Harmony Perception: Harmoniousness is More Than the Sum of Interval Consonance

Norman D. Cook
Kansai University, Takatsuki, Osaka, Japan

ATTEMPTS TO EXPLAIN HARMONY PERCEPTION SINCE Helmholtz (1877/1954) have relied primarily on psychoacoustical models of the dissonance among the partials of chord tones. Those models are successful in explaining interval perception and the interval structure of common scales, but do not account for even the basics of triadic harmony. By introducing a 3-tone "tension" factor, I show how the sonority of the triads of diatonic music can be explained. Moreover, the relative size of the intervals among the partials in triads determines the major/minor modality of chords: major chords have a predominance of larger lower intervals, while minor chords have a predominance of smaller lower intervals. Finally, by invoking the "frequency code" known from linguistics and ethology, the positive/negative valence of the major/minor chords is shown to have an acoustical basis. I conclude that the perception of harmony can be explained by the acoustical structure of triads, without invoking cultural factors.

Received April 2, 2008, accepted February 18, 2009.

Key words: harmony perception, major, minor, tension, dissonance

THE USAGE OF BOTH PITCH INTERVALS AND pitch triads (played either simultaneously as chords or sequentially as melodies) differs widely among various musical cultures, but the empirical facts concerning their common perception are well-established both from laboratory experiments and from statistics on their prevalence in Western classical and popular music. Among the triads, major and minor chords are perceived as relatively "pleasant," "consonant" and "beautiful," and they totally dominate most forms of popular music, globally. Other triads are of course employed, but most often in support of and in transition to major or minor key resolution.

Although the psychoacoustics underlying the perception of interval dyads is relatively well-understood, there is a notable absence of quantitative hypotheses that explain why the triads are perceived as they are or why the laws of traditional harmony theory work so well. Here, I show that straightforward considerations of the acoustical properties of chords can explain many of the outstanding questions about diatonic music ("diatonic" meaning the major and minor scales of Western music and the harmonic phenomena implied by those scales), provided that 3-tone configurations are brought into the picture.

As already understood by Helmholtz (1877/1954) and his contemporaries, the science of musical pitch begins with the wave structure of tones (e.g., Cook, 1999; Pierce, 1992). Because of the presence of upper partials (overtones or higher harmonics), chords, intervals and even isolated tones are phenomena of some acoustical complexity. This one technical detail about the upper partials is often unfamiliar to music listeners, but an understanding of the partial structure of tones adds a dimension that ultimately leads to a deeper appreciation of the phenomena of harmony. The crucial point is simply that what is heard as a "single tone" actually consists of a fundamental frequency (F0, defined in terms of cycles per second, or Hertz) and higher harmonic components (F1, F2, F3 and so on, described as sine waves that are integral multiples of the F0, with the partials usually having gradually weaker amplitudes). Eventually, the upper partials become so weak that they can be ignored, but discussions of musical tones usually consider the first 4-5 partials. Most musical sounds in the real world are composed of a set of these partials, which together determine their musical timbre.

Previous psychoacoustical work has considered the effects of upper partials, but has focused narrowly on interval dissonance, i.e., the relative separation of partial dyads. Such "sensory dissonance" is normally distinguished from other kinds of unsettled, "musically dissonant" pitch effects that are presumably due to factors other than the raw acoustical signal. Although the dissonance models have been largely successful in explaining the perception of intervals themselves and in providing a coherent, acoustical basis for understanding the interval structure of the diatonic scales

Music Perception VOLUME 27, ISSUE 1, PP. 25–41, ISSN 0730-7829, ELECTRONIC ISSN 1533-8312 DOI:10.1525/MP.2009.27.1.25

TABLE 1. Empirical Data and Theoretical Predictions on the Rank Ordering of the Sonority of the Common Triads.

		Empirical Sonority			Theoretical Sonority Based on Interval Consonance					Present Model
Chord Class	Interval Structure	Roberts (1986)	Eberlein (1994)	CFK (2007)	Helmholtz (1877/1954)	P&L (1965)	K&K (1969)	Parncutt (1989)	Sethares (1999)	C&F (2006)
	4-3	1	1	1	3	4	1	1	4	1
Major	3-5	3	3	4	**9**	**11**	**11**	6	**8**	5
	5-4	2	7	2	1	2	6	3	2	4
	3-4	4	2	3	3	4	1	4	4	2
Minor	4-5	6	4	5	1	2	6	6	2	3
	5-3	5	6	6	**9**	**11**	**11**	**10**	**8**	6
	3-3	7	9	7	13	13	6	9	12	12
Dim	3-6	8	5	12	11	8	9	5	10	7
	6-3	9	10	10	11	8	9	8	10	10
Aug	4-4	10	11	13	**5**	10	13	**2**	12	13
	5-2	—	8	9	**5**	**6**	**1**	—	**6**	8
Sus4	2-5	—	11	8	**5**	**6**	**1**	—	**6**	11
	5-5	—	11	11	**8**	**1**	**1**	—	**1**	9

Note: CFK: Cook, Fujisawa, & Konaka, 2007; C&F: Cook & Fujisawa, 2006; K&K: Kameoka & Kuriyagawa, 1969; P&L: Plomp & Levelt, 1965.

(and their pentatonic subsets), the successes of the dissonance models do not generalize even to the triads. This fact can be seen in Table 1, which summarizes the total consonance ("sonority") of the most common triads, as calculated from various well-known dissonance models.

Table 1 also shows two types of empirical evidence. The first is from laboratory experiments (Roberts, 1986) in which musicians and nonmusicians evaluated the "consonance" of the common triads in root and inverted positions when presented as isolated chords. For both groups of participants, the sequence was, unsurprisingly: major > minor > diminished > augmented (suspended fourth triads and other chords containing a semitone or whole-tone dissonance were not tested). We have reported similar experimental results that include the suspended fourth chords (Cook, Fujisawa, & Konaka, 2007). The second type of empirical data concerns the prevalence of harmonic triads in a large sample of chords from classical music (Bach, Handel, Mozart, Beethoven and Mendelssohn), as collected by Eberlein (1994). It was found that the relative prevalence of chord types is the same as the sequence of consonance obtained in laboratory experiments (with some variations among the triad inversions), suggesting that the frequency of usage can be taken as a proxy for the pleasantness/sonority of these chords. Again, there is no surprise in the empirical data, which conform to musical "common sense" about the sonority of the triads.

Unfortunately, Table 1 indicates that the theoretical models used to explain the relative consonance of the intervals of diatonic scales produce results concerning the total sonority of triads that are dramatically inconsistent with experimental results. Although the interval consonance models have sometimes been heralded as providing the scientific basis for an explanation of "Western harmony" (Terhardt, 1974; Parncutt, 1989; Tramo, Cariani, Delgutte, & Braida, 2001), calculation of the total consonance (≈"sonority," "tonality") of triads directly from the consonance curves produces anomalous values that are inconsistent with both laboratory measures of perceived consonance and empirical usage. It should be noted that the interval dissonance models were designed to explain interval perception itself, and not presumably to explain triadic sonority. In fact, with the exception of Parncutt (1989), advocates of the interval dissonance models have not even reported the triadic implications of their models. Nonetheless, it is significant that the various parameters of these models that provided the best results concerning interval perception predict that either the augmented or the suspended fourth triads are more consonant than some or all of the major or minor triads—contrary to empirical facts. While there are many other musical phenomena that a successful model of harmony perception might explain, replication of the empirical findings on triad perception constitutes a minimal threshold for theoretical models to cross in order to achieve *prima facie* plausibility.

A Psychoacoustical Model of Harmonic Instability

Experimentally, what is found in studies of interval perception is that normal listeners hear a mildly "unpleasant," "unsettled" dissonance at small intervals of 1-2 semitones and at larger intervals of 10, 11, or 13 semitones. Moreover, dissonance is consistently reported for an interval of 6 semitones (the tritone) (e.g., Cook, 2002; Kameoka & Kuriyagawa, 1969; Plomp & Levelt, 1965). To account for such experimental results, interval effects have been modeled many times since Helmholtz (1877/1954). Following the work of Sethares (1999), the relative dissonance (*D*) of any two tones can be defined most simply as:

$$D = v \cdot \beta_3 \left[\exp(-\beta_1 x) - \exp(-\beta_2 x) \right] \qquad \text{(Eq. 1)}$$

where v is the product of the relative amplitudes of the two tones (set to values of normalized amplitudes between 0.0 and 1.0), x is the interval size, defined as $x = \log(f_2/f_1)$, and the parameters are β_1 (–0.8) β_2 (–1.6), and β_3 (4.0). Parameter β_1 specifies the interval of maximal dissonance (0.8 being just less than one semitone) and parameter β_2 specifies the steepness of the fall from maximal dissonance. The variables f_1 and f_2 are the frequencies of the two tones (in Hertz). More complex models have been advocated by Plomp & Levelt (1965), Kameoka & Kuriyagawa (1969), Terhardt (1974), Parncutt (1989), and Tramo et al. (2001), but the results are qualitatively similar to those obtained by Sethares (1999). The dissonance model curve is shown in Figure 1A; the total dissonance of two tones plus their upper partials can be obtained by applying Eq. 1 to all pairs of partials and then summing the dissonance (Figure 1B).

The remarkable result of such modeling is that, simply by including the first few upper partials for each of the two F0s, the theoretical curves (Figure 1B) gradually come to resemble the experimental curve. In other words, a very simple model devised to explain the relative dissonance of 1-2 semitone intervals (Plomp & Levelt, 1965; Figure 1A) predicts that intervals of 12 and 7 semitones are the most consonant. These are the two intervals—an octave and a fifth—that are used in virtually every musical tradition and are heard by all normal listeners as consonant, "pleasant" intervals. Somewhat less clearly (and dependent on the parameters of the dissonance model and the relative amplitude of the upper partials), small decreases in dissonance are predicted at (or near to) many of the intervals of the pentatonic and diatonic scales (3, 4, 5, 7, 9 and 12 semitones; Figure 1B). Similarly, the relatively strong dissonance at intervals of 6, 10, 11 and 13 semitones also is explained as a consequence of small interval effects among upper partials. These theoretical results have been justifiably celebrated as a triumph of reductionist science and provide strong support for the psychoacoustical approach advocated by Helmholtz (1877/1954).

It should be noted, however, that the locations of the consonant intervals in Figure 1B do not correspond to any known musical scale: to obtain a scale that is actually used, notes need to be added, subtracted and/or slightly retuned. The fact that physical acoustics does not uniquely produce real scales, much less the Western diatonic scales or the 12-equitempered tones of the piano keyboard, indicates that some form of human manipulation must be included to establish a musical tradition. In this regard, Leonard Bernstein's (1976) famous declaration that the "physical universe" has provided us with diatonic music is clearly stating the case too strongly. The physical universe provides some nonarbitrary, acoustical

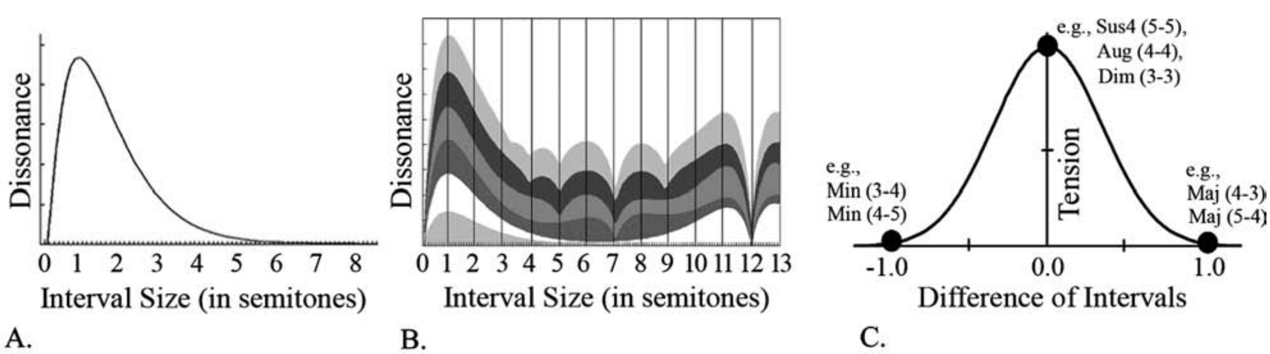

FIGURE 1. (A) The dissonance model curve of Eq. 1. (B) The more complex dissonance curves obtained from calculations including the effects of 1-6 partials. The amplitudes of the partials are assumed to decrease as 1/n. Note the decreases in total dissonance near to many of the intervals of the diatonic scales. (C) The tension model curve of Eq. 2. The effects of upper partials on the tension curve are illustrated on the "triadic grid" (Figure 6).

raw materials, but the human mind can and does assemble the raw materials in various ways.

Despite the successes of the interval dissonance models, the fact that the summation of interval consonance does not accurately predict triadic sonority (Table 1) has led us to include a second factor (for details of the psychoacoustical model, see Cook, 2001, 2002; Cook & Fujisawa, 2006; Cook & Hayashi, 2008; Cook, Fujisawa, & Konaka, 2007). The second factor is what we refer to as the "tension" inherent to any 3-tone combination, as calculated from the relative size of its intervals. Specifically, a tension value (T) is obtained from each triplet combination of partials from the three fundamental tones, as follows:

$$T = v \cdot \exp\left[-\left(\frac{y-x}{\alpha}\right)^2\right] \qquad \text{(Eq. 2)}$$

where v is again the product of the relative amplitudes of the three partials (each set to 0.0-1.0), and α (≈0.6) is a parameter that determines the steepness of the fall from maximal tension; x and y are, respectively, the lower and upper of the two intervals in each tone triplet, defined as $x = \log(f_2/f_1)$ and $y = \log(f_3/f_2)$, where the frequencies of the three partials are $f_1 < f_2 < f_3$ (in Hertz). The theoretical tension curve is as shown in Figure 1C. Similar to application of Eq. 1 for modeling interval perception, application of Eq. 2 to triad perception is made for every triadic combination of partials (see The Tension of Triads section, below).

The proposed tension factor is essentially formalization of the idea of "intervallic equidistance," as described by Leonard Meyer (1956). He noted that: "If the vertical [pitch] organization is undifferentiated as to intervallic distance, then there can be no focal point around which organization can take place . . . [Such chords] have no root and no shape and hence no tendency." (p. 166) and that "The . . . power of chromaticism arises . . . because uniformity of progression, if persistent, tends . . . to create ambiguity and hence affective tension. Moreover, ambiguity leads, particularly in the realm of harmonic progression, to a general tonal instability." (p. 218)

Assuming that "tension" is a consequence of the structure of 3-tone combinations and is distinct from 2-tone interval dissonance, a theoretical value for the overall perceptual "instability" of chords can be obtained if both the dissonance among tone pairs and the tension among tone triplets are added together. That is, the total instability (I) can be defined as the weighted sum of dissonance (Eq. 1) and tension (Eq. 2):

$$I = D + \delta T \qquad \text{(Eq. 3)}$$

where δ (≈0.2) reduces the contribution of the tension component relative to the dissonance component. This parameter means that interval effects in the model are perceptually five-fold stronger than the triadic effects. The total instability scores can be used for the rank ordering of the "sonority" of the triads (Table 1, column C&F). It is seen that there is approximate agreement with the empirical sonority scores. Those numerical results suggest that the resolved/unresolved character of triads ("harmonic stability") has a straightforward psychoacoustical basis that is quite distinct from conventional arguments based solely on the summation of interval dissonance. The musical significance of this model will be discussed below.

Implications of the Model

From what is known about the perception of both isolated tones and intervals, it can be expected that the perception of chords consisting of three or more tones also will be influenced by the number and strength of upper partials. Such effects will be considered in detail below, but the first topic in understanding harmony is the spacing among the F0s themselves. Similar to the phenomena of musical intervals, when a 3-tone chord is sounded, the frequencies with the greatest amplitude are usually those of the three distinct notes that are played with the fingers and that are specified in the musical score. The higher harmonics tag along for free and give the chord a depth and complexity that might be called its overall "sonority" (the effects of which are discussed below).

The set of all possible combinations of three tones can be conveniently represented on a "triadic grid," as shown in Figure 2. The vertical axis represents the lower interval and the horizontal axis represents the upper interval, so that the major chord in root position, for example, has grid position 4-3; inverted major chords are located at positions 3-5 and 5-4. The minor chords are found at grid positions 3-4, 4-5, and 5-3. (Note that different tuning systems—Pythagorean, just, mean-tone, etc.—would imply small shifts of the vertical and horizontal lines of the grid, but will not be discussed here.)

Although still recognizably major or minor, the sonority of the major and minor triads in their various inversions and when played over one or two octaves differs somewhat and their musical usages also differ, so we must take note of their interval substructure. For this purpose, larger grids of the same kind will be used to illustrate the computational results of the psychoacoustical model. As shown below, the merit of mapping acoustical properties onto the triadic grid is that the relative dissonance (tension, instability or modality) of

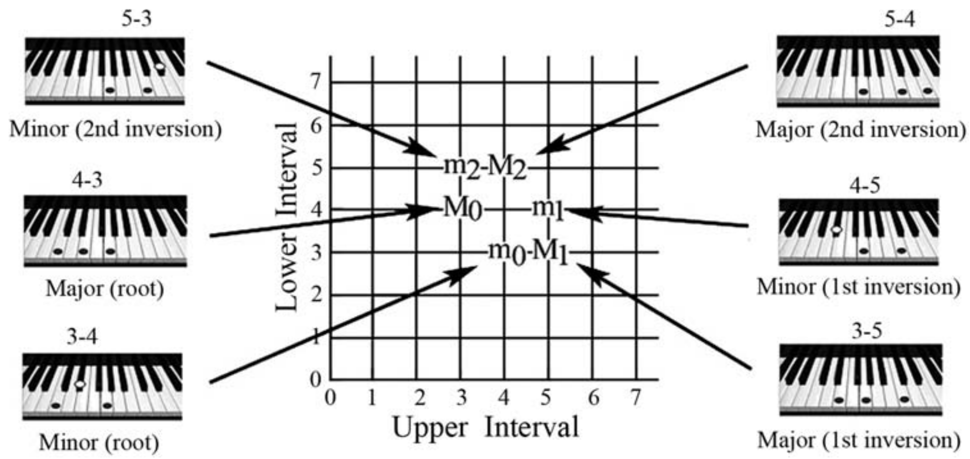

FIGURE 2. A small triadic grid showing the locations of the most familiar major (M_0, M_1, M_2) and minor (m_0, m_1, m_2) triads in their three inversions. The interval substructure of these triads is specified by their location at the intersection of lines indicating intervals of 3, 4 or 5 semitones. Examples of these chords in the key of C and their interval structures are also shown.

a large set of 3-tone chords can be viewed simultaneously. Moreover, changes in such values that occur at intervals of less than one semitone (10 cents in the present calculations) also can be visualized.

Clearly, the triadic grid is not a part of traditional harmony theory. Although traditional theory is arguably a coherent, self-consistent description of harmonic phenomena, it is a conceptual system that is not built on the foundations of acoustical physics. For this reason, terminology from music theory will be avoided in so far as possible in the present discussion and all chords will be described in terms of their acoustical (semitone) substructure. If the properties of triadic harmonies can be explained successfully in terms of psychoacoustics, then, in principle, it should be possible to reconstruct other aspects of traditional harmony theory on that basis. The present essay is intended as a first step in that direction.

Among the many possible triads that can be specified on the grid, the major and minor triads are the chords that provide the harmonic framework for the vast majority of Western classical and popular music; they are used again and again—either as triads or as triads with repetition of the triadic pitches an octave higher or lower. Other grid locations include triads of varying utility and beauty, as well as many chords that are simply avoided in most types of music. Whatever our subjective evaluation of the sonority of these chords, however, the chords themselves have certain structural properties that can be described objectively in terms of the spacing of the F0s and their partial components.

For the present purposes, only those triads that can be played over one or two octaves (as illustrated on grids with the axes running from 0 to 13 semitones, Figure 3) will be considered. Any recognizable triad—and many unrecognizable ones as well—can be specified on the grid as the point of intersection of the vertical and horizontal lines indicating semitone steps. In addition to the major and minor chords, there are of course other chords with established names in harmony theory, labels for which are also shown in Figure 3; it is seen that there are two clusters on the grid where the most common chords lie. In subsequent displays of the triadic grid, psychoacoustical properties will be color-coded and mapped onto the grids using the dimensions of dissonance, tension, instability and modality. Model calculations for dissonance, tension and instability produce non-negative real values, whereas modality can be zero, positive or negative (Figure 3B).

Given the known F0 interval structure of the triads, what can be said about their relative sonority? How can the stability or instability of the triads—their resolved or unresolved character—be explained? Once these questions have been answered in relation to dissonance, tension and instability, it will then be possible to ask how the commonly perceived positive and negative emotional valence of the major and minor chords might also be accounted for on a psychoacoustical basis.

The Dissonance of Triads

Since chords can be considered as the sum of their intervals, the obvious first step in trying to explain their overall sonority is to add up the dissonance of the intervals to obtain a "total dissonance" score. The inclusion of the dissonance among all pairs of upper partials will of course make the calculation of total dissonance

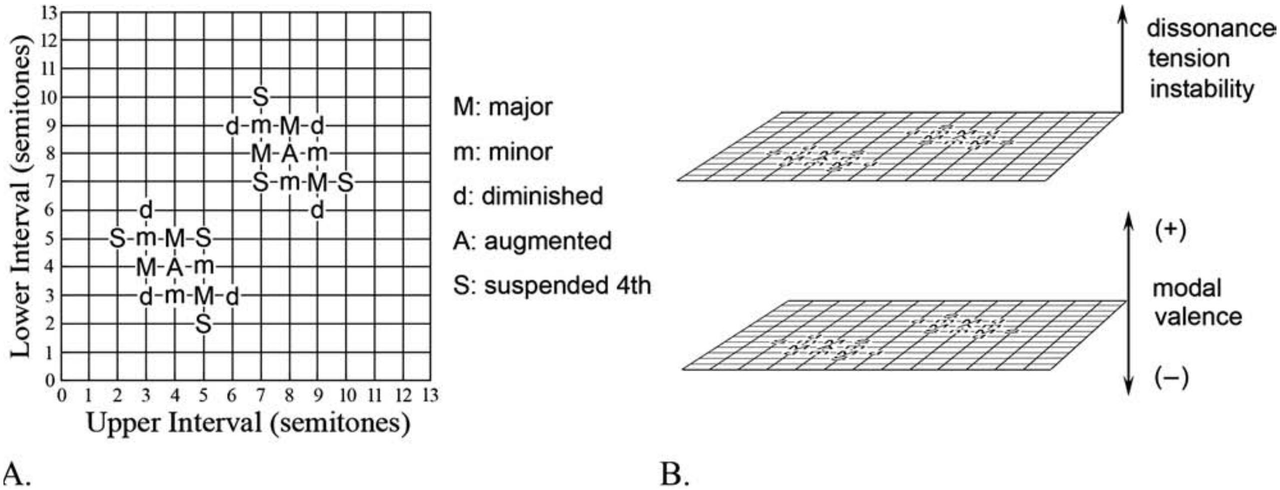

FIGURE 3. (A) A larger triadic grid showing the locations of the major, minor, diminished, augmented, and suspended fourth triads in their various inversions when played over 1-2 octaves. (B) Calculated acoustical features of the triads can then be plotted along a third dimension. The locations of the triads on the grid remain fixed, but the strengths of the acoustical features change with the inclusion of upper partials (see Figures 4, 6, 7 and 9).

somewhat complex, but with or without upper partials the question is essentially: Can interval consonance/dissonance explain the sonority of chords?

Figure 4A illustrates the summed dissonance of the intervals (calculated with the model in Figure 1A) when only the F0s are considered. It can be seen that there are two strips of relatively strong dissonance if either interval is one or two semitones in size. The remainder of the triadic grid is found to be a region of low dissonance. Figure 4B-D shows dissonance maps that include upper partials in the calculations. The fine-structure of the maps gradually gets more complicated as upper partials are added, but the general pattern remains approximately the same. That is, there are red peaks of strong

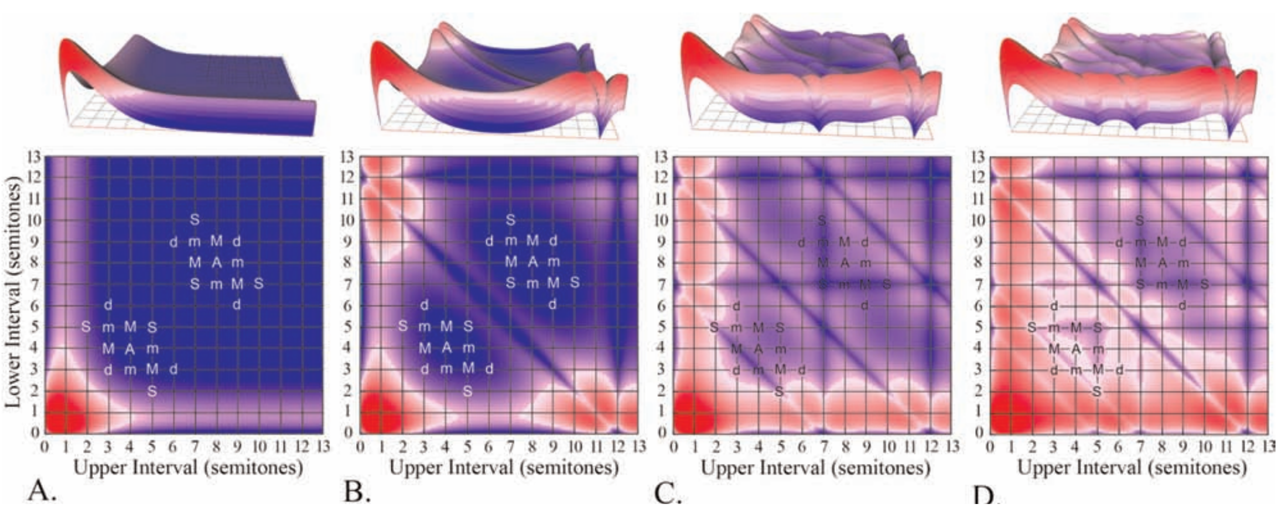

FIGURE 4. The total dissonance for all triads containing intervals of 0-13 semitones. The stipled regions (blue regions, in the online version) have low dissonance, whereas nonstipled regions (red regions online) indicate high dissonance. (A) Results when only the fundamental frequencies (F0s) are included in the calculations. When either interval is less than two semitones, dissonance is present, but this falls off rapidly so that all of the common triads are located in or at the edge of a broad "valley of consonance." (B) The dissonance map obtained when both the F0 and F1 are included in the calculations. The two ridges of dissonance along the x- and y-axes now have more complex structure and the "valley of consonance" is divided into two regions. (C & D) The dissonance maps when F0-F2 and F0-F3 are considered. The valleys of consonance show more complex structure, but remain as regions of relatively low dissonance bordered by mountain ranges of high dissonance. Solely on the basis of "total dissonance," a clear distinction cannot be made between the major and minor chords or between the resolved (M, m) and unresolved (d, A, S) chords. Note that there are several vertical, horizontal and diagonal strips of low dissonance—corresponding to regions where intervals of 12 semitones arise. All of the color-coded triadic grids are screenshots of the "Seeing Harmony" software, available at: www.res.kutc.kansai-u.ac.jp/~cook. (For a color version of this figure, please see the digital PDF at www.musicperception.net)

dissonance (when either of the intervals is small) and blue expanses of relatively strong consonance (where all of the common triads lie). Therefore, although we have good reason to think that small intervals make chords less stable because of dissonance effects, calculation of the total dissonance of triads does not indicate why people easily distinguish between the resolved (major and minor) chords and the unresolved (diminished, augmented, and suspended fourth) chords (e.g., Cook et al., 2007; Roberts, 1986). Judging solely from the total dissonance among the partials, all of the common triads are rather similar (Figure 4).

This "negative result"—i.e., the failure to explain the perceptual differences among the triads on the basis of dissonance and, indeed, the many counterintuitive results summarized in Table 1—has led some commentators to conclude that, in addition to so-called "sensory dissonance," the learning of musical traditions and habituation to the "acceptable" tone combinations in specific musical cultures must be invoked to account for the perceived sonority of the triads. In other words, in addition to the acoustical dissonance perceived by all human beings, there is also "musical dissonance" or "cultural dissonance" (Parncutt, 1989, pp. 56-60) that is culture-dependent. In what has become the conventional view of music perception, psychoacoustics is said to play a role up to and including 2-tone combinations, but more complex pitch phenomena are essentially "learned." (For example, "It may even be that acclimatization to a convention can completely override [the] acoustic facts," Ball, 2008; "Our emotional response to particular scales or chords seems likely to be acquired from exposure to a particular culture," McDermott, 2008; "The objective organization of sounds is only loosely related to how minds interpret those sounds," Huron, 2008; "Scale and harmonic structures depend on learning," Trainor, 2008). Despite the apparent popularity of such a view, however, it is instructive to examine the 3-tone structure of the triads before we conclude that culture trumps acoustics.

The Tension of Triads

The first topic in harmony perception beyond 2-tone effects is the 3-tone configuration of the triads. Whereas the structural question underlying interval perception was simply how close are each pair of tones (partials) to one another, the structural feature of the triads that contributes to their harmonic tension is the symmetry/asymmetry of triplets of partials (Figure 1C). The symmetry (≈"intervallic equidistance") of certain unresolved chords is of course well-known: (i) the diminished chord in root position (3-3 semitone structure), (ii) the augmented chord (4-4), and (iii) the so-called suspended fourth chord in second inversion (5-5). These triads evoke a sense of instability that is perceptually evident to both musicians and nonmusicians (Roberts, 1986), and is consistently found in laboratory experiments with people from the East and West (Cook, 2002). Traditional harmony theory explains their unsettled character in terms of the absence of the interval of a fifth, but Meyer's (1956) alternative explanation in terms of the Gestalt perception of "symmetrical" acoustical structures provides the possibility of quantitative evaluation (Eq. 2).

The idea that intervallic equidistance leads to tension seems to apply to the three symmetrical chords mentioned above, but what about the various inversions of those chords? Except for the augmented chord (which retains the same interval structure in its inversions), inversions of the diminished (3-6 & 6-3) and suspended fourth (2-5 & 5-2) chords—all of which have a notably unsettled, "tense" character—have unequal intervals. So does Meyer's (1956) argument break down already with these simple counterexamples? The answer is that, if the upper partials are brought into consideration, then an abundance of intervallic equidistance is found in all of the unresolved triads (diminished, augmented, and suspended fourth) in all of their inversions, while equal intervals are not found in any of the major and minor triads. Figure 5 illustrates this by showing these triads together with the first set of upper partials.

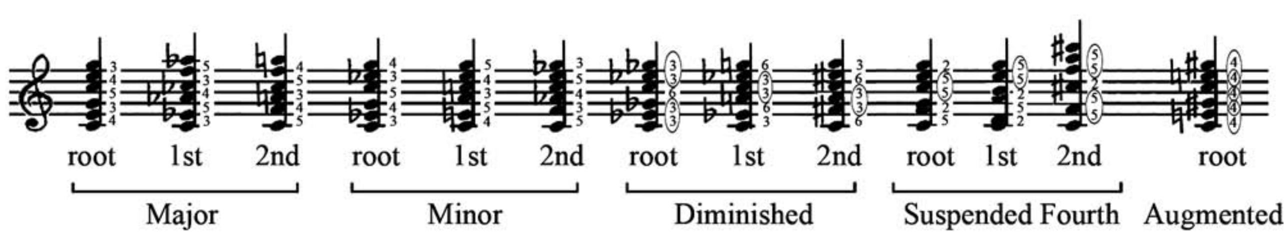

FIGURE 5. The interval substructure of the common triads with the first set of upper partials also shown. Interval sizes in semitones are indicated by small integers. None of the major and minor chords, but all of the diminished, augmented, and suspended fourth chords show repeating intervals of the same size (circled).

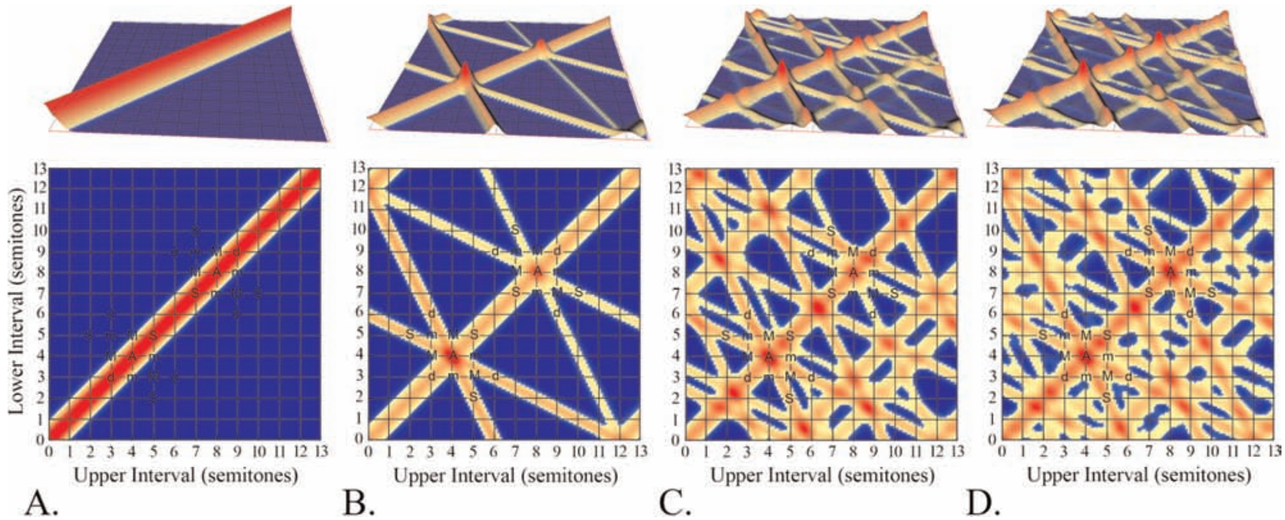

FIGURE 6. The theoretical "tension" of triads. (A) When only the F0 is considered, tension is found along a diagonal strip due to the presence of two intervals of the same size. The remainder of the grid corresponds to regions of low tension (stipled and red in the online version). This effect can explain the unsettled character of the augmented chord (A), one of the diminished chords, and one of the suspended fourth chords, but leaves the perceptual tension of other unresolved triads (d, S) unexplained. (B) The theoretical "tension" of triads when both F0 and F1 effects are considered. The unsettled tensionof all inversions of the diminished, augmented, and suspended fourth triads is predicted (unstipled or red regions). The augmented chords at both locations show the highest tension. (C & D) More complex tension maps obtained when F0-F2 and F0-F3 are included in the calculations. All of the tension chords (d, A, S) are located on the tension strips, while the major and minor triads (M, m) lie in regions of relatively low tension (blue or stipled). (For a color version of this figure, please see the digital PDF at www.musicperception.net)

Depiction of the first set of upper partials of the triads in musical notation already distinguishes between the resolved and unresolved triads, but the generality of the tension effect is more easily understood in the acoustical depiction shown on the triadic grids. Specifically, if the Gaussian curve (Figure 1C) is used to model the tension effect, maximal tension is obtained when the difference in the size of the two intervals is zero; when the symmetry is broken and the intervals are unequal, the tension is reduced. Because the results of the tension calculations differ depending upon how many of the upper partials are included, several "tension maps" are shown in Figure 6.

As was the case with the dissonance maps, the tension maps gradually become more complex when more upper partials are included. It is noteworthy, however, that, with the addition of only the first set of upper partials (Figure 6B), oblique lines indicating higher tension fall on all interval combinations that correspond to the augmented, diminished, and suspended fourth triads in all of their inversions. Note that the major and minor triads lie at locations just off of the high-tension strips.

The Stability/Instability of Triads

The success of the tension calculations in indicating that the diminished, augmented, and suspended fourth triads have high tension—in all of their inversions and when played over one or two octaves—suggests that the total "harmonic instability" of 3-tone combinations is a consequence of two independent acoustical factors. The first is a 2-tone effect—interval dissonance ("sensory dissonance")—and has been acknowledged to be an important part of music perception at least since Helmholtz (1877/1954). The second factor is triadic tension—and is explicitly a 3-tone effect (Meyer, 1956).

Computational results for the overall stability/instability of triads are shown in Figure 7. As was the case for the maps of both dissonance and tension, the absolute values of "instability" change as upper partials are added, but regions of relative stability (low instability) are found for the major and minor chords, in all of their inversions and when played over one or two octaves. The unresolved tension chords are located at regions of slightly greater instability surrounding the major and minor chords (Figure 7B-D). Moreover, by combining the dissonance and tension effects into an overall "instability" score, it is found that the experimental sequence of relative stability (major > minor > diminished ≈ suspended fourth > augmented) is reproduced (column C&F, Table 1). The obvious conclusion to draw is that there is a straightforward acoustical basis for the perceptual stability of the major and minor chords. As shown in Table 1, none of the well-known dissonance models achieves even this modest level of agreement with perceptual findings.

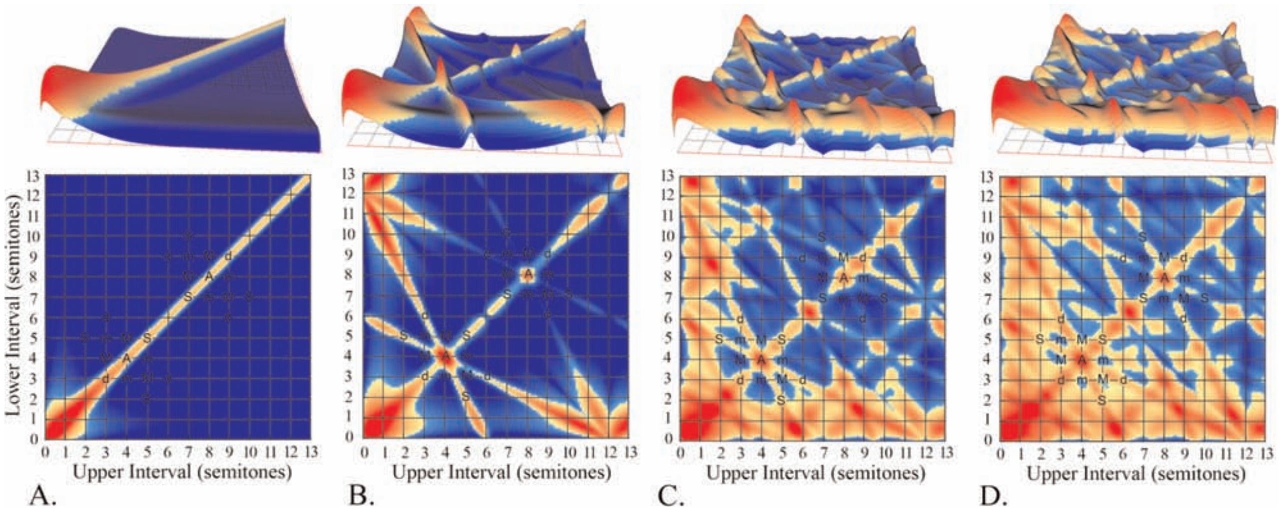

FIGURE 7. The instability maps. (A) When only F0 is included in the calculations, many of the unresolved chords are located in regions of high stability (stipled or blue in the online version). When upper partials are included in calculations (F0-F1, B: F0-F2, C: F0-F3, D), only the major (M) and minor (m) triads remain in regions of relative stability. The unresolved tension chords (d, A, S) lie in surrounding regions of greater instability (non-stipled or red). (For a color version of this figure, please see the digital PDF at www.musicperception.net)

What the instability calculations imply about the history of Western diatonic music is that the musicians of the 14th century who first employed 3-tone chords had become sensitive to the symmetry/asymmetry in the acoustical patterns of 3-tone configurations. Where their Medieval predecessors had been enthralled by lower-level interval effects, Renaissance musicians became more interested in the configuration of polyphonic chords and less interested in the “perfection” of the various consonant intervals. The relative contribution of dyadic dissonance and triadic tension (“sensory” and “musical” dissonance) remains an issue in music perception today, but it is a misunderstanding to maintain that either effect alone explains musical sonority. When music employs intervals, their “perfect” tuning is important, but when chords containing three or more simultaneous tones are used, then the interval effects become secondary, and the tuning of the chord as a chord (i.e., relative interval size) becomes the more salient perceptual phenomenon. Of course, if there is a starkly dissonant semitone interval in a triad, the instability caused by the dissonance will be the most salient feature of the chord. But when the intervals in a triad are not overwhelmingly dissonant, the relative size of the neighboring intervals, not the location of the tones relative to the tonic, becomes more salient, and the phenomenon of harmony itself dominates.

Of historical interest is the fact that, while Renaissance musicians were busy inventing new kinds of polyphonic music that employed primarily asymmetrical 3-tone configurations, Renaissance theorists remained obsessed with the effects of intervals—and devoted their theoretical energies to justifying why 3- and 4-tone chords and chord progressions are “consonant” or “dissonant” due to the use of certain intervals (e.g., Rameau, 1722/1971). Rather than address triadic harmony on its own terms as a “3-tone phenomenon,” the theorists used the theoretical framework that had sufficed for discussion of scales and intervals, i.e., the relative “perfection” of tone dyads. Rameau (1722/1971), for example, maintained that, “the power of the major and minor chords is obtained by the use of the major or minor third. . . . Thus, we can attribute the power of harmony to these intervals” (p. 123). This overemphasis on intervals to the complete exclusion of 3-tone psychophysics remains a problem in more modern attempts at explaining harmony perception (discussed in Cook, 2002; Cook & Fujisawa, 2006; Cook et al., 2007; Cook & Hayashi, 2008).

Harmonic Modality

The previous sections concerning the dissonance, tension, and instability of chords suggest that there are acoustical grounds for considering the major and minor triads to be musically more sonorous than other triads. Undoubtedly, the learning of musical styles and habituation to the different kinds of harmonies, scales and tuning systems in various musical traditions also influence how pitch combinations are perceived. But whatever additional effects are essentially the result of culture, there are clearly structural features of 3-tone chords that contribute to their overall stability and that have less to do with culture than with acoustics.

All of the major and minor triads are rather stable chords—because they lie in broad valleys of low dissonance and, moreover, in smaller pockets of low tension. But, if interval dissonance and triad tension were the only factors determining harmonic sonority, we would expect all of the major and minor chords to sound rather similar—all being small variations on the theme of "triadic stability". That is perceptually not the case, and the musical labels "major" and "minor" were invented presumably because there was something affectively different about these two classes of sonorous chords.

Empirically, the difference between major and minor harmonies is recognized by musicians and nonmusicians, adults and children as young as 4-7 years (Kastner & Crowder, 1990), and peoples from the West, the Indian subcontinent (Bharucha, 1993), and the Far East—despite vastly different musical experience. Typical results of three such experiments in our laboratory testing Japanese undergraduates are shown in Figure 8A.

So, what is the structural feature of these chords that allows people to so unambiguously distinguish between the three major chords and the three minor chords? The textbook explanation of major and minor chords normally is framed in terms of the major and minor scales, and the roles of the intervals of a major or minor third in relation to the tonic, in accordance with traditional harmony theory. As accurate as that description may be, the acoustical factors that determine harmonic perceptions have not previously been identified—and the perception of major and minor modality is often attributed to learned familiarity with the "Western idiom." Indeed, there is even a longstanding contention that the perception of major and minor may be little more than an arbitrary, culture-specific "habit" with no claim to universality—a local custom that has insidiously infiltrated to all parts of the musical world for no reason other than Eurocentric hegemony! On the contrary, however, if the perceived stability of diatonic harmonies can be deduced solely from the acoustical structure of 3-tone chords, then perhaps the perception of major and minor also has an acoustical basis.

A Psychoacoustical Model of Modality

From a state of "intervallic equivalence," there are two directions of pitch movement that can eliminate the structural symmetry and thereby reduce the perceptual tension (Figure 1C). As soon as the two intervals in a (nondissonant) triad differ by one semitone, the tension disappears—and the asymmetrical chord "resolves." Since the only directions available for resolution from the tension of intervallic equivalence correspond to major and minor harmonies (lower interval > upper interval, or vice versa), it is possible to reformulate the tension model such that moving away from tension will result in a quantitative measure of the degree of "majorishness" or "minorishness" of any 3-tone chord (Cook, Fujisawa, & Takami, 2006).

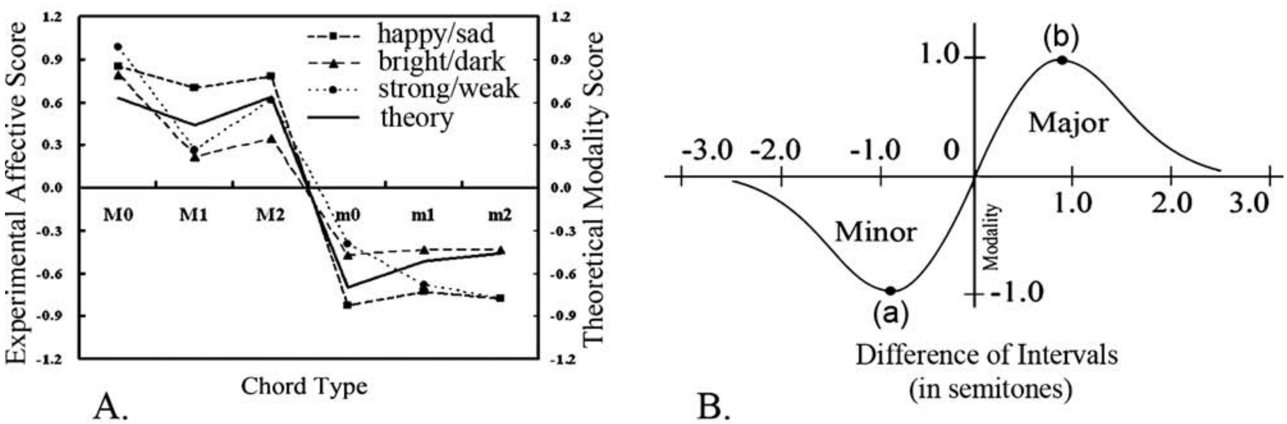

FIGURE 8. (A) The results of three experiments in which 20 (18 or 66) undergraduate nonmusicians evaluated the bright/dark (happy/sad or strong/weak) quality of 72 (24 or 12) isolated major (M0, M1, M2) and minor (m0, m1, m2) chords presented in random (pseudorandom or fixed) order in various keys and at various pitch heights. Indications of differences among the inversions of these chords are also of interest, but, in any case, the affective distinction between major and minor is clear. The thick solid line shows predictions of the theoretical model (Cook & Fujisawa, 2006). (B) The theoretical modality curve. The difference in the magnitude of the intervals (lower minus upper) of a triad will determine its positive (major, b) or negative (minor, a) modality. Note that, for comparison with empirical results, the effects of the upper partials must be plotted on the triadic grids (Figure 9).

The degree of harmonic "tension" was calculated using the "difference of intervals" in any triad of partials (Eq. 2, Figure 1C), but it is also possible to calculate the positive/negative valence of the major and minor chords in relation to interval size (Eq. 4, Figure 8B). When the lower interval is larger than the upper interval, a positive modality score, typical of major chords, is obtained, whereas a smaller interval below a larger interval gives a negative value, typical of minor chords. Specifically, modality (*M*) is defined as:

$$M = -v \cdot \left[\frac{2(y-x)}{\varepsilon}\right] \exp\left\{-\left[\frac{-(y-x)^4}{4}\right]\right\} \quad \text{(Eq. 4)}$$

where v again expresses the product of the amplitudes of the three partials (0.0-1.0), x and y are again the lower and upper intervals, respectively, and the parameter, ε, 1.6 is set to give a positive modality score of 1.0 for the major chord in root position and a negative modality score of −1.0 for the minor chord in root position. Similar to calculation of the total tension of tone combinations, calculation of the total modality (*M*) for any triad with upper partials requires application of Eq. 4 to all triplet combinations of the partials of the three tones. Being based on the difference of interval size, the modality score is mathematically related to the tension score, such that a tension score of 1.0 necessarily implies a modality score of 0.0. The significance of the modality calculation, however, is that, in producing both positive and negative values, it can be used to distinguish between the positive and negative valences of major and minor chords.

Implications of the Modality Model

The modality scores obtained when only the F0s of the three tones are considered are shown on the triadic grid of Figure 9A. An orange ridge of "major modality" (positive valence) is seen when the lower interval is one semitone larger than the upper interval—with a blue valley of "minor modality" (negative valence) running in parallel below it. The vertical heights (depths) indicate that most of the resolved major and minor chords have appropriate (positive and negative) modality scores, respectively, but some are located in (pale yellow and pale blue) regions with modality scores near to zero—theoretically, neither major nor minor, which is contrary to what we know from musical experience (Figure 9A).

When upper partials also are included (Figure 9B-D), however, the match between the theoretical modality maps is remarkably consistent with the perceptual facts. Already with consideration of only the first set of upper partials, peninsulas of positive modality arise at all inversions of the major triad, and complementary troughs of minor modality arise at the minor triads. In contrast, the tension chords (diminished, augmented, and suspended fourth) lie exclusively in in-between regions where there is neither major nor minor modality (Figure 9B). (Note that the small differences in interval structure implied by various tuning systems would have only miniscule effects on the overall modality

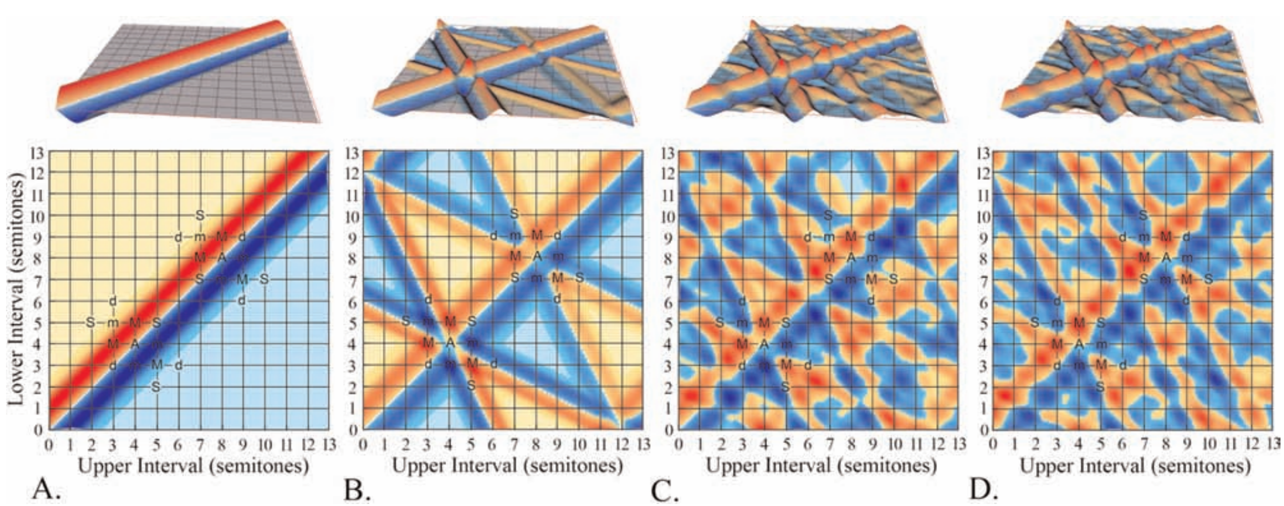

FIGURE 9. The modality maps. (A) When only F0 is included in the calculations, several of the major and minor chords have modality scores near to zero. (B) When the first set of upper partials are included, all of the major and minor chords obtain appropriate positive (nonstipled or red) or negative (stipled or blue) modality scores. (C & D) The modality scores for all major (M) and minor (m) triads remain correct with the addition of further upper partials, F0-F2 and F0-F3, respectively. (For a color version of this figure, please see the digital PDF at www.musicperception.net)

TABLE 2. The Relationships Among the Major, Minor, and Tension Chords.

	Chord after Lowering One Tone				Chord after Raising One Tone		
	n-n-d	n-d-n	d-n-n	Tension Chord	n-n-u	n-u-n	u-n-n
	3-2	2-4	4-3	3-3	3-4	4-2	2-3
	4-3	3-5	5-4	4-4	4-5	5-3	3-4
Interval	5-4	4-6	6-5	5-5	5-6	6-5	4-5
Structure	6-5	5-7	7-6	6-6	6-7	7-5	5-6
(in semitone	7-6	6-8	8-7	7-7	7-8	8-6	6-7
units)	5-1	4-3	6-2	5-2	5-3	6-1	4-2
	2-4	1-6	3-5	2-5	2-6	3-4	1-5
	8-7	7-9	9-8	8-8	8-9	9-7	7-8
	9-8	8-10	10-9	9-9	9-10	10-8	8-9
	3-5	2-7	4-6	3-6	3-7	4-5	2-6
	6-2	5-4	7-3	6-3	6-4	7-2	5-3
	n-n-d	n-d-n	d-n-n		n-n-u	n-u-n	u-n-n
	dom7	dom7	major	dim	minor	—	minor7
	major	major	major	aug	minor	minor	minor
Labels	major	dom7	—	sus4	—	—	minor
from	—	—	—	tritones	—	—	—
Music	—	dom7	major	sus4	minor	—	—
Theory	—	major	dom7	sus4	minor	—	—
(alternatives	dom7	—	major	sus4	—	minor	—
are possible)	major	major	major	aug	minor	minor	minor
	major	dom7	dom7	dim	minor7	minor	minor
	major	dom7	dom7	dim	minor7	minor	—
	major	major	dom7	dim	—	minor7	minor

Note: The cells shaded in lightest grey are all tension triads. The dark grey cells are all major-related (major or dominant-seventh) triads and the mid-tone grey cells are all minor-related (minor or minor-seventh) triads. The unshaded cells do not have common labels from harmony theory; n: no change; d: downward semitone change; u: upward semitone change.

scores of these triads.) The addition of further upper partials makes the modality maps more complex (Figure 9C and D), but the major and minor triads consistently show positive and negative modality scores, respectively, and the tension chords have modality scores near to zero.

The conclusion that can be drawn from the modality calculations is that, among the upper partials of all of the major chords, there is a predominance of triadic structures where the lower interval is one semitone larger than the upper interval. Minor chords show the opposite structural feature. This regularity of the major and minor chords is already apparent when the F0s and F1s of these chords are written in musical notation (Figure 5). The simplicity of this structural feature should come as a surprise to anyone familiar with the complexities of traditional harmony theory, for the conventional view requires the entire edifice of Western music theory—with special consideration of the roles of major and minor thirds in relation to the tonic, and therefore an understanding of key and the use of scales to establish key. Although traditional harmony theory successfully describes modality through such complex arguments, the simplicity of the acoustical explanation suggests that there may be a more direct route to understanding.

Given the structural features that contribute to tension and modality, what can be said about the dynamic relationship between the modal (major and minor) chords and the amodal (tension) chords? This topic is most easily understood in relation to the augmented chord because all of its nearest neighbors on the triadic grid are major or minor. Starting with the unresolved tension of the augmented chord, lowering or raising any of its tones by one semitone will transform its unstable tension into the resolved stability of, respectively, a major or minor chord. In fact, as shown in Table 2, a notable regularity of diatonic harmony in general is that pitch changes in any of the tension chords give similar results. Raise any tone, and one proceeds to a minor (or minor seventh) chord; lower any tone, and one ends up with a major (or dominant seventh) chord. In other words, given the starting point of

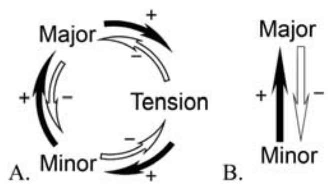

FIGURE 10. (A) The Cycle of Modes. The plus and minus symbols indicate semitone rises or falls. If chords containing interval dissonances are avoided, semitone rises lead from tension to minor to major and back to tension harmonies indefinitely (the solid arrows), whereas semitone falls show the reverse cycle (the open arrows). (B) The traditional view of the major and minor chords is only part of the cycle.

the tension of intervallic equidistance, the rising or falling direction of semitone movement determines the mode of resolution: major (downward) or minor (upward). Although this simple pattern is a direct consequence of the lawfulness of traditional diatonic harmony, it is not discussed in any of the classic texts on harmony theory because the 3-tone psychoacoustical feature described here as harmonic tension is not a part of traditional theory.

The pattern of modality among the triads can be illustrated succinctly as a Cycle of Modes (Figure 10A). The traditional view of mode relationships is that the essential difference between the major and minor chords is the semitone shift that can transform major chords into minor chords, and vice versa (Figure 10B). That view is of course correct, as far as it goes, but implicitly dismisses all other triads as irrelevant "dissonances"—which, technically, is not correct.

By bringing the (unresolved, but not dissonant!) diminished, augmented, and suspended fourth chords into a broader theory of harmony, it is clear that there is an endless cycle of affective modality that can be traversed by raising or lowering triad tones one at a time. Provided that chords containing dissonances are avoided, the cycling will entail repeated transitions from tension-to-major-to-minor and back to tension (with falling tones) or transitions from tension-to-minor-to-major-to-tension (with rising tones). An example is shown in Figure 11.

The Affective Valence of Major and Minor

What has not yet been addressed, however, is why the two modes have their characteristic affective valences. Is it nothing more than a musical custom, a habit, or an entrenched bias for hearing the major chord as somehow strong, bright and positive, and the minor chord as weak, dark and negative—but for no good reason, other than that diatonic music since the Renaissance has always been that way? Could it as easily be reversed?

In discussing the affective valence of chords, musicians rightly object to simple "happy/sad" characterizations of the major and minor triads for the very good reason that many other factors also contribute to the affect in real music (as distinct from chords played in the psychology laboratory). Perhaps a few major chords to celebrate a sports victory or a few minor chords to lament a lost love can be played effectively as a relentless sequence of chords in the same major or minor mode, but music that musicians would describe as interesting, subtle, nuanced, and ultimately effective as music normally includes both major and minor chords, many moments of tension and dissonance, and of course the effects of rhythm, tempo and lyrics. Such music is purposely constructed to elicit an affective atmosphere with twists-and-turns, intimations of positive or negative affect, and the highlighting of internal contradictions before inevitably resolving to the affect of unambiguous major or minor harmonies. Music without tonal resolution can still be interesting—from the trance music of Bali to the geometrical intellectualisms of Schoenberg and Webern—but most music, including modern classical and jazz and virtually all of popular music, uses specifically the major and minor chords to resolve tensions and produce moments of release and composure.

The emotional response to major and minor music has been evaluated experimentally in many previous studies (see, Scherer, 1995, and Gabrielsson & Juslin, 2003, for reviews) and is often discussed in the framework of classical Western music (Cooke, 1959; Scruton, 1997). The emotional effects of mode can of course be suppressed

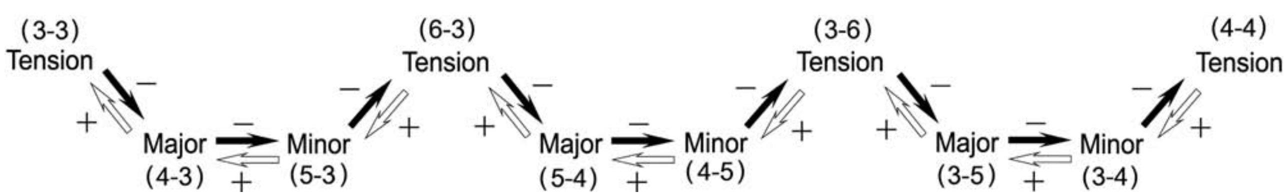

FIGURE 11. Possible transitions among the modes with rises or falls of one semitone in each triad. Many alternatives are possible, and the sequence extends indefinitely. The numbers in parentheses indicate the size of the intervals in each triad, and the + and – symbols indicate semitone steps.

and even reversed through rhythms, timbres and lyrics that tell a different story, but a direct comparison of major and minor triads consistently points in the direction of positive versus negative affect (Kastner & Crowder, 1990). Whatever its origins, that difference is a crucial element of diatonic music because the positive and negative affect associated with major and minor harmonies are focal points where music has meaning: We "feel" major and minor chords at an emotional level that is clearly more significant than the unfocused, meandering of birdsong, the chortle of splashing water in a mountain stream or the harmless whistling of children trying to keep the ghosts away! Individual tones may be pure and isolated intervals may be consonant, but without harmonic focus such auditory candy does not rise to the level of art.

For an explanation of the acoustical origins of the affective response to harmony, the Cycle of Modes can be put to good use. Of the three types of harmonies that do not entail interval dissonance, the tension chords are typically perceived as affectively neutral, inherently ambiguous "amodal" triads. They certainly elicit the emotions associated with ambiguity and uncertainty, and leave the listener in a state of anticipation expecting some sort of resolution, but the tension itself does not point in the direction of either a positive or negative, happy or sad, or weak or strong resolution. Tension is an unresolved starting point from which we await progression to the settled affect of major or minor (Cooke, 1959, p. 64) through some sort of tonal change.

As summarized in Table 2 and Figure 10, it is an empirical fact of diatonic music that (if dissonant chords are avoided) semitone rises in pitch from such affective ambiguity imply the negative affect of the minor mode, whereas semitone falls in pitch from the tension chords imply the positive affect of the major mode (and never the contrary). Of course, multiple pitch rises and falls can move any triad from one mode to any other mode, but the nearest "local" phenomenon in triadic pitch space from a stance of unresolved neutrality to one of emotionality involves semitone steps from tension to major or minor resolution. The simplest formulation of the old puzzle of harmonic modality is therefore to ask why the human ear attaches emotional significance to such changes in auditory frequency? From a harmonic state of affective ambiguity, why do pitch rises lead to negative emotional valence and pitch falls lead to positive valence (Figure 10A)?

Surprisingly, the answer to this question is already known in fields unrelated to music and referred to as the "frequency code." From the study of animal vocalizations, it is known that there is a strong tendency for animals to signal their strength, aggression, and territorial dominance using vocalizations with a low and/or falling pitch and, conversely, to signal weakness, defeat, and submission using a high and/or rising pitch (Morton, 1977). Concrete examples of the frequency code are familiar to most people from the low-pitched growling of aggressive dogs and the high-pitched yelp of injured or retreating dogs, but it is known to be true for species as diverse as gorillas, cats, dogs, elephants, frogs, horses, and birds— virtually any species that uses changes in vocal frequency for the purposes of social communication.

Ohala (1983, 1984, 1994) has been one of the leading advocates of the idea concerning the inherent "sound symbolism" of rising or falling pitch. He has noted that:

> Animals in competition for some resource attempt to intimidate their opponent by, among other things, trying to appear as large as possible (because the larger individuals would have an advantage if, as a last resort, the matter had to be settled by actual combat). Size (or apparent size) is primarily conveyed by visual means . . . As Morton (1977) points out, however, the F0 of voice can also indirectly convey an impression of the size of the signaler, since F0, other things being equal, is inversely related to the mass of the vibrating membrane (vocal cords in mammals, syrinx in birds), which, in turn, is correlated with overall body mass . . . To give the impression of being large and dangerous, then, an antagonist should produce a vocalization as . . . low in F0 as possible. On the other hand, to seem small and non-threatening, a vocalization which is tone-like and high in F0 is called for . . . Morton's (1977) analysis, then, has the advantage that it provides the same motivational basis for the form of these vocalizations as had previously been given to elements of visual displays, i.e., that they convey an impression of the size of the signaler. I will henceforth call this cross-species F0-function correlation "the frequency code." (p. 330)

A perceptible increase or decrease in pitch signifies a change in the vocalizing animal's assumed social position. Other behaviors are also used to signal or establish dominance, but the "frequency code" is the primary auditory means for affective signaling. While most other behaviors have species-specific significance, rising or falling F0 has cross-species generality and profound meaning for any animal within earshot, regardless of night-time obscurity, visual angle, or jungle obstructions. A falling F0 implies that the vocalizer is not in retreat, has not backed down from a direct confrontation, may become a physical threat and has assumed a stance of social dominance. Conversely, a rising F0 indicates defeat, weakness, submission, an unwillingness to challenge, and signals the vocalizer's acknowledgement of nondominance. As shown by Morton (1977) and Ohala (1984, 1994), a falling voice signals strength because—all else being equal—a low auditory frequency indicates a larger object than a high frequency. This is as true for vocal cord resonances as for church bells: large vibrating cavities produce low sounds, small ones high sounds. As a consequence, this instinctively understood signal is widely used by many species as a hopeful ploy to scare away potential rivals without engaging in actual combat.

There is an extensive and fascinating academic literature on how and why these F0 signals have evolved, their correlations with facial expressions (smiles are correlated with higher frequencies) and the related, inherent "sound symbolism" of vowel sounds [see Bolinger (1978); Cruttendon (1981); Juslin & Laukka (2003); Ladd (1996); Levelt (1999); Morton (1977); Ohala (1983, 1984, 1994); Scherer (1995); Scherer et al. (2003) for further discussion]. But the important point in the present context is simply that the pervasive use of vocal pitch changes by diverse animal species to indicate social status is empirically well-established.

If the "frequency code" were merely a peculiarity of animal communications, it could possibly be dismissed as irrelevant to human behavior in general and music in particular, but in fact the universality of such sound symbolism is known to have spilled over into human languages: Rising and falling voice intonations have related, if greatly attenuated, meanings concerning social status in human interactions. Across diverse languages, falling pitch is again used to signal social strength (commands, statements, dominance) and rising pitch to indicate weakness (questions, politeness, deference and submission): "in both speech and music, ascending contours convey uncertainty and uneasiness, and descending contours certainty and stability" (Brown, 2000, p. 289). As argued most forcefully by Ohala (1983, 1984, 1994), the inherent meaning of pitch rises or falls is one of a very small number of crosslinguistic constants that have been found in all human languages. As a consequence, although we may have no idea what the babbling foreigner is trying to communicate, his tone of voice will clearly indicate whether it is a command or a question—and whether his assumed social status is one of strength or weakness.

The most common statement of the frequency code in linguistics is in relation to the rising auditory frequency used in interrogatives. Although it is possible to ask questions with a falling tone of voice (usually carrying some implication of strength or authority about the speaker), an inquiry that indicates a lack of information and a desire for an answer from another person is most frequently stated with a noticeable rise in the vocal pitch of the speaker. Since "establishing dominance" is only one of many aspects of human communication, rising and falling pitch is only part of the sound symbolism of speech utterances, but its crosscultural prevalence demonstrates the importance of our biological roots— extending even to the realm of language.

The frequency code as known from both animal communications and human speech can be stated simply as: Falling pitch signals strength, rising pitch signals weakness. In both cases, the pitch context is provided by the tonic or "natural frequency" of the individual's voice, so that the meaning of the frequency code is apparent simply from the direction of rising or falling pitch. In the context of diatonic music, however, the "meaning" of pitch changes can be deciphered only within a specific musical context. Musical key and the location of the tonic are not "givens," but must be established within the context of the ongoing music. Normally, that is done gradually—sometimes with intended ambiguities and delays, but nearly always evolving toward a definite key within which the listener can appreciate the musical significance of any pitch movement.

The question concerning the affective valence of diatonic harmonies then becomes: What is the minimal musical context from which pitch movement will allow the listener to hear unambiguous musical meaning? In diatonic music, the affect of a major or minor key can be established simply and unambiguously through the use of a resolved harmonic triad. Since a modal triad requires a pitch range of at least 7 semitones, a modally ambiguous triad over a range of 6 (or more) semitones provides a sufficient context from which a semitone rise or fall will establish a major or minor key. It is a simple consequence of the regularities of diatonic harmony that, given this minimal context, a semitone rise can resolve to a minor key and a semitone fall can resolve to a major key, but not vice versa. Remarkably, pitch movement from any 3-tone (nondissonant) combination that is neither inherently major nor inherently minor shows this same general pattern (see Table 2).

It is a noteworthy fact that the direction of tonal movement from the ambivalence of amodal tension to a major or minor triad is the same as the direction of pitch changes with inherent affect in animal vocalizations and language intonation (Figure 12). In the long history of animal evolution, upward pitch movement has come to imply the negative affect of social weakness,

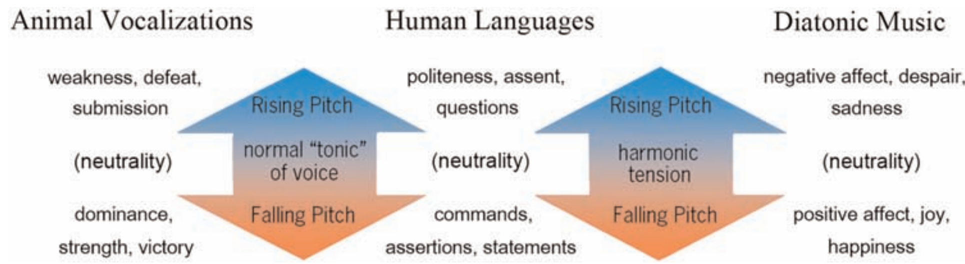

FIGURE 12. The "frequency code" manifests itself in animal calls, language and music. Given an appropriately neutral starting point, rising or falling pitch has related meanings in all three realms: positive affect with pitch decreases, negative affect with pitch increases. (For a color version of this figure, please see the digital PDF at www.musicperception.net)

whereas downward pitch movement implies the positive affect of social strength. It is therefore a plausible hypothesis that, within the realm of diatonic pitch space, when a 3-tone combination shifts away from the unresolved acoustical tension of intervallic equidistance toward resolution, we inevitably infer an affective valence from the detection of the direction of tonal movement: a semitone shift up is weak (submissive, retreating), a semitone shift down is strong (dominant, assertive). The fact that we feel anything at all—and do not experience such harmonic phenomena as cold and emotionless "information processing"—is indication that our biological beings have been activated. When embedded in a richly embroidered musical context, the emotional response to modal harmony can occur together with piloerection, lacrimation, tachycardia, or brachycardia and other indications of autonomic arousal (Gabrielsson & Juslin, 2003). These are not signs of cool headed, detached listening, but, on the contrary, indications of emotional involvement.

The similarity of the binary pattern of affect in response to pitch changes in all three realms (Figure 12) suggests an ancient evolutionary history underlying the common perception of major and minor chords. In this view, the positive/negative affect of falling/rising musical pitch is another manifestation of the evolutionary roots of human behavior in general. Since the frequency code has previously been identified as a low-level, universal "code" for signaling social status relations, it would be parsimonious to argue that the same pitch phenomena in those realms are at work when human beings perceive the affect of harmony. Definitive evidence for the commonality of these three phenomena will probably require brain-imaging indication of the same cortical regions being involved in various manifestations of the frequency code.

The implications of the sound symbolism hypothesis are complex and far-reaching. Suffice it to say that, even for music without lyrics and without imitation of the sounds of nature (i.e., for the inherent meaning of "absolute music"), music has at least one form of biological grounding in the use of instinctively understood pitch movement. In this view, each and every pitch rise or fall contains a miniscule implication of strength or weakness among competing animal species. Within the framework of diatonic music, in which the meandering of pitch typical of birdsong is replaced by the musical context known as key, the usage of harmonies again provides an affective biological grounding in relation to the neutrality of harmonic tension. In addition to the various ups-and-downs of melody, whenever 3-tone combinations produce the ambiguity of intervallic equidistance, a subsequent fall in pitch conjures up the biological twinge of "social strength," whereas a rise suggests "social weakness." Effective music interweaves these evolutionarily ancient, instinctively understood pitch signals with changes in rhythm, timbre, and lyrics, such that music will virtually never be heard as a statement of victory or defeat, but the grounding to our biological beings—the faint probes of the autonomic nervous system that we feel on hearing well-crafted music—is arguably a consequence of the frequency code woven into what might otherwise be heard as nothing more than arid, inherently emotionless changes in auditory frequency.

Conclusion

The perceptual stability/instability and major/minor modality of triads have clear-cut acoustical foundations that can be explicated in terms of the 3-tone partial structure of chords. Precisely how the auditory nervous system might undertake the 2-tone dissonance and 3-tone tension/modality calculations and combine them into musical percepts remains to be studied in terms of brain activity.

It also remains to be seen how much of the complexity of traditional harmony theory can be reconfigured solely on a psychoacoustical basis. Outstanding questions include whether or not the empirical evidence on the usage of harmonic cadences in classical and popular music (Eberlein, 1994; Huron, 2006) can be explained, whether there are additional acoustical features of 4- and 5-tone harmonies (Kuusi, 2002) that are more than the sum of 2- and 3-tone effects, and of course what the relative effects of exposure, learning, and training in musical traditions may be. However those more complex questions may eventually be answered, it can be said that the core phenomena of diatonic harmony, built from pitch triads, have an acoustical simplicity that becomes apparent as soon as we look beyond the inherent limitations of 2-tone psychophysics.

Author Note

My thanks go to Toshihiko Amemiya for pointing out the relevance of the frequency code, Gene Narmour for advice on radical reductionism in music perception, Takefumi Hayashi for computer graphics guidance, Takashi Fujisawa for devising the modality curve, Kazuaki Takami for first drawing the Cycle of Modes, and Hiroo Konaka, Chikashi Fujimoto, and Shun Fukushima for programming work. Our combined efforts have resulted in a computer graphics software program, "Seeing Harmony," that is available at: http://www.res.kutc.kansai-u.ac.jp/~cook

Correspondence concerning this article should be addressed to Norman D. Cook, Department of Informatics, Kansai University, Takatsuki, Osaka, 569 Japan. E-MAIL: cook@res.kutc.kansai-u.ac.jp

References

Ball, P. (2008). Facing the music. *Nature,* 453, 160-162.

Bernstein, L. (1976). *The unanswered question: Six talks at Harvard*. Cambridge, MA: Harvard University Press.

Bharucha, J. J. (1993). Tonality and expectation. In R. Aiello (Ed.), *Musical perceptions* (pp. 213-239). Oxford: Oxford University Press.

Bolinger, D. L. (1978). Intonation across languages. In J. H. Greenberg, C. A. Ferguson, & E. A. Moravcsik (Eds.), *Universals of human language: Phonology* (pp. 471-524). Palo Alto, CA: Stanford University Press.

Brown, S. (2000). The 'musilanguage' model of music evolution. In N. L. Wallin, B. Merker, & S. Brown (Eds.), *The origins of music* (pp. 271-300). Cambridge, MA: MIT Press.

Cook, N. D. (2001). Explaining harmony: The roles of interval dissonance and chordal tension. *Annals of the New York Academy of Science, 930*, 382-385.

Cook, N. D. (2002). *Tone of voice and mind: The connections between music, language, cognition and consciousness.* Amsterdam: John Benjamins.

Cook, N. D. (2007). The sound symbolism of major and minor harmonies. *Music Perception, 24*, 315-319.

Cook, N. D., & Fujisawa, T. X. (2006). The psychophysics of harmony perception: Harmony is a three-tone phenomenon. *Empirical Musicology Review, 1*, 106-126.

Cook, N. D., Fujisawa, T. X., & Konaka, H. (2007). Why not study polytonal psychophysics? *Empirical Musicology Review, 2*, 38-43.

Cook, N. D., Fujisawa, T. X., & Takami, K. (2006). Evaluation of the affective valence of speech using pitch substructure. *IEEE Transactions on Audio, Speech and Language Processing, 14*, 142-151.

Cook, N. D., & Hayashi, T. (2008). The psychoacoustics of harmony perception. *American Scientist*, *96*, 311-319.

Cook, P. R. (Ed.) (1999). *Music, cognition and computerized sound: An introduction to psychoacoustics*. Cambridge, MA: MIT Press.

Cooke, D. (1959). *The language of music.* Oxford: Oxford University Press.

Cruttendon, A. (1981). Falls and rises: Meanings and universals. *Journal of Linguistics, 17*, 77-91.

Eberlein, R. (1994). *Die Entstehung der tonalen Klangsyntax* [The emergence of the syntax of tonal harmony]. Frankfurt: Lang.

Gabrielsson, A., & Juslin, P. N. (2003). Emotional expression in music. In R. J. Richardson, K.R. Scherer & H. H. Goldsmith (Eds.), *Handbook of affective sciences* (pp. 503-534), Oxford: Oxford University Press.

Helmholtz, H. L. F. (1954). *On the sensations of tone as a physiological basis for the theory of music*. New York: Dover. (Original work published in 1877)

Huron, D. (2006). *Sweet anticipation.* Cambridge, MA: MIT Press.

Huron, D. (2008). Lost in music. *Nature, 453*, 456-457.

Juslin, P. N., & Laukka, P. (2003). Communication of emotions in vocal expression and music performance: Different channels, same code? *Psychological Bulletin, 129*, 770-814.

Kameoka, A., & Kuriyagawa, M. (1969). Consonance theory: Parts I and II. *Journal of the Acoustical Society of America, 45*, 1451-1469.

Kastner, M. P., & Crowder, R. G. (1990). Perception of major/minor: IV. Emotional connotations in young children. *Music Perception, 8*, 189-202.

KUUSI, T. (2002). Theoretical resemblance versus perceived closeness: Explaining estimations of pentachords by abstract properties of pentad classes. *Journal of New Music Research, 31*, 377-388.

LADD, D. R. (1996). *Intonational phonology*. Cambridge, UK: Cambridge University Press.

LEVELT, W. J. M. (1999). Producing spoken language: A blueprint of the speaker. In C. M. Brown & P. Hagoort (Eds.), *The neurocognition of language* (pp. 83-122). Oxford: Oxford University Press.

MCDERMOTT, J. (2008). The evolution of music. *Nature, 453*, 287-288.

MEYER, L. B. (1956). *Emotion and meaning in music*. Chicago, IL: Chicago University Press.

MORTON, E. S. (1977). On the occurrence and significance of motivation-structural roles in some bird and mammal sounds. *American Naturalist, 111*, 855-869.

OHALA, J. J. (1983). Cross-language use of pitch: An ethological view. *Phonetica, 40*, 1-18.

OHALA, J. J. (1984). An ethological perspective on common cross-language utilization of F0 in voice. *Phonetica, 41*, 1-16.

OHALA, J. J. (1994). The frequency code underlies the sound-symbolic use of voice-pitch. In L. Hinton, J. Nichols, & J. J. Ohala (Eds.), *Sound symbolism* (pp. 325-347). Cambridge, UK: Cambridge University Press.

PARNCUTT, R. (1989). *Harmony: A psychoacoustical approach*. Berlin: Springer.

PIERCE, J. R. (1992). *The science of musical sound.* New York: Freeman.

PLOMP, R., & LEVELT, W. J. M. (1965). Tonal consonances and critical bandwidth. *Journal of the Acoustical Society of America, 38*, 548-560.

RAMEAU, J.-P. (1971). *Treatise on harmony* (P. Gossett, Trans.). New York: Dover. (Original work published in 1722)

ROBERTS, L. (1986). Consonant judgments of musical chords by musicians and untrained listeners. *Acustica, 62*, 163-171.

SCHERER, K. R. (1995). Expression of emotion in voice and music. *Journal of Voice, 9*, 235-248.

SCHERER, K. R., JOHNSTONE, J., & KLASMEYER, K. (2003). Vocal expression of emotion. In R. J. Richardson, K. R. Scherer, & H.H. Goldsmith (Eds.), *Handbook of affective sciences* (pp. 433-456). Oxford: Oxford University Press.

SCRUTON, R. (1997). *The aesthetics of music.* Oxford: Oxford University Press.

SETHARES, W. A. (1999). *Tuning, timbre, spectrum, scale*. Berlin: Springer.

TERHARDT, E. (1974). Pitch, consonance and harmony. *Journal of the Acoustical Society of America, 55*, 1061-1069.

TRAINOR, L. (2008). The neural roots of music. *Nature, 453*, 598-599.

TRAMO, M. J., CARIANI, P. A., DELGUTTE, B., & BRAIDA, L. D. (2001). Neurobiological foundations for the theory of harmony in Western tonal music. *Annals of the New York Academy of Sciences, 930*, 92-116.