

Why Not Study Polytonal Psychophysics?

NORMAN D. COOK [1]
Kansai University

TAKASHI X. FUJISAWA [2]
Kwansei Gakuin University

HIROO KONAKA [3]
Kansai University

ABSTRACT: The relative consonance/dissonance of 2-tone intervals is well understood both experimentally and theoretically and provides a strong foundation for explaining why diatonic scales or their subsets are used in most musical cultures. Frequent textbook assertions notwithstanding, however, the consonance of intervals fails to account for the basic facts of harmony (3 or more tone combinations). We have recently shown (Cook & Fujisawa, 2006) how consideration of 3-tone psychophysics can explain the fundamental regularities of diatonic harmony. Distinct from the dissonance of 2-tone intervals, 3-tone combinations introduce an effect described by Leonard Meyer (1956) as harmonic “tension”: when a third tone is located midway between an upper and a lower tone, the chord takes on an unresolved, unstable, tense character – a psychoacoustical property inherent to the diminished and augmented chords. If the effects of the upper partials are included in a formal model that includes both 2-tone and 3-tone effects, the perceived sonority of the triads (major>minor>diminished>augmented) is easily explained.

Submitted 2006 December 2; accepted 2006 December 17

KEYWORDS: *harmony, psychophysics, dissonance, tension, major mode, minor mode*

WHILE noting the fact that there are yet unresolved issues in the psychophysics of harmony perception, Parncutt (2006a) did not address the main question that we raised concerning the possible role of 3-tone combinations (Cook & Fujisawa, 2006). We have often cited his well-known work (*Harmony: A psychoacoustical approach*, 1989) because it is one of a small number of psychophysical models that was developed explicitly to explain the perception of diatonic harmony. Although many authors have implied as much, Parncutt and a few others (Terhardt, 1974; Tramo et al., 2001) have explicitly stated that the perception of chords is principally an issue of the perception of the summed effects of the 2-tone intervals (and their upper harmonics). This strongly reductionist approach has in fact had great success in explaining interval perception itself, but the inability to account for even the most basic facts about harmony (in the usual sense of combinations of three or more pitch classes) is a serious short-coming for any model that claims to provide the scientific basis for the perception of diatonic music.

Experiments (with musicians and non-musicians from diverse cultures) and the statistics on the prevalence of triads in 18th and 19th Century classical music indicate that diminished chords are more *consonant* than augmented chords, but the Parncutt (1989) model incorrectly predicts that they should be more *dissonant*. This small discrepancy between theory and experiment is, however, only one flaw in a pattern of incorrect model predictions that we have shown (Cook & Fujisawa, 2006) to be characteristic of the interval-based models of harmony perception. More fatal than the incorrect prediction of the relative instability of diminished and augmented chords is the fact that Parncutt’s model (Table 1) also predicts that the augmented chord is more consonant than 5 of the 6 major and minor triads, and that the 2nd inversion of the minor triad is the most dissonant of all the triads! Such model predictions are, perceptually, nonsense. Since all of the harmony models that are based on interval dissonance make similarly poor predictions about triad perception, we have challenged the claim that even an approximate understanding of Western harmony can be obtained from the psychoacoustics of interval perception.

Table 1 shows the theoretical predictions of several of the well-known consonance models, as well as two types of empirical evidence. The first kind of empirical data is from laboratory experiments by Roberts (1986), in which she had musicians and non-musicians evaluate the “consonance” of the common triads in root and inverted positions (presented as isolated chords). For both groups of subjects, 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 (Cook, 1999; Cook & Fujisawa, 2006; Fujisawa, Konaka & Cook, 2006). The second type of data concerns the prevalence of harmonic triads in a sample of 1538 chords in classical music (from Bach, Handel, Mozart, Beethoven and Mendelssohn), as collected by Eberlein (1994; cited by Parncutt, 2006a). The prevalence of chord types is identical to the sequence of consonance as report by Roberts and us (with some variations among the triad inversions), and with the suspended fourth chords being slightly more prevalent than the augmented chord. Again, there is no surprise in the empirical data, which conform to musical “common sense” about the relative sonority of the triads.

Table 1. The relative sonority of common triads in root and inverted positions.

Chord Class	Interval Structure	Empirical Sonority			Theoretical Sonority					
		Roberts (1986)	Eberlein (1994)	FKC (2006)	C&F (2006)	Sethares (1999)	Parncutt (1989)	K&K (1969)	P&L (1965)	Helmholtz (1877)
I. Major	4-3	1	1	1	1	4	1	1	4	3
	3-5	3	3	4	5	<u>8</u>	6	<u>11</u>	<u>11</u>	<u>9</u>
	5-4	2	7	2	4	2	3	6	2	1
II. Minor	3-4	4	2	3	2	4	4	1	4	3
	4-5	6	4	5	3	2	6	6	2	1
	5-3	5	6	6	6	<u>8</u>	<u>10</u>	<u>11</u>	<u>11</u>	<u>9</u>
III. Dim	3-3	7	9	7	12	<u>12</u>	9	6	<u>13</u>	<u>13</u>
	3-6	8	5	12	7	10	5	9	8	11
	6-3	9	10	10	10	10	8	9	8	11
IV. Aug	4-4	10	11	13	13	12	<u>2</u>	13	10	<u>5</u>
	5-2		8	9	8	<u>6</u>		<u>1</u>	<u>6</u>	<u>5</u>
V. Sus 4 th	2-5		11	8	11	<u>6</u>		<u>1</u>	<u>6</u>	<u>5</u>
	5-5		11	11	9	<u>1</u>		<u>1</u>	<u>1</u>	8

Experimental data are from Roberts (1986, musicians) and Fujisawa, Konaka & Cook (2006, non-musicians); empirical prevalence values are from Eberlein (1994, p. 421). The theoretical sonority values are from Cook & Fujisawa (2006), Sethares (1999, p. 92), Parncutt (1989, p. 140), Kameoka & Kuriyagawa (1969), Plomp & Levelt (1965), and Helmholtz (1877, p. 193). Striking anomalies in the sonority ranking of these models are underlined in bold type.

How well do models of the consonance/dissonance of diatonic *dyads* explain the relative sonority of the harmonic *triads*? The comparison between empirical and theoretical values in Table 1 clearly indicates that, whatever merits these models may have with regard to explaining why there exist diatonic scales, they fail dramatically to explain harmony perception. Note that the model failures are not confined to small reversals of the order of sonority of chord inversions, but are stark contradictions with musical common sense and the historical usage of triads. Specifically, the Helmholtz model predicts that the augmented and suspended fourth chords should be more consonant than two of the major and minor chords: the Sethares, Kameoka & Kuriyagawa, and Plomp & Levelt models predict that the suspended fourth triads are more consonant than the major or minor chords; and the Parncutt model predicts greater sonority of the augmented chord than most of the major and minor chords. To recover from such clearly anomalous predictions, theorists must then invoke “learning effects” and qualitative arguments from traditional harmony theory (“root clarity”, Parncutt, 2006a) that are said to somehow trump the underlying psychoacoustics.

In contrast to models that are built on the summation of 2-tone effects, the psychophysical model we advocate (the C&F column in Table 1; Cook & Fujisawa, 2006; Cook, Fujisawa & Takami, 2006) contains both the effects of 2-tone interval consonance/dissonance and the effects of 3-tone “tension”, as calculated from the relative size of the two intervals contained in any 3-tone combination (see below). So doing, the predicted sequence of the “harmoniousness” or “sonority” of chords is identical to the experimental sequence (major>minor>diminished>augmented). There is also surprisingly good agreement with the sequence of inversions within each chord category in the Eberlein data (1994) – of which we were not aware until Parncutt’s recent commentary (2006a). Because of small differences in the evaluation of the sonority of the various chord inversions between musicians and non-musicians in the Roberts data and the inevitable bias in the sample of the music in the Eberlein data, it would be pointless to adjust model parameters to obtain a perfect fit with empirical data, but any model that purports to explain the perception of diatonic harmony certainly must be able to reproduce the gross sequence: major>minor>diminished>augmented. This rather modest demand is not met by any of the interval-based models, but is immediately achieved when the tension effects of 3-tone combinations are brought into consideration.

Our main contention is that the psychoacoustics of pitch perception does not come to an abrupt halt at the level of 2-tone combinations, but can be studied using psychophysical techniques at least up to the level of 3- and 4-tone combinations. On the one hand, Parncutt (2006a) now states his belief that harmony is more than the summation of interval effects, but, on the other hand, the interval masking and fusion effects emphasized in his earlier work (Parncutt, 1989) are essentially 2-tone effects that do not go substantially beyond the previous interval models. Be that as it may, the reluctance to consider 3-tone configurations by several generations of psychophysicists has meant that the core phenomena of musical harmony – tension and resolution, and the affective valence of the major and minor modes – have remained scientific mysteries, while of course being central to nearly all classical and popular music (Narmour, 1990) and, indeed, lying at the heart of the subjective experience of diatonic music.

We maintain that the perception of the unresolved “tension” of a chord is a universal, psychophysical phenomenon distinct from the summed dissonance of its intervals. Both interval dissonance and triadic tension contribute to the overall “instability” of pitch combinations – and hence to listener evaluations of the “pleasantness”, “beauty” or “sonority” of chords, but tension is not due to dissonance (nor vice versa) and it is a disservice to the field of music psychology to conflate these issues. Not only are there psychophysical and music theoretical grounds for distinguishing between 2- and 3-tone effects, but we have found distinct patterns of brain activation in response to chords containing a semitone dissonance and chords with a 3-tone pattern implying tension (Cook, 2002; Cook, Callan & Callan 2002). Perhaps the strongest support for making a distinction between 2- and 3-tone effects are brain-damage studies that indicate a dissociation between dissonance perception and evaluation of the affective valence of chords. Peretz et al. (2001) have reported a patient who was unable to distinguish between consonance and dissonance, but exhibited normal responses to the major and minor modes. Clearly, modality cannot be the summation of interval effects for such patients, if the ability to detect consonant and dissonant intervals is lost, but the sonority of chords has remained intact. Since the sonority of chords is not identical to the summed “consonance” of its intervals, we advocate usage of the terms “dissonance” and “consonance” to describe the clash or agreement between *two tones*, and “tension”, “harmoniousness” and “sonority” to describe the instability/stability inherent to a *three-tone* chord.

Introducing three-tone effects into models of music perception increases the complexity of theoretical calculations, but such effects are both conceptually and computationally tractable. Contrary to Parncutt’s assertion that our model is mathematically abstruse, we reiterate that the calculation of harmonic tension is a simple matter of calculating the relative size of the intervals in triads. As all psychophysicists know, transforming even a simple idea into a formal model requires mathematical formulation that can be daunting to the uninitiated, but a distinction between the mathematical statement of intuitively-plausible ideas, on the one hand, and “mathematics for the sake of mathematical beauty”, on the other, should be made. Parncutt’s (2006b, p. 207) dismissal of our tension calculation as “mysterious” mathematics – somehow on a par with the 1263 page mathematical *tour de force* of Mazzola (2002) – is therefore quite unwarranted. The essential attractiveness of Meyer’s idea of “intervallic equidistance” lies in its intuitive clarity. When compared to the Ptolemaic complexities of traditional harmony theory, the underlying idea of harmonic tension is simplicity itself – and quantifiable in the form of a single equation (see Cook, 2002; Cook & Fujisawa, 2006; Cook, Fujisawa & Takami, 2006; Fujisawa & Cook, 2006). Specifically, we use Meyer’s idea (1956) that tension is perceived when three tones are equally spaced in pitch space. Intuitively obvious examples not requiring any mathematics at all are the diminished chord in root position (two

intervals of 3-semitones) and the augmented chord (two intervals of 4-semitones), but the overall harmonic tension of any triad can be obtained easily by calculating the relative size of the intervals contained. Consideration of the relative spacing and amplitudes of the upper partials of the fundamental frequencies makes the calculated “tension curves” graphically more complex, but the concept and its perceptual clarity could hardly be more straightforward.

Our negative argument regarding previous models is that, although interval-based psychophysics is an important starting point for understanding music perception, it does not suffice to explain harmony. Our constructive argument is that *3-tone psychophysics* both clarifies the issue of the relative “stability” (harmoniousness, prevalence) of the triads, *and* indicates the structural features that are at the heart of the affective valence of the major and minor modes (Cook, 2002, 2007; Cook & Fujisawa, 2006). Rather than consider the possibility of 3-tone effects, Parncutt retreats to qualitative arguments concerning “root clarity” – interesting and possibly valid within the context of traditional harmony theory, but no longer within the Helmholtzian scientific tradition of psychoacoustics.

It is here that we find an internal contradiction in the Parncutt approach. On the one hand, he makes use of the conventional dissonance model from Plomp & Levelt (1963) and calculates a rather complex continuous function of the total dissonance of intervals that is dependent on the number and amplitude of upper partials (see Figure 2 in Cook & Fujisawa, 2006). Similar to all of the other psychophysical models of dissonance, he does *not* maintain that some intervals are categorically dissonant and others are categorically consonant, but rather models dissonance/consonance as a continuous function. On the other hand, he objects to an analogous continuous function for the calculation of harmonic tension – insisting that tension is heard categorically, with no “in between,” and consequently finds a problem in the steepness of the slope of our Gaussian tension curve. In response, we note that the calculation of the tension effect using curves with steep or gentle slopes (simulating categorical versus continuous perception) does not notably change the calculated tension effects (Figure 1).

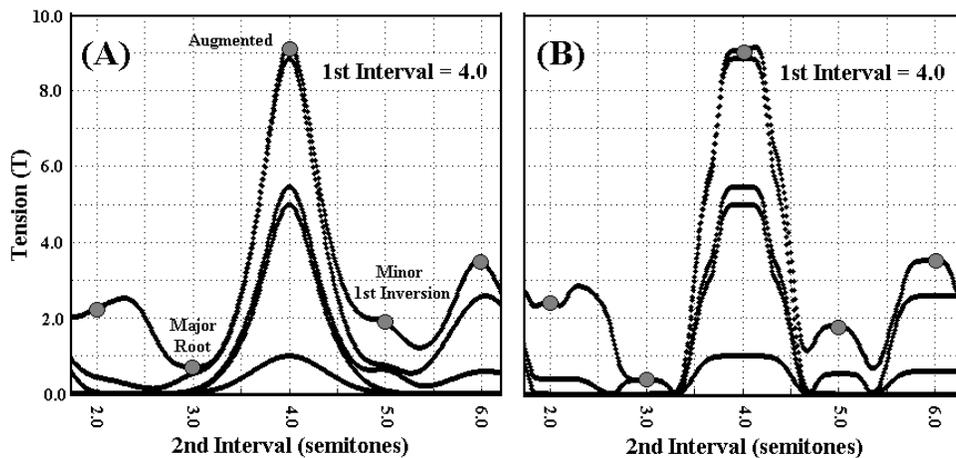


Fig. 1. A comparison of curves used to compute the tension function. In (A) is shown the original Gaussian curve from Cook & Fujisawa (2006), together with the more complex curves resulting from the inclusion of the effects of upper partials. In (B) is shown a steeper, nearly step-wise function (the Gaussian curve to the third power) that more closely simulates categorical perception of tension as “present or absent”. The total tension curves are obtained by including the effects of 1-4 upper partials (amplitudes = $1/n$). In this example, the calculations assume a pitch triad with a lower interval of 4-semitones, and an upper interval that is allowed to vary continuously from 2 to 6 semitones. In both (A) and (B), notable peaks are seen at the augmented chord (two intervals of 4 semitones), and notable troughs of tension are seen at the major chord in root position (intervals of 4 and 3 semitones) and the minor chord in 1st inversion (intervals of 4 and 5 semitones). It is noteworthy that small adjustments in the shape of the tension function or the number of upper partials do *not* alter the location of the peaks and troughs in the total tension curves.

Why has three-tone psychophysics been given so little attention in music psychology? Several reasons come to mind, but none provides a true answer. For example, some music theorists have argued that normal listeners simply cannot “hear” three distinct tones. That is undoubtedly true in the sense that,

unlike some of the musical savants who can identify up to 10 distinct tones when played simultaneously, most musicians and non-musicians struggle to identify even three or four simultaneous tones. But tone-identification aside, it must be asked: “What do we hear when we listen to chords in the major or minor mode, if not three-tone effects?” The characteristic modality of major and minor chords has an affective ring that children as young as 4 years can identify (Kastner & Crowder, 1990) and that is absent in isolated pitch intervals. Similarly, the tension inherent to an augmented or diminished chord is perceptible to normal listeners from diverse musical backgrounds, and the lack of resolved sonority (“stability”) of such chords is clearly *not* due to the dissonance of small intervals (Table 1). Our basic contention with regard to the scientific study of music perception is that music psychologists must move beyond the desperate – and demonstrably unsuccessful – reductionism of the interval models and their tacit assumption that the only science of musical pitch is the psychophysics of two-tone combinations. The musical importance of three-tone combinations is evident from the fact that pitch triads are the beginning of the phenomena of resolved vs. unresolved harmonies and of the major and minor modalities. Conceptually, experimentally and computationally, three-tone combinations are not overwhelmingly complex, and can be studied prior to addressing more difficult questions concerning the perception of “real” musical compositions – with the inevitable complications of musical training, familiarity and culture.

In summary, we agree with Parncutt in his belief that “there is a need to compare the predictions of such models... with statistical distributions of real musical structures” (Parncutt, 2006a, p. 207) and that “The proof ... [is]... in the pudding: the model must first successfully account for a wide variety of empirical data, and the appropriateness of its mathematical formulations should primarily be evaluated on that basis” (Parncutt, 2006b, p. 203). But we must point out that interval model predictions on triad perception are inconsistent with the Eberlein (1994) data, which Parncutt himself cites, but are largely consistent with a psychophysical model that includes 3-tone effects (Table 1). Meanwhile, the more difficult challenge of explaining the sonority and prevalence of the harmonic *tetrads* remains to be taken up.

ACKNOWLEDGMENT

This research was supported in part by an Academic Frontiers Project at Kansai University (2003-2007).

NOTES

[1] N.D. Cook can be reached at the Department of Informatics, Kansai University, Takatsuki, Osaka 569-1095 Japan and at: cook@res.kutc.kansai-u.ac.jp

[2] T.X. Fujisawa can be reached at the Department of Informatics, Kwansai Gakuin University, Sanda, Hyogo 669-1337 Japan and at: fujisawa@ksc.kwansei.ac.jp

[3] H. Konaka can be reached at the Department of Informatics, Kansai University, Takatsuki, Osaka 569-1095 Japan and at: m7105@gra.kutc.kansai-u.ac.jp

REFERENCES

Cook, N.D. (1999). Explaining harmony: the roles of interval dissonance and chordal tension. *Annals of the New York Academy of Science*, Vol. 930, pp. 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* (in press).

- Cook, N.D., Callan, D.E., & Callan, A. (2002). Frontal areas involved in the perception of harmony. *8th International Conference on Functional Mapping of the Human Brain*, Sendai, Japan.
- 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*, Vol. 14, pp. 142-151.
- Cook, N.D., & Fujisawa, T.X. (2006). The psychophysics of harmony perception: Harmony is a three-tone phenomenon. *Empirical Musicology Review*, Vol. 1, pp. 106-126.
- Eberlein, R. (1994). *Die Entstehung der tonalen Klangsyntax*. Frankfurt: Lang.
- Fujisawa, T.X. & Cook, N.D. (2006). How to calculate “harmoniousness” and draw the curves: Dissonance, tension and modality. *Journal of Informatics*, Vol. 25, pp. 35-51.
- Fujisawa, T.X., Konaka, H., & Cook, N.D. (2006). An fMRI study of brain activation in response to two-chord musical progressions (submitted).
- Helmholtz, H.L.F. (1877/1954). *On the Sensations of Tone as a Physiological Basis for the Theory of Music*. New York: Dover.
- Kameoka, A., & Kuriyagawa, M. (1969). Consonance theory: Parts I and II. *Journal of the Acoustical Society of America*, Vol. 45, pp. 1451-1469.
- Kastner, M.P., & Crowder, R.G. (1990). Perception of major/minor: IV. Emotional connotations in young children. *Music Perception*, Vol. 8, pp. 189-202.
- Mazzola, G. (2002). *The Topos of Music – Geometric Logic of Concepts, Theory and Performance*. Berlin: Birkhauser.
- Meyer, L.B. (1956). *Emotion and Meaning in Music*. Chicago: University of Chicago Press.
- Narmour, E. (1990) *The Analysis and Cognition of Basic Melodic Structures*. Chicago: University of Chicago Press.
- Parncutt, R. (1989). *Harmony: A Psychoacoustical Approach*. Berlin: Springer.
- Parncutt, R. (2006a). Commentary on Cook & Fujisawa’s “The psychophysics of harmony perception: Harmony is a three-tone phenomenon.” *Empirical Musicology Review*, Vol. 1, pp. 204-209.
- Parncutt, R. (2006b). Commentary on Keith Mashinter’s “Calculating sensory dissonance: Some discrepancies arising from the models of Kameoka & Kuriyagawa, and Hutchinson & Knopoff.”. *Empirical Musicology Review*, Vol. 1, pp. 201-203.
- Peretz, I., Blood, A.J., Penhune, V., & Zatorre, R. (2001). Cortical deafness to dissonance. *Brain*, Vol. 124, pp. 928-940.
- Plomp, R., & Levelt, W.J.M. (1965). Tonal consonances and critical bandwidth. *Journal of the Acoustical Society of America*, Vol. 38, pp. 548-560.
- Roberts, L. (1986). Consonant judgments of musical chords by musicians and untrained listeners. *Acustica*, Vol. 62, pp. 163-171.
- Sethares, W.A. (1999). *Tuning, Timbre, Spectrum, Scale*. Berlin: Springer.
- Terhardt, E. (1974). Pitch, consonance and harmony. *Journal of the Acoustical Society of America*, Vol. 55, pp. 1061-1069.

Tramo, M.J., Cariani, P.A., Delgutte, B., & Braida, L.D. (2001). The neurobiological foundations for the theory of harmony in Western tonal music. In, R.J. Zatorre & I. Peretz (Eds.), *The Biological Foundations of Music* (Vol. 930, pp. 92-116), New York: The New York Academy of Sciences.