is (again) 10. The resultant normalized stability profiles are shown below (again, with elements rounded to the nearest integer, and values corresponding to actual notes in bold):

semitones	0	1	2	3	4	5	6	7	8	9	10	11
major key	$\hat{1}$		$\hat{2}$		$\hat{3}$	$\hat{4}$		$\hat{5}$		$\hat{6}$		$\hat{7}$
profile	**33**	0	**10**	1	**17**	**15**	2	**24**	1	**11**	0	**5**
minor key	$\hat{1}$		$\hat{2}$	$\hat{3}$		$\hat{4}$		$\hat{5}$	$\hat{6}$			$\hat{7}$
profile	**28**	3	**9**	**21**	3	**9**	2	**17**	**12**	3	8	**6**

To account for the effect of prevailing tonality on the root, the pc-salience profile of a chord is rotated around the pc-cycle until its first element corresponds to the prevailing tonic, and then added to the stability profile of the prevailing tonality. The resultant profile is the predicted goodness-of-fit tone profile of the chord in context. The peak of the resultant profile is the predicted root of the chord in context.

The procedure is inspired by experimental results of Järvinen (1995). He obtained 12-element profiles for the frequency-of-occurrence of pcs that were played during improvised jazz solos over a blues progression. Profiles for individual chords were influenced by a combination of the chord itself (the local hierarchy) and the prevailing tonality (the global hierarchy).

If the tonality of a chord progression is ambiguous, modal, or in a process of change (during a modulation), the global hierarchy will not correspond simply to one of the Krumhansl profiles. In this case, the model may be adapted by replacing the Krumhansl profiles by *aggregate pc-salience profiles*, calculated by summing across time the profiles of several preceding chords and accounting appropriately for the decay of sensory memory (Huron & Parncutt, 1993).

Local voice leading. This effect has not yet been formulated quantitatively in the model and will not be considered further in the present article.

Interpretation of output. Elements in the profiles output by the above procedure typically vary in the range 0 to 80. Given their inherent imprecision, different elements should not always be regarded as significantly different. In the following applications, I will assume that values are different only if they differ by 5 or more. Values differing by 4 or less will be regarded as the same. So the root of a chord will be regarded as ambiguous if the profile has two or more peaks whose heights differ by less than 5 units. The value 5 is arbitrary and may be estimated more precisely on the basis of future experimental data.

6 Predictions

Let us first of all check that the model makes reasonable predictions for diatonic triads. The table below shows the profile generated by the model for each of the seven root-position diatonic triads in a major key. The top line of the table (*sem*) shows the number of semitones above the tonic. The second line shows major scale degree numbers. Calculated salience values for the notes of each chord are printed bold, bass notes are italicized, and predicted roots are underlined:

sem	0	1	2	3	4	5	6	7	8	9	10	11
	$\hat{1}$		$\hat{2}$		$\hat{3}$	$\hat{4}$		$\hat{5}$		$\hat{6}$		$\hat{7}$
Ia	<u>***87***</u>	0	16	7	**36**	26	6	**43**	7	25	2	5
IIa	35	6	<u>***59***</u>	3	21	**40**	2	39	1	**30**	16	9
IIIa	48	4	12	7	<u>***66***</u>	17	6	**48**	1	27	0	**24**
IVa	**52**	6	23	3	17	<u>***69***</u>	2	29	7	**30**	12	9
Va	44	4	**29**	7	30	17	2	<u>***78***</u>	1	17	6	**24**
VIa	**58**	0	25	1	**36**	30	6	26	7	<u>***60***</u>	2	9
VIIa	35	10	**29**	3	30	**34**	2	<u>43</u>	1	13	16	<u>***44***</u>

Note: The chord symbols in the left-hand column are commonly used in Britain, but with the gradual acceptance of the Schenkerian approach they are becoming less popular. The symbols are based purely on the principle of stacked thirds: the Roman numeral corresponds to the scale-step where the stacking begins (here, only upper case is used (I, II, III, etc.) regardless of the quality of the chord), and the letter shows which note is in the bass (a = root position, b = first inversion, etc.). For present purposes, the symbols are intended to specify only the notes of which the triads are composed (e.g., VIIa consists of the notes $\hat{7}$, $\hat{2}$ and $\hat{5}$, with $\hat{7}$ in the bass), and the symbols are not intended to imply roots or functions (e.g., the root of VIIa is not necessarily $\hat{7}$).

Summarizing the values in the above table, the model predicts that the perceptual root corresponds to the bass for triads Ia, IIa, IIIa, IVa and Va. The perceptual root of VIa is predicted to be ambiguous (it is either the stacked-thirds root $\hat{6}$ or the tonic $\hat{1}$), consistent with Riemann's idea that VIa is a kind of tonic (Tp). Also ambiguous is the root of VIIa (it is either the bass note $\hat{7}$ or the missing stacked-thirds root $\hat{5}$), supporting the notion that VIIa is a species of dominant.

Perhaps the easiest way to interpret these predictions, and further predictions to be presented below, is to define the predicted root as the most appropriate note to double in the chord, in the absence of considerations of voice leading. The table may thus be interpreted as a set of recommended doublings. Doubling the predicted root tends to render the chord more stable or consonant, whereas doubling some other note might make it less stable; so doubling the predicted root is less likely to upset a chord's harmonic character. In this interpretation, the table suggests that the best note to double in a root-position diatonic triad in a major key is its stacked-thirds root, except if the triad is built on the submediant or leading note of the major scale. This is in broad agreement with conventional harmonic theory and the tonal repertoire. Of course, the leading note cannot be doubled (regardless of what the model predicts) if it is constrained to rise to the tonic and parallel octaves are forbidden. The model also suggests in the case of the diminished triad that an implied root on the dominant may be added without changing the chord's harmonic character.

On the basis of dualist harmonic theories (von Oettingen, 1866, and later adherents) – not to mention the many other historic theories based on the harmonic series that, for one reason or another, have regarded the major triad as somehow more natural than the minor – one might expect the salience profiles for diatonic chords in minor tonalities to look quite different from those in major tonalities. But in actual music, the rules according to which scale degrees function in major and minor are surprisingly similar. For example, major and minor modes may be mixed on the same tonic without upsetting the prevailing tonal hierarchy (for examples see Forte, 1974, p.464-465), and the patterns of relative prevalence of specific chord doublings and progressions (defined relative to the diatonic rather than the chromatic scale) are much the same in major and minor tonalities (McHose, 1947).

The predictions of the present model are more consistent with expectations based on real music than are the predictions of more simplistic theories based the harmonic series. Root-position diatonic triads in a minor key turn out to be rather similar to analogous predictions for the major key. In fact, the perceptual roots of all root-position diatonic triads in a minor key are predicted to lie in the bass:

sem	0	1	2	3	4	5	6	7	8	9	10	11
	$\hat{1}$		$\hat{2}$	$\hat{3}$		$\hat{4}$		$\hat{5}$	$\hat{6}$			$\hat{7}$
Ia	***76***	5	13	**46**	3	24	2	**36**	27	7	9	12
IIa	30	18	***48***	23	12	**28**	4	30	**31**	3	27	6
IIIa*	41	5	9	***76***	3	15	8	**36**	23	7	**27**	12
IVa	**47**	18	13	23	8	***57***	4	21	**37**	3	23	6
Va	39	7	**28**	27	16	11	2	***71***	12	9	13	**25**
VIa	**47**	14	13	**40**	8	22	4	17	***66***	3	13	12
VIIa	30	13	**28**	23	16	**28**	2	36	12	5	23	***45***

Here and in the following the asterix (*) against chord III means that its 5th is perfect – corresponding to the natural 7th degree of the minor scale – not augmented. In all other chords, $\hat{7}$ is sharpened (from 10 to 11 semitones above the tonic).

We turn now to inverted triads. Here are predictions for first-inversion diatonic triads in a major key:

sem	0	1	2	3	4	5	6	7	8	9	10	11
	$\hat{1}$		$\hat{2}$		$\hat{3}$	$\hat{4}$		$\hat{5}$		$\hat{6}$		$\hat{7}$
Ib	**67**	0	16	7	***56***	26	6	**43**	7	25	2	5
IIb	35	6	**39**	3	21	***60***	2	39	1	**30**	16	9
IIIb	48	4	12	7	**46**	17	6	***68***	1	27	0	**24**
IVb	**52**	6	23	3	17	**49**	2	29	7	***50***	12	9
Vb	44	4	**29**	7	30	17	2	**58**	1	17	6	***44***
VIb	***78***	0	25	1	**36**	30	6	26	7	**40**	2	9
VIIb	35	10	***49***	3	30	**34**	2	43	1	13	16	**24**

The perceptual roots of Ib and Vb correspond to their stacked-thirds roots. All three notes of IVb are potential roots. The perceptual roots of the first-inversion minor triads IIb, IIIb and VIb correspond to their bass notes. The root of VIIb is also in the bass.

The corresponding predictions for the minor key are as follows:

sem	0	1	2	3	4	5	6	7	8	9	10	11
	$\hat{1}$		$\hat{2}$	$\hat{3}$		$\hat{4}$		$\hat{5}$	$\hat{6}$			$\hat{7}$
Ib	**56**	5	13	***66***	3	24	2	**36**	27	7	9	12
IIb	30	18	**28**	23	12	***48***	4	30	**31**	3	27	6
IIIb*	41	5	9	**56**	3	15	8	***56***	23	7	**27**	12
IVb	**47**	18	13	23	8	**37**	4	21	***57***	3	23	6
Vb	39	7	**28**	27	16	11	2	**51**	12	9	13	*45*
VIb	***67***	14	13	**40**	8	22	4	17	**46**	3	13	12
VIIb	30	13	***48***	23	16	**28**	2	36	12	5	23	**25**

Here, the perceptual roots of Ib, IIb, IVb, VIb and VIIb are all predicted to lie in the bass. To take one of many examples of how the bass of these chords can function as roots in music: the chord progression VIb-I in a minor key

occurs repeatedly in Debussy's *La Mer*; the VIb chord functions as I (that is, with root $\hat{1}$) with a suspended 6th that resolves to the 5th of the following I chord. The root of IIIb* is ambiguous: it is either the stacked-thirds root $\hat{3}$ or the bass $\hat{5}$. Only the root of Vb corresponds to stacked-thirds theory.

According to the model, then, voicing the third instead of the root in the bass affects the root of all diatonic triads except I and V in major keys and V in minor. In this sense, the model is more consistent with a figured-bass approach to harmony (such as Schenker's) than with a stacked-thirds approach (after Rameau) – but the balance could easily be tipped back in the direction of Rameau by reducing the size of the added constant for the bass note from 20 to, say, 10 or 15.

The predicted profiles for second-inversion triads in major keys are as follows:

sem	0	1	2	3	4	5	6	7	8	9	10	11
	$\hat{1}$		$\hat{2}$		$\hat{3}$	$\hat{4}$		$\hat{5}$		$\hat{6}$		$\hat{7}$
Ic	**67**	0	16	7	**36**	26	6	***63***	7	25	2	5
IIc	35	6	**39**	3	21	**40**	2	39	1	***50***	16	9
IIIc	48	4	12	7	**46**	17	6	**48**	1	27	0	***44***
IVc	***72***	6	23	3	17	**49**	2	29	7	**30**	12	9
Vc	44	4	***49***	7	30	17	2	**58**	1	17	6	**24**
VIc	**58**	0	25	1	***56***	30	6	26	7	**40**	2	9
VIIc	35	10	**29**	3	30	***54***	2	43	1	13	16	**24**

The predicted profiles for second-inversion triads in minor keys are:

sem	0	1	2	3	4	5	6	7	8	9	10	11
	$\hat{1}$		$\hat{2}$	$\hat{3}$		$\hat{4}$		$\hat{5}$	$\hat{6}$			$\hat{7}$
Ic	**56**	5	13	**46**	3	24	2	***56***	27	7	9	12
IIc	30	18	**28**	23	12	**28**	4	30	***51***	3	27	6
IIIc*	41	5	9	**56**	3	15	8	**36**	23	7	***47***	12
IVc	***67***	18	13	23	8	**37**	4	21	**37**	3	23	6
Vc	39	7	***48***	27	16	11	2	**51**	12	9	13	**25**
VIc	**47**	14	13	***60***	8	22	4	17	**46**	3	13	12
VIIc	30	13	**28**	23	16	***48***	2	36	12	5	23	**25**

In both major and minor modes, the perceptual root of Ic (the cadential $^{6}_{4}$) is predicted to be ambiguous – it is either the stacked-thirds root $\hat{1}$ or the bass $\hat{5}$. In the major key, the root of Vc is $\hat{5}$, as expected from stacked-thirds theory; the root corresponds to the bass for IIc, IVc and VIIc; all three notes of IIIc, as well as the tonic, are root candidates; and the root of VIc is the tonic $\hat{1}$ or the bass $\hat{3}$. In the minor key, the root of IIIc* is $\hat{3}$ as expected from stacked-thirds theory; of Vc, it is either $\hat{5}$, the stacked-thirds root, or $\hat{2}$, the bass; and the roots of IIc, IVc, VIc and VIIc all lie in the bass.

In general, the perceptual roots of second-inversion diatonic triads are predicted to be more ambiguous than the roots of first-inversion triads. This observation alone cannot, however, account for the rarity of second inversions in common-practice tonal music. A fuller explanation of that phenomenon is beyond the present scope.

The root of the dominant-7th chord in a major key is predicted to be the stacked-thirds root $\hat{5}$ in all inversions except V^7d, for which the root is ambiguous (either $\hat{5}$ or the bass, $\hat{4}$):

sem	0	1	2	3	4	5	6	7	8	9	10	11
	$\hat{1}$		$\hat{2}$		$\hat{3}$	$\hat{4}$		$\hat{5}$		$\hat{6}$		$\hat{7}$
V^7a	41	7	**24**	7	27	**31**	2	***72***	1	16	12	**19**
V^7b	41	7	**24**	7	27	**31**	2	***52***	1	16	12	***39***
V^7c	41	7	***44***	7	27	**31**	2	***52***	1	16	12	**19**
V^7d	41	7	**24**	7	27	***51***	2	***52***	1	16	12	**19**

Much the same applies for the minor key. The root is predicted to be $\hat{5}$ in all cases, but with increasing competition from the bass note in the higher inversions:

sem	0	1	2	3	4	5	6	7	8	9	10	11
	$\hat{1}$		$\hat{2}$	$\hat{3}$		$\hat{4}$		$\hat{5}$	$\hat{6}$			$\hat{7}$
V^7a	36	10	**23**	27	13	**25**	2	***66***	12	7	19	**21**
V^7b	36	10	**23**	27	13	**25**	2	***46***	12	7	19	***41***
V^7c	36	10	***43***	27	13	**25**	2	***46***	12	7	19	**21**
V^7d	36	10	**23**	27	13	***45***	2	***46***	12	7	19	**21**

These results are consistent with the free use of all inversions of dominant 7ths in common-practice tonal music, and the maintenance of the chord's *dominant function* in all inversions.

Predictions look quite different for the diatonic supertonic 7th chord in a major key (II^7, a minor-7th chord). The chord is tonally weaker and more ambiguous than the dominant seventh, and the root is predicted to change whenever the chord is inverted. The predicted root corresponds to the bass in all inversions except for II^7c, where the root may be either the bass $\hat{6}$ or the tonic $\hat{1}$. The principle of stacked thirds, according to which the root should always be $\hat{2}$, is completely overpowered by the competing principles of voicing and tonality:

sem	0	1	2	3	4	5	6	7	8	9	10	11
	$\hat{1}$		$\hat{2}$		$\hat{3}$	$\hat{4}$		$\hat{5}$		$\hat{6}$		$\hat{7}$
II^7a	**49**	4	***54***	2	20	**41**	2	35	6	**26**	13	8
II^7b	**49**	4	**34**	2	20	***61***	2	35	6	**26**	13	8
II^7c	***49***	4	**34**	2	20	**41**	2	35	6	***46***	13	8
II^7d	***69***	4	**34**	2	20	**41**	2	35	6	**26**	13	8

A similar picture emerges in the minor key, where the root of the II^7 (now a half-diminished 7th chord) is predicted to correspond to the bass in all inversions, with some competition from the tonic in the root-position chord II^7a:

sem	0	1	2	3	4	5	6	7	8	9	10	11
	$\hat{1}$		$\hat{2}$	$\hat{3}$		$\hat{4}$		$\hat{5}$	$\hat{6}$			$\hat{7}$
II^7a	***43***	14	***46***	23	10	**30**	4	27	**30**	3	23	6
II^7b	**43**	14	**26**	23	10	***50***	4	27	**30**	3	23	6
II^7c	**43**	14	**26**	23	10	**30**	4	27	***50***	3	23	6
II^7d	***63***	14	**26**	23	10	**30**	4	27	**30**	3	23	6

Let us move on now to some of the more common chromatic chords. The perceptual root of the Neapolitan 6th is predicted to lie in the bass in both major and minor key contexts:

sem	0	1	2	3	4	5	6	7	8	9	10	11
major key	33	**34**	10	7	23	***54***	14	27	**20**	17	14	7
minor key	28	**37**	9	27	8	***48***	14	21	**31**	9	21	8

In a major-key context, the root of the augmented 6th chord is predicted to be ambiguous. It is either the stacked-thirds root $\hat{6}$, or the tonic $\hat{1}$:

sem	0	1	2	3	4	5	6	7	8	9	10	11
	$\hat{1}$		$\hat{2}$		$\hat{3}$	$\hat{4}$		$\hat{5}$		$\hat{6}$		$\hat{7}$
Italian	***52***	10	20	1	25	24	**23**	24	***50***	11	6	15
German	***47***	9	17	**15**	23	25	**18**	24	***50***	11	5	17
French	***49***	7	**31**	1	26	22	**18**	31	***43***	11	9	12

In a minor key, the root of the augmented 6th corresponds to the bass, $\hat{6}$:

sem	0	1	2	3	4	5	6	7	8	9	10	11
	$\hat{1}$		$\hat{2}$	$\hat{3}$		$\hat{4}$		$\hat{5}$	$\hat{6}$			$\hat{7}$
Italian	**47**	13	18	21	10	18	**23**	17	***60***	3	13	16
German	**42**	12	16	**36**	8	19	**18**	17	***60***	3	12	18
French	**43**	10	**30**	21	11	16	**18**	24	***53***	3	16	13

In a major key, the root of the diminished-7th chord on $\hat{7}$ lies in the bass, with competition from both $\hat{5}$ (*missing* stacked-thirds root) and $\hat{1}$ in the case of VII^7a and VII^7d:

sem	0	1	2	3	4	5	6	7	8	9	10	11
	$\hat{1}$		$\hat{2}$		$\hat{3}$	$\hat{4}$		$\hat{5}$		$\hat{6}$		$\hat{7}$
VII^7a	***34***	14	**24**	2	31	**29**	4	***38***	**16**	13	15	***39***
VII^7b	34	14	***44***	2	31	**29**	4	38	**16**	13	15	**19**
VII^7c	34	14	**24**	2	31	***49***	4	38	**16**	13	15	**19**
VII^7d	***34***	14	**24**	2	31	**29**	4	***38***	***36***	13	15	**19**

In the minor key, the root of the diminished 7th on $\hat{7}$ is less ambiguous, with the bass note predominating in all cases:

sem	0	1	2	3	4	5	6	7	8	9	10	11
	$\hat{1}$		$\hat{2}$	$\hat{3}$		$\hat{4}$		$\hat{5}$	$\hat{6}$			$\hat{7}$
VII^7a	29	17	**23**	23	17	**23**	4	31	**26**	5	22	***41***
VII^7b	29	17	***43***	23	17	**23**	4	31	**26**	5	22	**21**
VII^7c	29	17	**23**	23	17	***43***	4	31	**26**	5	22	**21**
VII^7d	29	17	**23**	23	17	**23**	4	31	***46***	5	22	**21**

Finally, consider the (diatonic) half-diminished 7th chord in root position on $\hat{7}$ in a major key. The root of VII^7a is predicted to be either the bass $\hat{7}$ or the *missing* stacked-thirds root $\hat{5}$:

sem	0	1	2	3	4	5	6	7	8	9	10	11
	$\hat{1}$		$\hat{2}$		$\hat{3}$	$\hat{4}$		$\hat{5}$		$\hat{6}$		$\hat{7}$
VII^7a	34	7	**31**	2	27	**33**	2	39	1	**27**	12	***42***

The half-diminished 7th chord in root position occurs diatonically in a minor key as II^7a. The predicted root is either the stacked-thirds root $\hat{2}$ or the tonic $\hat{1}$:

sem	0	1	2	3	4	5	6	7	8	9	10	11
	$\hat{1}$		$\hat{2}$	$\hat{3}$		$\hat{4}$		$\hat{5}$	$\hat{6}$			$\hat{7}$
II^7a	**43**	14	***46***	23	10	**30**	4	27	**30**	3	23	6
TC	29	13	**27**	21	18	9	**18**	28	***49***	5	15	**28**

The perceptual root of the Tristan chord (TC in the above table, enharmonically equivalent to a half-diminished 7th in root position) is predicted to correspond to its bass $\hat{6}$. On the one hand, this result should be taken with a grain of salt, because the model assumes an established minor key context, and the context in which the Tristan chord appears is arguably too tonally uncertain to warrant a Krumhansl minor-key profile. On the other hand, the argument that the bass is the root is strengthened by the observation that the bass and alto voices move in parallel (augmented 6th to minor 7th) as the chord resolves. According to principle (iv) above, this parallel motion reinforces the bass note as a perceptual root.

7 Conclusion

The model systematically makes predictions that are broadly in accord with music theory, and that promise to allow conventional diatonic theory to be extended to accommodate any chord in the chromatic scale, in any chromatic context. The theory may thus be able to bridge the gap tonal and atonal theory, forging a smooth transition where previously there was impassable abyss. It would appear to have considerable potential for shedding light on the harmonic procedures of tonal modernists such as Debussy and Stravinsky, and on the harmony of jazz – particularly that of the bebop period.

The model suggests a new, systematic approach to harmonic function in the Riemann tradition. Riemann pointed out that the function of a chord is often not clear-cut: A chord may have a mixture of tonic, dominant and subdominant functions, in varying proportions. The present model offers a new means of estimating these proportions. The degree to which a given chord represents one of the three primary functions may be predicted by correlating the pc-salience profile of the chord with the pc-salience profiles of suitable tonic, dominant, and subdominant archetypes. For practical purposes, these archetypes may be represented acoustically as triads (I, V, IV) or tetrads (IV^6, V^7) of octave-complex tones. In a similar vein, the theory may shed light on the theory of chord substitutions, especially in jazz. It may be that a chord is a good substitution for another chord if the correlation coefficient between the pc-salience profiles of the two chords exceeds a certain value. Quantitative exploration of these ideas are beyond the present scope.

The main problem with the model as it stands is the paucity of systematic tests. Further improvements are unlikely without detailed comparisons between model predictions and the results of perceptual and cognitive experiments, and of computer-based analyses of musical scores.

The model currently neglects the soprano voice. McHose (1947, p.72) observed that, in major and minor triads in the Bach Chorales, the soprano is doubled more often than the bass, and the bass more often than the middle parts. Given that doubling is one basis for a definition of the root, this observation strongly suggests that the soprano, like the bass, is a root candidate by virtue of its voicing alone, consistent with the idea that the soprano and bass voices are the most perceptually salient voices in a 4-part texture (presumably due to relative lack of masking by other voices). The present model's addition of a constant to account for the salience of the bass should, therefore, be accompanied by similar acknowledgment of the added salience of the soprano. The reason for neglecting the soprano in the present model is that it soprano voicing is also neglected in mainstream music theory, with which the predictions of the model are intended to be compared. It is anticipated that the calculated salience of the soprano voice will need to be boosted when comparing the predictions of the model with future experimental data.

A final drawback of the present model is its neglect of local voice leading – clearly an important influence on the root. It is not yet clear how local voice leading may be appropriately formulated.

References

Agmon, E. (1993). Towards a theory of diatonic intonation. *Interface – Journal of New Music Research, 22*, 151–163.

Bregman, A. (1990). *Auditory scene analysis: The perceptual organization of sound.* Cambridge, MA: The MIT Press.

Ferrar, W. (1941). *Algebra: A text-book of determinants, matrices, and algebraic forms.* Oxford: Oxford University Press.

Forte, A. (1974). *Tonal harmony in concept and practice* (2nd ed.). New York, NY: Holt, Rinehart and Winston.

Fyk, J. (1995). *Melodic intonation, psychoacoustics, and the violin.* Zielona Gora: Organon Press.

Huron, D. (1994). *The Humdrum toolkit reference manual.* Menlo Park, CA: Centre for Computer Assisted Research in the Humanities.

Huron, D., & Parncutt, R. (1993). An improved model of tonality perception incorporating pitch salience and echoic memory. *Psychomusicology, 12*, 152–169.

Järvinen. (1995). Tonal hierarchies in jazz improvisation. *Music Perception, 12*, 415–437.

Krumhansl, C., & Kessler, E. (1982). Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychological Review, 89*, 334–368.

McHose, A. (1947). *The contrapuntal harmonic technique of the 18th century.* Englewood Cliffs, NJ: Prentice-Hall.

Parncutt, R. (1988). Revision of Terhardt's psychoacoustical model of the roots of a musical chord. *Music Perception, 6*, 65–94.

Parncutt, R. (1993). Pitch properties of chords of octave-spaced tones. *Contemporary Music Review, 9*, 35–50.

Parncutt, R. (1994). Template-matching models of musical pitch and rhythm perception. *Journal of New Music Research, 23*, 145-168.

Parncutt, R., & Stuckey, R. (1992). Towards a standard alternative notation and terminology based on the chromatic scale. *Musikometrika, 4*, 117–143.

Pritschet, H. (1992). *Wahrnehmung des Grundtons von Dreiklängen: Eine experimentelle Untersuchung.* Regensburg: S. Roderer.

Terhardt, E. (1982). Die psychoakustischen Grundlagen der musikalischen Akkordgrundtöne und deren algorithmischen Bestimmung. In C. Dahlhaus & M. Krause (Eds.), *Tiefenstruktur der Musik.* Berlin: Technical University of Berlin.

Terhardt, E., Stoll, G., & Seewann, M. (1982). Algorithm for extraction of pitch and pitch salience from complex tonal signals. *The Journal of the Acoustical Society of America, 71*, 679–688.

Thomson, W. (1993). The harmonic root: A fragile marriage of concept and percept. *Music Perception, 10*, 385–416.

von Oettingen, A. (1866). *Harmoniesystem in dualer Entwicklung: Studien zur Theorie der Musik.* Dorpat, Leipzig: W. Gläser.