Contemporary Music Review,
1993, Vol. 9, Parts 1 & 2, pp. 35–50
Photocopying permitted by license only

Pitch properties of chords of octave-spaced tones

Richard Parncutt

Department of Music Acoustics, Royal Institute of Technology, Stockholm, Sweden

Listeners were presented with simultaneities of 1, 2, 3, or 4 octave-spaced (Shepard tones). In Experiment 1, they were asked how many tones they heard in each chord (its *multiplicity*). In Experiment 2, they heard a chord followed by a tone, and were asked how well the tone went with the chord; this resulted in a *tone profile* for each chord. In Experiment 3, they heard successive pairs of chords, and were asked to rate their *similarity*. The experiments may be regarded as octave-generalized versions of experiments reported in Parncutt (1989).

Results were modelled by adjusting and extending a psychoacoustical model for the root of a chord (Parncutt, 1988). The model predicts the multiplicity of a chord, the *salience* (probability of noticing) of each tone in a chord, and the strength of harmonic relationships between chords (*pitch commonality*). Implications for the theory of roots, implied scales, and harmonic relationships are discussed.

KEY WORDS: Pitch salience, chord, root, tone profile, pitch commonality, similarity, multiplicity.

Introduction

In Western music theory chords have roots, and imply scales. For example, a C added sixth chord (CEGA) normally has the root C, and implies the scale of C major. Roots and implied scales are usually somewhat ambiguous: a C^6 chord may have other roots (such as A), or imply other scales (such as G major), depending on context.

This paper investigates the roots and scales implied by musical chords by comparing the results of listening experiments with calculations according to a psychoacoustical model. The model accounts for roots and implied scale tones by means of a single parameter, *pitch salience*. Roots are supposed to have high pitch salience, or perceptual importance; additional, implied scale tones have intermediate pitch salience. The model further accounts for harmonic relationships, measured by similarity judgments of pairs of chords, by means of a parameter called *pitch commonality*. The model explains the sensory origins of roots, implied scales, and harmonic relationships, but neglects culture-specific effects such as conditioning by particular, arbitrary chord sequences.

Octave-spaced tones were used in the experiments so as to enable octave-generalized aspects of music theory to be investigated as directly as possible. By building chords from octave-spaced tones, effects of octave register (pitch height) and voicing were minimized. Remaining register effects (such as the "tritone paradox" investigated by Deutsch, 1987) were avoided in Experiments 2 and 3 by random transposition of trials (see procedure sections).

In this research, octave equivalence is regarded firstly as an axiom of music theory. As a perceptual phenomenon, it is assumed to be primarily learned from music (see e.g. Burns, 1981). It would appear to be unnecessary to postulate a neurophysiological basis for active equivalence, as e.g. Ohgushi (1983) has done, in order to account for octave-generalized aspects of music theory.

The paper begins by considering the number of simultaneously noticed tones in a chord, here called its *multiplicity*.[1] This parameter is later used in the model to scale pitch saliences as absolute values, representing probabilities of noticing.

Experiment 1: Multiplicity

The number of tones simultaneously noticed in a musical chord does not necessarily correspond to the number of pure tone components (Thurlow and Rawling, 1959) or complex tone components (DeWitt and Crowder, 1987; Parncutt, 1989). In the present experiment, sounds were constructed from octave-spaced tones (Shepard, 1964) and listeners were asked how many such tones they heard.

Apart from providing some new experimental data, the experiment aimed to test the algorithm for *pitch ambiguity* in Parncutt (1988). A similar algorithm (for *multiplicity*) is presented below as part of a model for the salience of a chroma (pitch class) in an octave-generalized chord, e.g. a chord made of octave-spaced tones. The multiplicity algorithm allows saliences to be expressed as absolute values: "probabilities of noticing" which may be compared across different chords.

Method

Listeners. 26 adults participated in the experiment. Their musical experience (here measured in terms of the number of years spent regularly practising or performing music, either instrument or voice) had a mean of 11 years and a standard deviation of 10 years.

Equipment. Waveforms were calculated by adding pure tone components in alternating cosine and sine phase (to reduce maximum amplitude) and transferred by analog signal to a digital sampling synthesizer (*Casio FZ-1*). During the experiment, sounds were called via MIDI by a Le_Lisp program running on a *Macintosh II* personal computer. They were amplified and reproduced over a loudspeaker in a sound-isolated room. Listeners responded by pressing keys on the computer keyboard.

Sounds were compared of octave-spaced tones of equal amplitude. By contrast to the tones used by Shephard (1964), pure tone components had equal amplitude (before amplification) across the range 16 Hz to 16 kHz.[2] All pure tone components were tuned to the standard equally-tempered scale, with A = 440 Hz and no octave stretching.

Twenty different sounds were presented, each in two different transpositions, six semitones apart. Pitches were chosen to produce a balance around the chroma cycle, and so not to emphasize any particular pitch. The 20 sounds consisted of one single octave-spaced tone (monad), six dyads of octave-spaced tones (spanning intervals 1 to 6 semitones), five triads (037, 047, 048, 036 and 057) and eight tetrads (047Q, 037Q, 047L, 037L, 0369, 036Q, 057Q and 046Q).[3]

Sounds had durations of 0.2 s. All components in each sound started exactly simultaneously; this was important, as the auditory system is remarkably sensitive to asynchrony in onset times, and uses asynchrony to discriminate musical tones in performance (Rasch, 1978). Overall loudness was adjusted to a comfortable level by each listener.

Procedure. In each trial, one of the 20 sounds was presented twice, with a pause of 0.5 s between presentations. The task was to indicate how many tones they heard in the sounds. No upper limit was set on their responses. Listeners could take as long as they wished to respond. They were asked, however, to respond spontaneously, without thinking too hard. It was stressed that this was not a test of musical ability. Each sound was presented twice, making a total of 40 trials. To avoid serial effects, trials were presented in a random order which was different for each listener.

The experiment was preceded by a practice session. During the practice, listeners were told after they responded whether the chord had contained one, or more than one, (octave-spaced) tone. In the experiment proper, no feedback was given.

Results

The non-musicians initially had some difficulty with the task, but after some practice found they were able to distinguish a small number of response categories (e.g. 1 to 3). Still, they were not sure that they actually heard all the tones that they guessed were present. Musicians' responses generally covered larger ranges. Many said afterwards that they had responded on the basis of musical experience (e.g. responding "4" on recognizing a seventh chord). All results correlated positively with the actual number of tones in the sounds, so none were eliminated from the analysis.

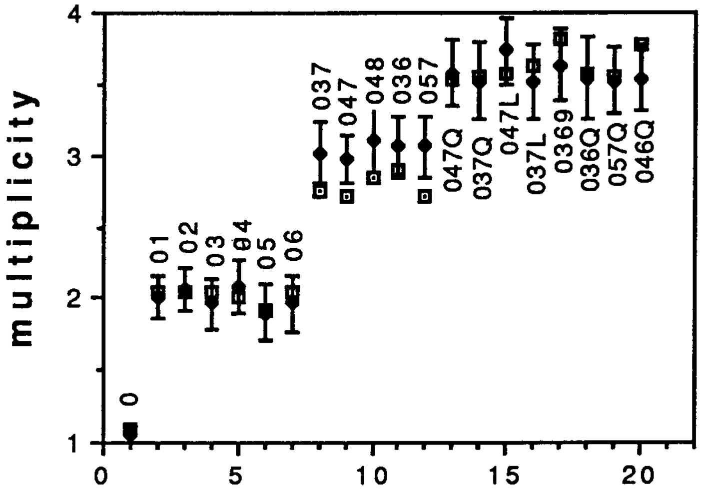

Figure 1 Points: mean responses (52 data per point). Bars: 95% confidence intervals. Squares: calculations according to (7) with $k_M = 45$, $k_W = 3.9$, and $k_S = 0.91$ (see below). Chord classes are indicated by intervals in semitones above the nominal root (e.g. "047" = major triad).

Results are graphed in Fig. 1 as means and 95% confidence intervals of responses.[4] Responses corresponded closely to the "actual" number of octave-spaced tones in each sound (but clearly not to the number of pure tone components) for the monad, dyads and triads. In the case of the monad, the result was not surprising, as listeners had been taught to recognize monads during the practice.[5] The tetrads were heard to contain about 3.5 tones, agreeing with Huron's (1989) finding that the accuracy of identifying the number of concurrent voices in polyphonic music drops markedly at the point where a three-voice texture is augmented to four voices.

There was an additional tendency for consonant sounds to have lower, and dissonant to have higher, multiplicity. So, for example, dyad 05 (perfect fourth/ fifth) was heard to have significantly less ($p < 0.05$) tones than dyad 04 (M3/m6), triad 047 (major) less than triad 048 (augmented), and tetrad 047Q (major-minor) less than tetrad 047L (major) with its dissonant semitone (L0). It is not clear whether this was a "real" effect, due to the number of tones actually heard, or an artifact: many listeners appear to have used dissonance as a basis upon which to guess the number of tones in each sound. Neither of these approaches explains why the mean response for the dissonant 037L (minor–major) tetrad was lower than that for the 047L (major) tetrad.

Experiment 2: Pitch analysis

Krumhansl and Kessler (1982) measured the relative musical importance of the twelve chroma in octave-generalized musical chords by presenting a chord (constructed from octave-spaced tones) followed by a tone (also octave-spaced) and asking how well, in a musical sense, the two went with each other. In the resultant *tone profiles*, the highest peak normally corresponded to the root of the chord, and subsidiary peaks corresponded to other notes. In the experiment to be described, I repeated this paradigm for five different chord classes, and compared results with calculations according to a psychoacoustical model for the root of a chord (Parncutt, 1988).

Method

Listeners. 32 people participated. Of these, 5 were later eliminated, as their results correlated negatively with the actual presence or absence of tones in the chords. The eliminated listeners had very little or no experience of practising or performing music. The remaining 27 had mean 12 years, standard deviation 9 years musical experience.

Equipment was the same as in Experiment 1.

Sounds. Five of the chords presented in the previous experiment were analyzed: the major triad 047, the minor triad 037, the major-minor tetrad 047Q, the half-diminished tetrad 036Q and the diminished tetrad 0369. As before, the chords were composed of octave-spaced tones. Comparison tones were also octave-spaced.

Procedure. The listeners considered this experiment more difficult than Experiments 1 and 3. So in most cases they did if after Experiment 3, to allow themselves more time to become more familiar with the kinds of sounds used, and the general procedure.

The experiment began with a practice session, during which listeners were told (after each response) whether the tone in each chord-tone pair had actually been part of the chord. In the experiment proper, no feedback was given. There were 60 trials, in which 5 chords were each compared with tones at 12 chromatic intervals above their (nominal) roots. Trials were presented in a random order that was different for each listener. Each trial consisted of a chord followed by a tone, then the same chord-tone pair repeated. Both chords and tones had durations of 0.2 s. The time interval between chord and tone was 0.35 s; between repetitions, 0.55 s. Each chord-tone pair was transposed through a random chromatic interval, with the exception of the 12 pairs including the diminished tetrad 0369, which were held at the same pitch (so as to allow investigation of absolute pitch effects: see blow). As before, listeners could take as long as they wished to respond, but were asked to do so as spontaneously as possible.

The task was to indicate how well the tone went with the chord on a scale from 0 (very badly) to 3 (very well). If listeners thought that the tone was actually in the chord, they were asked to select the response "3".

Results

Results for all five analysed chords are shown in Fig. 2.

Mean responses were generally higher for notes actually in the chord. In the case of the major–minor tetrad 047Q and the half-diminished tetrad 036Q, there was a further clear peak in the responses at the root (0 and 3 respectively). So 036Q was most often heard as a minor sixth chord (0379). Surprisingly, in the case of the major (047) and minor (037) triads, the response at the root (0) was not significantly different from that at the fifth (7). In ordinary musical voicings of these two chords, the root is clearly more salient than other chroma (Parncutt, 1989; Terhardt, Stoll, and Seewann, 1982a).

A pitch height (or absolute pitch) effect occurred for the diminished tetrad 0369, which (unlike the other chords) was not subjected to random transpositions. Of the four results corresponding to actual notes, the lowest occurred at interval 9, which in this case was always the note A. This effect is consistent with the finding of Terhardt, Stoll, Schermbach, and Parncutt (1986) that the distribution of the main pitches of octave-spaced tones is centred around 300 Hz, or between D_4 and Eb_4 (Eb_4 lies opposite A on the chroma cycle).

Results for notes not actually present in the chords also showed some structure. The major triad 047 went better with notes in the corresponding major scale than with other notes: specifically, it went better with 2 than 1, 5 than 6, 9 and 8, and 11 than 10. Similarly, the minor triad 037 went better with notes in the corresponding harmonic minor scale (2 than 1, 8 than 6, 10 than 11). The major–minor tetrad 047Q went better with notes in the major scale on 5, the scale in which it is a dominant seventh chord (2 better than 1, 5 than 6, 9 than 8). The half-diminished triad 036Q went particularly well with 8, a note which, if actually present, would function as its root, producing a dominant ninth chord.

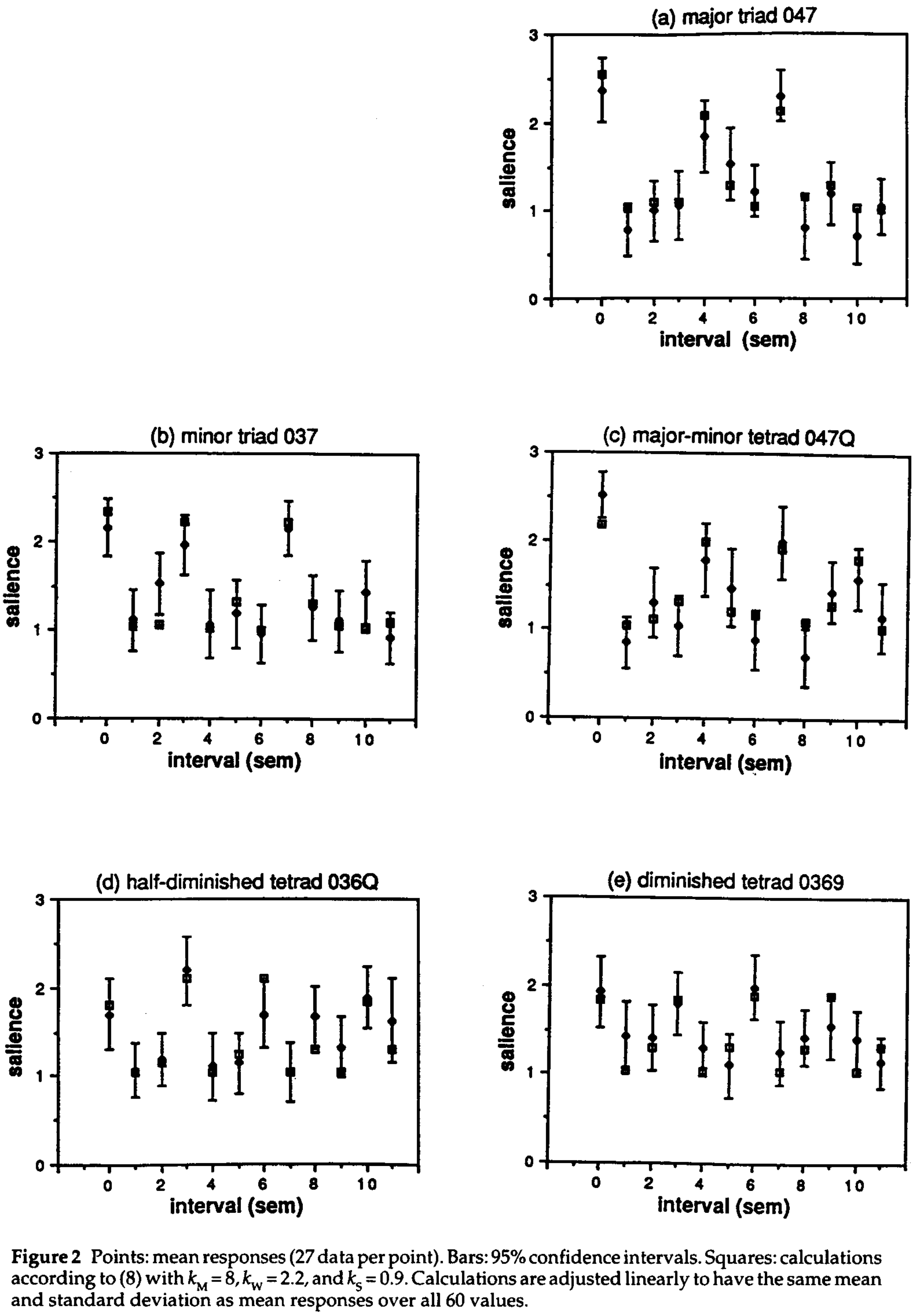

Figure 2 Points: mean responses (27 data per point). Bars: 95% confidence intervals. Squares: calculations according to (8) with $k_M = 8$, $k_W = 2.2$, and $k_S = 0.9$. Calculations are adjusted linearly to have the same mean and standard deviation as mean responses over all 60 values.

According to music theory (as well as the model of Parncutt, 1988), the diminished tetrad 0369 should go well with the notes 1, 4, 7 and Q (all of which can function as the root of the chord). In the experiment, however, these notes went no better with the chord than the remaining chroma (2, 5, 8, L).

Experiment 3: Similarity

The harmonic relationship between musical chords may be modelled by their pitch commonality, or the extent to which they have pitches in common, and measured by similarity judgments (Parncutt, 1989). The aim of the current experiment was to measure the strength of harmonic relationships for some common chord progressions in music theory.

Method

Listeners. 26 people took part, with mean 11 years and standard deviation 10 years' experience of regular musical practice.

Equipment and sounds were as previously.

Procedure. Five chord pairs were each compared with 12 different chromatic intervals between the two roots. In the first 12 trials, both chords were major triads, so the actual chord pairs were 047–047, 047–158, 047–269, . . . 047–L36. In the next 12 trials, major triads were compared with minor triads (047–037, 047–148, . . . 047–L26). Then, major-minor tetrads were compared with major triads (047Q–047, 047Q–158, etc.), and half-diminished tetrads with major–minor tetrads (036Q–047Q, 036Q–158L, etc.). Finally, diminished tetrads were compared with major triads (0369–047, 0369–158, 0369–269), minor triads (0369–037, –148, –259), major–minor sevenths (0369–047Q, – 158L, –2690) and half-diminished tetrads (0369–036Q, –147L, –2580). All but the major–major (047–047, etc) pairs were also presented in the reverse order (e.g. 037–047 as well as 047–037), making a total of 60 + 48 = 108 trials. Each listener heard the trials in a different random order, and each chord pair was transposed through a random chromatic interval.

Listeners were asked to rate the similarity of the chords on a 4-point scale from 0 (very different) to 3 (very similar). Musicians were told that "similarity" may be interpreted as meaning "harmonically related"; they were nevertheless asked to avoid thinking in terms of music theory and to respond instead according to how related they *perceived* the chords to be. As before, listeners were allowed to practice for as long as they wished.

Results

Figure 3 shows the results for all trials. Note that the confidence intervals in part (a) of the figure are larger than in parts (b) to (e), as pairs of major triads (part a) were presented only once in the experiment, while the other chord pairs were presented in two different orders.

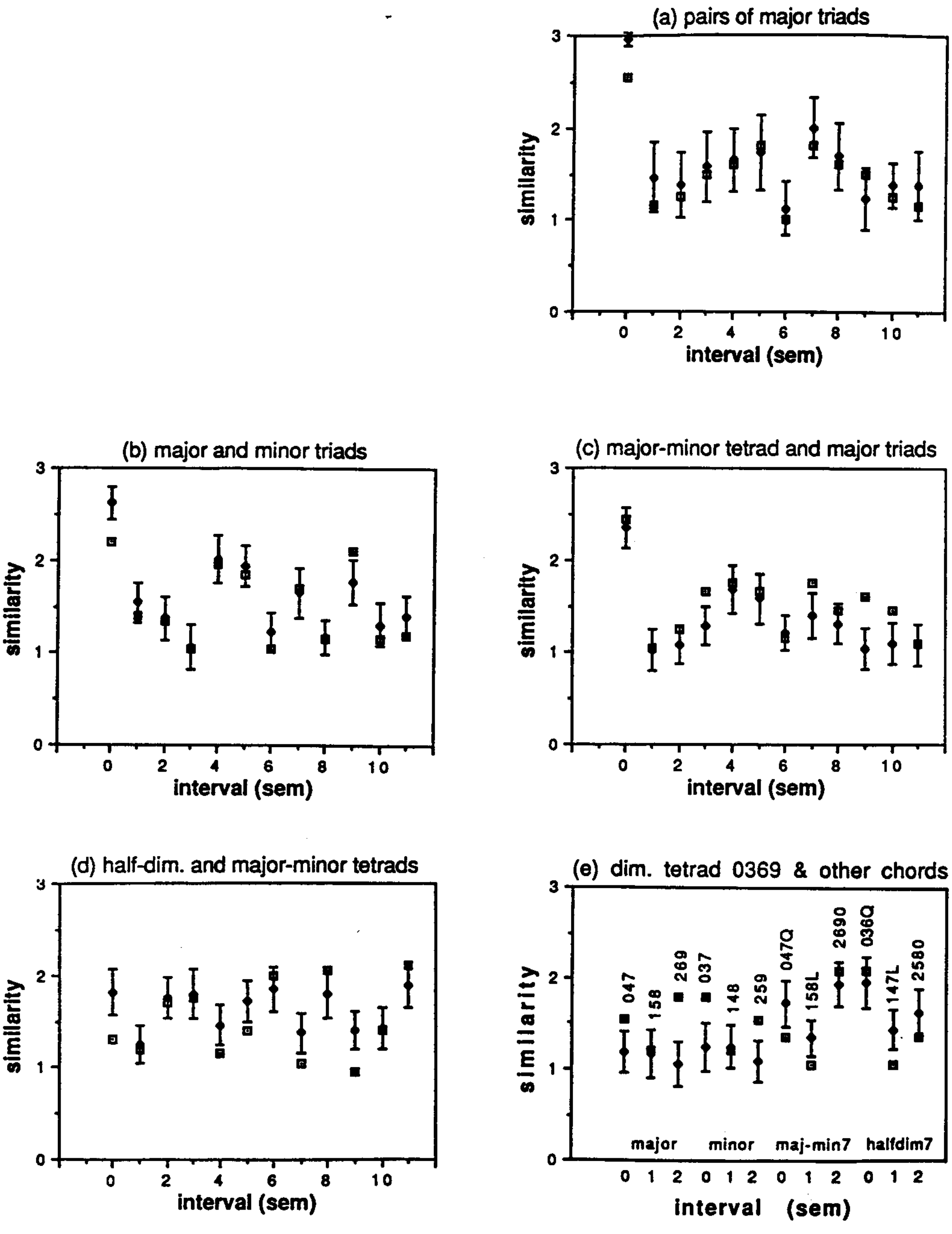

Figure 3 Points: mean responses (26 data per point in part **a**, 52 in parts **b**–**e**). Bars: 95% confidence intervals. Squares: correlation coefficients between tone profiles according to (8), with $k_M = 4$ and $k_W = 2.1$, adjusted linearly to have the same mean and standard deviation as mean responses over all 60 values. In part **e**, the comparison chord is 047 (major triad) for the first three points 037 (minor triad) for the second three, 047Q (major-minor tetrad) for the third three, and 036Q (half-diminished tetrad) for the last three.

Figure 3a contains the only trial in which chords were identical. Almost all listeners responded "very similar" to this trial. Next in the rank order to mean responses were pairs of major triads (047) with intervals of 5 and 8 semitones between their roots. Surprisingly, the rising fifth (7) was judged more similar than the musically more usual rising fourth (5) (cf. Bharucha and Krumhansl, 1983). Next in line were third relationships (3, 4, 8 – but, surprisingly, not 9 – semitones). Like fifth relationships, these involve one common tone (for example, the tone 7 is common to 047 and 37Q). The low response at the tritone interval (6) is in line with music theory.

Part (b) of the figure shows results for pairs of major (047) and minor (037) triads. The "tonic minor/major" relationship (interval 0) was judged most similar. Next comes interval 4 – the mediant relationship in the major key, or the submediant in the minor (both called *Gegenklang* by Riemann, 1893; De la Motte, 1976) – and the familiar dominant-tonic cadence in a minor key (interval 5). Only after these comes the relative major/minor relationship (9), together with the dominant minor or subdominant major relationship (5). The response for the unusual progression 047–148 (interval 1) was presumably enhanced by the presence of a common tone (4).

The results in Figure 3c are headed by interval 0 (for the almost identical chords 047Q and 047). Next, surprisingly, comes interval 4 (the "German augmented sixth" relationship in the major key), along with the familiar dominant-tonic (interval 5). After this comes interval 7 (cf. the progression II – V^7), intervals 8 (cf. V^7 – III in the minor key), and interval 3 (e.g. I – VI^7).

The results in Figure 3d are less structured, due to the relatively high pitch ambiguity of the half-diminished tetrad 036Q (allowing for many equally satisfactory resolutions or preparations). Similarity with the major–minor tetrad 047Q is highest for intervals 0, 2, 3, 6, 8 and 11 semitones between the nominal roots. Of these, the progression at interval 2 is unusual, apparently because it involves parallel fifths; interval 11 corresponds to the resolution of the *Tristan* chord (from the opening of Wagner's *Tristan und Isolde*); and the other progressions tend in musical contexts to sound like chromatic shifts over the same root. The most functionally important resolution of the half-diminished tetrad, ii^7 – V^7 in a minor key (corresponding to interval 5 in the figure), is next in the rank order to similarity.

Figure 3e shows resolutions of (or preparation for) the diminished tetrad 0369. For the major and minor triads (the first six points), similarity is higher for interval 0 (e.g. 0369–047) due to pitches in common, and for interval 1 (e.g. 0369–158) which is functionally important due to its good voice-leading and V–I implication. Overall, however, similarity is low at these six points due to the difference in *timbre* (or consonance) between the major/minor triads and the diminished tetrad. In the case of the major-minor tetrad 047Q, similarity is highest for interval 2 (0369—2690: cf. the resolution of a minor ninth interval to an octave over a dominant seventh chord) due to pitches in common. The high result for the half-diminished tetrad 036Q at interval 0 is also due to common pitches.

Model

Results were simulated by a model of pitch salience in chords of octave-spaced tones. The model is based on the model for the root of a musical chord of Parncutt (1988). The model may also be regarded as an octave-generalized version of the model of pitch salience and tonal relationship of Parncutt (1989). These models are in turn based on Terhardt's (1982) model for the root of a chord, and his algorithm for the pitch and pitch salience of complex tonal signals (Terhardt *et al.* 1982b).

In Parncutt (1988), I aimed primarily to predict the root of a chord class (i.e. a chord built from full harmonic complex tones, which may appear in different inversions, spacings and doublings). The present model aims instead to simulate the results of the above experiments on pitch properties of octave-spaced tones. The experiments suggested that pitch salience in chords of octave-spaced tones is similar to, but not exactly the same as, musical chord-root salience (consider, for example, the responses at intervals 0 and 7 in Figure 2, parts a and b). Accordingly, the present model differs from its predecessor in a number of way. It accounts for masking, places less emphasis on the recognition of harmonic pitch-patterns, and contains some free parameters.

Input

As the stimuli were realized from octave-spaced tones, it is appropriate to restrict the model to a single octave register. This register may be supposed to lie in the most important region of pitch perception – say, 500 to 1000 Hz (Fletcher and Galt, 1950), or register 5. Within this register, chroma (pitch classes) c take values 0, 1, 2, . . . 9, Q and L, corresponding to the musical note names C, C#/Db, D, ... A, Bb/A#, and B.

Pure tone components are assumed to have SPLs of 50 dB. The sounds in the experiment were not as quiet as this may suggest, as they included pure tone components in many critical bands, covering almost the entire range of hearing (see Zwicker, Flottorp, and Stevens, 1957). Note also that the predictions of the model are almost independent of input level over quite a wide range of levels.

Masking and audibility

Masking is accounted for in the model between all pure tone components less than one octave apart. Each octave-spaced tone masks every pure tone component of every other octave-spaced tone from both sides. The extent to which a component at chroma c' masks another component c is expressed in terms of the effective reduction (in dB) of the level at c due to c':

$$\mathrm{ml} = 50 - k_M \,\mathrm{mod}_{12}\,(c - c') \text{ from one side,} \tag{1a}$$

and $$\mathrm{ml} = 50 - k_M \,\mathrm{mod}_{12}\,(c' - c') \text{ from the other.} \tag{1b}$$

Here, "mod_{12}" means "modulo 12" (clock arithmetic). The "masking parameter" k_M is the first free parameter in the model. Its value is expected to lie in the vicinity of 9 dB per semitone, i.e., 27 dB per critical band (Zwicker and Feldtkeller, 1967), given that critical bandwidths is approximately equal to 3 semitones in the region above 500 Hz.[6]

Contributions to masking of component c by component(s) c' are combined by adding amplitudes:

$$\mathrm{ML}(c) = 20 \log_{10} \Sigma\, 10^{\mathrm{ml}/20}, \tag{2}$$

where the summation is carried out over all values of c not equal to c'.

The audible level AL (in dB) of each pure tone component – its level above masked threshold – is then simply:

$$AL(c) = \max\{50 - ML(c), 0\}, \tag{3}$$

where the "max" function prevents negative values of AL.

The audibility A_p of each pure tone component is assumed to saturate with increasing audible level:

$$A_p(c) = 1 - \exp[-AL(c)/15]. \tag{4}$$

Harmonic pitch pattern recognition

The recognition of harmonic pitch patterns among pure tone sensations is simulated by means of a template of root-support weights $w(i)$, where i stands for "interval" and ranges from 0 to 11 semitones. In Parncutt (1988), these weights were estimated at 1, 0, 1/5, 1/10, 1/3, 0, 0, 1/2, 0, 0, 1/4, 0, for intervals 0, 1, 2, . . ., 11. Here, two changes are made to these values. First, the value at interval 3 is set at zero (instead of 1/10), as interval 3 is not among the first 10 harmonics (0, 0, 7, 0, 4, 7, 10, 0, 2, 4 semitones).[7] Second, the weights at the remaining root-support intervals 2, 4, 7 and 10 are adjusted to optimize the fit between calculations and the results of the above experiments. First, they are treated as independent free parameters (Table 1 below). Then, the weights listed above are all raised to the power k_w (the "harmonic weight parameter"): the second free parameter in the model (Table 2).

The effect of harmonic pattern recognition on the overall audibility A of tone components is accounted for by the following template matching procedure:

$$A(c) = \sum_i w(i)\, A_p(\text{mod}_{12}[c + i]). \tag{5}$$

Multiplicity and salience

A first estimate of the number of tones heard in a simultaneity of octave-spaced tones – its multiplicity – is:

$$M' = \sum_c A(c)/A_{max}, \tag{6}$$

where A_{max} is the maximum value of A, i.e. the audibility of the most audible tone component. M' is scaled in the model by raising it to a power less than one:

$$M = (M')^{k_s}, \tag{7}$$

where the "simultaneity perception parameter" k_s is the third free parameter in the model.

The perceptual salience S of each tone sensation is defined as its probability of being noticed. If S is proportional to the audibility A, and the sum of the saliences S of all tone sensations in a sound equals its multiplicity M, then:

$$S(c) = \frac{A(c)}{A_{max}} \cdot \frac{M}{M'}$$

A graph of calculated pitch salience against chroma (for c = 0 to 11) is called a calculated tone profile (cf. Krumhansl and Kessler, 1982).

Modelling of experimental data

The mean responses for the 20 sounds in Experiment 1 (multiplicity), the 60 chord-tone pairs in Experiment 2 (pitch analysis) and the 60 chord pairs in Experiment 3 (similarity) were compared with calculated responses according to the model. Calculated values for Experiment 1 were the multiplicities M (equation 7); for Experiment 2, the pitch saliences S (8); and for Experiment 3, correlation coefficients r between calculated tone profiles of pairs of chords.

In Experiment 1, the values of the free parameters were independently adjusted by small steps until the root-mean-square difference between the 40 experimental and 40 theoretical values reached a minimum. In the other two experiments, parameters were adjusted until the correlation coefficient between 60 experimental and 60 theoretical values was a maximum.[8] The entire adjustment and minimizing procedure was performed automatically by computer.

First six parameters were varied: the masking parameter k_M (equation 1), the weights w (i) for i = 2, 4, 7 and 10, and the simultaneity perception parameter k_S (equation 7). Results are shown in Table 1.

Table 1 Optimal values of six free parameters

Experiment	k_M	w (2)	w (4)	w (7)	w (10)	k_S	r
1 (Multiplicity)	51	.00	.00	.04	.05	.91	.98
2 (Pitch Analysis)	7	.11	.00	.30	.07	.8	.89
3 (Similarity)	4	.08	.18	.25	.00	–a	.77

a. No value is given for k_S in Experiment 3. This parameter only affects the scaling of pitch saliences, so it has no influence on correlation coefficients between tone profiles of chords, i.e. on predicted chord similarities.

As shown in the Table, optimal values of the weights w in the different experiments were unstable, and did not permit improvement on previous, theoretical estimates (Parncutt, 1988). So the previous estimates were retained, and subsequently only three parameters (k_M, k_W and k_S) were varied (Table 2).

Table 2 Optimal values of three free parameters

Experiment	k_M	k_M	k_S	r
1 (Multiplicity)	45	3.9	.91	.98
2 (Pitch Analysis)	8	2.2	.9	.88
3 (Similarity)	4	2.1	—	.76
"typical"	6	2	1	

The impossibly high values of k_M and k_W for Experiment 1 (in both tables) support the contention that listeners estimated numbers of tones by recognizing whole sounds, guessing on the basis of musical experience, rather than by counting tone sensations (as assumed in the model). Values of k_M for the other two experiments are more plausible: assuming a critical bandwidth (cb) of 3 semitones in the dominance region of spectral pitch perception, they correspond to gradients in the range 12–24 dB/cb. A "typical" value of 6 (i.e. 18 dB/cb) is proposed for music-theoretic applications (see footnote 6).

Values of k_W for Experiments 2 and 3 are high by comparison to the effective value of 1 in Parncutt (1988). This suggests an essential difference between the salience (probability of noticing) of a tone in a chord of octave-spaced tones, and the probability that a tone will function as the root of a chord class in music theory.

Values of k_S in Experiments 1 and 2 are also high, compared to the value of 0.5 proposed in Parncutt (1989). This presumably compensates for the octave-generalized model's failure to consider all the tones in a sound. The model, in effect, only looks at one octave register, neglecting the possibility that tones in other octaves might also be noticed.

Conclusion

The model was found to simulate the results of the experiments quite accurately, in spite of the listeners' relatively low confidence in their responses, and the unavoidable influences of cultural conditioning. This confirms the validity of the model in music-theoretical applications, as described in Parncutt (1988).

The model explains important aspects of octave-generalized harmony theory such as the root of a chord, chord-scale compatibility, and harmonic relationship, in terms of a psychoacoustical theory of pitch perception. As described in Parncutt (1989), the model may be used to simulate – quite accurately – the experimental key profiles of Krumhansl and Kessler (1982), by summing tone profiles of individual chords. This confirms Butler's (1989) claim that these key profiles, at least in the case of chord progressions, are largely artifacts of effects of short-term memory: they may be essentially "sensory" rather than "cognitive". Note, however, that the model does not directly explain tone profiles obtained from scales (Krumhansl and Shepard, 1979) or other tone sequences (e.g., Cuddy and Badertscher, 1987). Tone profiles in these cases appear to be a result of cultural conditioning by exposure to music containing chords (including broken chords) and chord sequences in major and minor keys. These latter profiles may be described as *indirectly* sensory, or "sensory in origin" (Parncutt, 1989).

The model may be relevant for the automated performance of music: it provides a possible psychoacoustic basis for the concepts *melodic charge* and *harmonic charge* in Sundberg's (1988) performance rules. In stochastic composition, the model could be used to compose chord progressions where the probability of occurrence a particular chord depends in some consistent fashion on its multiplicity, pitch commonality with previous and following chords, or pitch commonality with a nominal tonic chord, or some combination of these.

Like Parncutt (1989), the present study addresses the issue of whether harmonic relationships in music are sensory or cognitive in nature. According to Shepard (1982, p. 346), there is:

> a fundamental limitation inherent in any purely unidimensional representation of pitch... there is no way in which such a unidimensional scale can represent the fact that under appropriately musical conditions, two tones separated by an especially significant interval, such as the octave..., are perceived to be more closely related than two tones separated by a slightly smaller interval, such as the major seventh. In order to accommodate an increase in similarity between all tones separated by a particular interval, the rectilinear scale must be deformed into some more complex structure requiring... a higher-dimensional embedding space.

The present research argues against Shepard's claims, in that most of the results could be accounted for by means of a sensory model based on a one-dimensional pitch scale. It was not necessary to invoke cognitive-structural representations. This casts some doubt on the existence of such structures.

Later in the same article (p. 350), Shepard makes the important point that:

> considerations of the sensory limitations of the input transducer – such as the reduced efficiency of the ear in discriminating nearby amusical pitches at the low end of the continuum of audible frequencies, which such psychophysical scales of pitch as the mel scale had been designed to represent – are essentially irrelevant to the problem of the representation of the cognitive structures that underlie the interpretation of musical sequences.

In other words, the mel scale has little or nothing to do with musical pitch relationships. In the terminology of the present study, musical pitch relationship depend instead on pitch commonality and familiarity. Both of these involve information processing in some way, and so may be regarded as cognitive, as well as sensory, in nature.

Acknowledgments

I am grateful to Johan Sundberg for the use of his laboratory and equipment, for friendly support, and for comments on the manuscript. Special thanks also to Sten Termström for calculating the waveforms of the sounds, and for general technical assistance. This research was supported by a scholarship from the Swedish Institute. Helpful suggestions from an anonymous reviewer are gratefully acknowledged.

Notes

1. Other possible terms for multiplicity are numerosity, ambiguity, and complexity.
2. Shepard's (1964) tones had a bell-shaped amplitude envelope, tailing off at low and high frequencies. The tones used here had flat amplitude envelopes across most of the audible spectrum (rounded by the frequency response of the loudspeaker). In spite of this difference, they sounded practically identical to Shepard's tones (Pollack, 1978). The lowest-pitched components of the tones used here were inaudible due to masking, and both the highest and the lowest audible components were irrelevant for pitch perception, in particular for the formation of virtual pitch (complex tone sensation), due the dominance effect in spectral pitch perception (Fletcher and Galt, 1950; Terhardt, et al., 1982b). The main pitch of both Shepard's tones and the tones used here lies in the vicinity of 300 Hz (Terhardt et al., 1986). A difference between the sound of the two kinds of tone did, however, become apparent when tones were superposed to form chords: chords composed of the tones used here sounded rougher, as they had more low-ptiched components. This presumably had no effect on their pitch salience patterns, which were determined mainly by pure tone components in the dominance region (near 700 Hz).
3. *Note on terminology*. In this paper, chords are specified in terms of the number of semitones between the music-theoretical root (or other reference pitch) and the other notes (Parncutt, 1990). So, for example, a major triad is written "047", where "0" denotes the root, "4" the major third and "7" the perfect fifth. Intervals of 10 or 11 semitones above the root are specified by the symbols Q and L respectively (obtained by superimposing the component symbols 0 and 1 of these numbers; L also stands for "leading note"). So the dominant seventh chord is called 047Q. Confusion between diatonic (traditional) and chromatic interval labels (e.g. "seventh" versus "7 semitones") is avoided by consistent use of the terms *monad, dyad, triad* and *tetrad* for single tones, (simultaneous) intervals, chords of three tones, and chords of four tones, respectively. The minor seventh chord 037Q, for example, is called a *minor tetrad*, and 0369 is called a *diminished tetrad*. Incidentally, the term "dominant tetrad" is reserved for

047Q chords actually on the dominant scale degree (7 semitones above the tonic); when tonal context is not specified, the 047Q chord is called a *major-minor tetrad* (cf. major-minor seventh). Confusion with conventional interval names (third, fifth, etc.) is further avoided by the formulation "interval 0" for unison, "interval 1" for one semitone etc.

4. To a good approximation, the mean responses for two different trials are significantly different ($p > 0.05$) if they differ by more than a 95% confidence deviation (i.e. the half-width of a 95% confidence interval) divided by root 2, provided the confidence deviation for the two experiments is about the same.
5. In a different context (Parncutt, 1989), single octave-spaced tones were mostly heard to comprise between two and three tones. This is an example of the general rule that the perceived number of tones in a musical sound depends on context.
6. At medium to high sound levels, the masking pattern of a pure tone is considerably higher and longer on its upper than its lower side. However, at low levels the pattern is almost symmetrical with respect to critical-band rate (Zwicker and Jaroszewski, 1982).
7. In Parncutt (188), interval 3 was assigned a small root-support weight (1/10), as the third harmonic of interval 3 (3 + 7 = 10) corresponds to a harmonic (the seventh) of interval 0. It may therefore contribute to the root sensation. Here, octave-spaced tones were used, containing only harmonics 1, 2, 4, 8, 16, etc., so interval 3 could not contribute to the root – except, perhaps, by cultural conditioning.
8. Note that each parameter may be regarded as a measure of how analytically sound is perceived: k_M at the level of spectral analysis (discrimination of pure tone components), k_W at the level of "hearing out" of pure (as opposed to complex) tone components, and k_S at the level of simultaneous perception of tones in a sound.

References

Bharucha, J. & Krumhansl, C.L. (1983) The representation of harmonic structure in music: Hierarchies of stability as a function of context. *Cognition*, **13**, 63–102.

Burns, E.M. (1981) Circularity in relative pitch judgments for inharmonic complex tones: The Shepard demonstration revisited, again. *Perception & Psychophysics*, **30**, 467–472.

Butler, D. (1989) Describing the perception of tonality in music: A critique of the tonal hierarchy theory and a proposal for a theory of intervallic rivalry. *Music Perception*, **6**, 219–242.

Cuddy, L.L. & Badertscher, B. (1987) Recovery of the tonal hierarchy: Some comparisons across age and musical experience. *Perception and Psychophysics*, **41**, 609–620.

de la Motte, D. (1976) *Harmonielehre*. Kassel: Bärenreiter.

Deutsch, D. (1987) The tritone paradox: Effects of spectral variables. *Perception & Psychophysics*, **41**, 563–575.

DeWitt, L.A. & Crowder, R.G. (1987) Tonal fusion of consonant musical intervals: The Oomph in Stumpf. *Perception & Psychophysics*, **41**, 73–84.

Fletcher, H. & Galt, R.H. (1950) The perception of speech and its relation to telephony. *Journal of the Acoustical Society of America*, **22**, 89–151.

Huron, D. (1989) Voice denumerability in polyphonic music of homogeneous timbres. *Music Perception*, **6**, 361–382.

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

Krumhansl, C.L. & Shepard, R.N. (1979) Quantification of the hierarchy of tonal functions within a diatonic context. *Journal of Experimental Psychology: Human Perception and Performance*, **5**, 579–594.

Ohgushi, K. (1983) The origin of tonality and a possible explanation of the octave enlargement phenomenon. *Journal of the Acoustical Society of America*, **73**, 1694–100.

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

Parncutt, R. (1989) *Harmony: A Psychoacoustical Approach*. Berlin: Springer-Verlag.

Parncutt, R. (1990) Chromatic chord symbols. *Computer Music Journal*, **14** (2), 13–14.

Pollack, I. (1987) Decoupling of auditory pitch and stimulus frequency. The Shepard demonstration revisited. *Journal of the Acoustical Society of America*, **63**, 202–206.

Rasch, R.A. (1978) The perception of simultaneous notes such as in polyphonic music. *Acustica*, **40**, 21–33.

Riemann, H. (1893) *Vereinfachte Harmonielehre*. London: Augener.

Shepard, R.N. (1964) Circularity in judgments of relative pitch. *Journal of the Acoustical Society of America*, **36**, 2346–2353.

Shepard, R.N. (1982) Structural representations of musical pitch. In D. Deutsch (Ed.), *The Psychology of Music*, pp. 344–390. London: Academic Press.

Sundberg, J. (1988) Computer synthesis of music performance. In J.A. Sloboda (Ed.), *Generative Processes in Music*. Oxford: Clarendon.

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

Terhardt, E., Stoll, G., Schermbach, R., & Parncutt, R. (1986) Tonhöhenmehrdeutigkeit, Tonverwandschaft und Identifikation von Sukzessivintervallen. *Acustica*, **61**, 57–66.

Terhardt, E., Stoll, G., & Seewann, M. (1982) Pitch of complex tonal signals according to virtual pitch theory: Tests, examples and predictions. *Journal of the Acousticacl Society of America*, **71**, 671–678 (a).

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

Terhardt, E., Stoll, G., Schermbach, R., & Parncutt, R. (1986) Tonhöhenmehrdeutigkeit, Tonverwandschaft und Identifikation von Sukzessivintervallen. *Acustica*, **61**, 57–66.

Thurlow, W.R. & Rawling, L.L. (1959) Discrimination of number of simultaneously sounding tones. *Journal of the Acoustical Society of America*, **31**, 1332–1336.

Zwicker, E. & Feldtkeller, R. (1967) *Das Ohr als Nachrichtenempfänger*, 2nd ed. Stuttgart: Hirzel-Verlag.

Zwicker, E., Flottorp, G., & Stevens, S.S. (1957) Critical bandwidth in loudness summation. *Journal of the Acoustical Society of America*, **29**, 548–557.

Zwicker, E. & Jaroszewski, A. (1982) Inverse frequency dependence of simultaneous tone-on-tone masking patterns at low levels. *Journal of the Acoustical Society of America*, **71**, 1508–1512.