Richard Parncutt

Survey of Recent Research on the Perception of Harmony and Tonality

The 1970s and 1980s saw the consolidation of three influential approaches to the perception of harmony and tonality: the psychoacoustical approach of Ernst Terhardt (1976), the music-theoretical approach of Fred Lerdahl (Lerdahl & Jackendoff, 1983), and the cognitive-empirical approach of Carol Krumhansl (1990a). Inevitably, each approach shed light on a significant subset of relevant issues, while neglecting or failing to account for others. Cross, West, and Howell (1991) have critically reviewed recent progress in the understanding of tonality.[1] They concluded that "the cognitive representation of tonality appears far more structured than purely psychoacoustical accounts would seem to allow, somewhat differently structured from what might be implied by consideration of diatonic relations alone and yet somewhat more dynamic than seems implicit in the cognitive-structuralist account" (p. 238). The aim of the present article is to overview recent research both within and between these areas.

David Huron has continued his research on the perception of harmony and counterpoint. In Huron (1991) he analyzed keyboard scores of J.S. Bach and demonstrated that, while in general favoring consonant intervals, Bach additionally avoids intervals that fuse (octaves, fifths) – consistent with an objective of maintaining the perceptual independence of voices, and suggesting that fusion and consonance are distinct perceptual attributes of sonorities. Huron and Sellmer (1992) investigated the relationship between critical bandwidth and the spacing of chords. Huron (in preparation) summarized the results of these and several other perceptual studies bearing on the conventions of voice leading in tonal music.

An experimental study by Pritschet (1992) aimed to find concrete evidence for the psychological reality of the root of a chord – and failed. The author concluded that neither roughness nor pitch perception can account for the roots of chords as defined in music theory.[2] In a similar vein, Platt and Racine (1990) investigated the main pitch class perceived in triads of pure tones, and found that only some musicians (and no nonmusicians at all) chose the music-theoretical root.[3] Thomson (1993) discussed the definition and origin of harmonic roots, stressing the role of tonal context but nevertheless favoring the context-independent approach of Terhardt.

Relationships perceived between successive sonorities were investigated by Tekman and Bharucha (1992). Sensitivity to mistuning in a chord was enhanced if the chord was preceded by another, harmonically related chord, for inter-onset intervals ranging from 50 to 2500 ms. In the light of earlier results, the authors conjectured that overlapping frequency spectra were probably not directly responsible for the priming effect.[4] Chordal relationships were also investigated and modeled by Parncutt (1993).

The literature on tonality perception has been marked by open and sometimes bitter disagreement between the adherents of the approach of Carol Krumhansl (1990a, b) and her opponents led by David Butler (1990, 1992) and Helen Brown (1992). Brown and Butler have continued to emphasize the importance of temporal aspects of tonality, that is, the order in which the tones of a melody occur — an aspect that had already been considered at some length by Krumhansl and by Bharucha in earlier studies on asymmetrical relationships between successive tones and chords in tonal contexts. While their point is undoubtedly correct, they have not presented a testable algorithm capable of predicting the effect of the order of tones on the tonic of a melody. Butler and Brown have also repeatedly emphasized the importance of rare intervals in the diatonic set – especially the tritone (occurring only once in the major diatonic set). The intervallic rarity theory, however, has serious problems of its own: The harmonic minor set contains two tritones, and chromaticism is the rule rather than the exception in most tonal music of the "common practice" period. Even assuming a clear-cut distinction between diatonic and non-diatonic tones, "the rarity of the tritone may allow a listener to infer the entire diatonic collection from a limited number of pitches; yet with regard to the location of a tonic scale-step, it has nothing to say" (Agmon, 1993, p. 76).

Given that Western tonality is based on harmonic progressions, one might expect the tonality of a melody to be determined by the harmonic progression that it implies. However, no general model for the harmonic progression implied by a melody has yet emerged. Platt and Racine (1994) explored the conditions under which changes of harmony are perceived in simple and somewhat atypical melodies (tones from one triad followed by tones from a different triad). Vos, Geenen and Gremmen (1994) devised an algorithm for the tonality of a melody in which melodic and harmonic aspects are processed in parallel.

Wolfgang Auhagen (1994), in a study not yet widely known in English speaking circles, made a quantum leap in the understanding of the tonality of melodies. On the basis of extensive new experimental data, he concluded that listeners determine the tonality of a melody by first orienting themselves to specific, learned melodic patterns. His results suggest that the tonality of a melody can only be predicted reliably on the basis of an individual listener's accumulated knowledge of melodic-tonal syntax; thus, both the cognitive-structural and interval-rarity approaches can only be valid in special cases. The problem reduces to one of understanding the historical evolution of tonal-harmonic syntax, an area in which Roland Eberlein (1994) has recently made considerable progress.

Neural nets have had a decisive influence on many academic fields in recent years, and the perception of harmony and tonality is no exception. For example, Eberlein (1990) found the neural net to be the most promising model for the processing of complex sounds in speech and music, able to account for a wide range of experimental data related to the perception of musical sonorities. Jamshed Bharucha (1992) described how neural nets can explain many aspects of the learning and subsequent perception of harmonic patterns, determining the roots of chords and the tonality of chord progressions; note, however, that Bharucha's model (with individual nodes corresponding to specific tones, chords and keys specified relative to the chromatic scale) is vastly oversimplified with respect to the complexity of neurophysiological reality. Neural net models can also reliably predict the pitch of individual complex tones (Bharucha, 1992; Laden, 1994; Laden & Keefe, 1991). Neural net accounts of music perception drawing on Kohonen feature maps were developed by Griffith (1993) and Leman (1990). Leman (1994, 1995) presented a schema-based approach incorporating a physiological model of the peripheral auditory system that can account for pitch perception (cf. Meddis & Hewitt, 1991), data-driven learning, and schema-based perception (including location of tonal centers). The model is potentially capable of extracting the harmonic functionality of chords from the audio signal of a harmonic progression.

Bharucha (1994) extended the neural net approach to questions of musical expectation and aesthetics. Melodic expectation in tonal contexts was also the subject of studies by Povel (1994) and Larson (1994). In Povel's study, listeners heard a I-IV-V7-I cadence followed by a single test tone, and sang the tone that they expected would follow. Some subjects focused on the tonal context, resolving a test tone to the nearest tone of the tonic triad, while others tended to ignore context altogether, singing a tone that was a fifth lower or a fourth higher than the test tone. Larson modeled the results of similar experiments by a computer algorithm that assumes the existence of three basic musical "forces": musical gravity (unstable tones tend to descend), musical magnetism (unstable tones tend to move to the nearest stable pitch) and music inertia (tones tend to continue in the same fashion, as in the Gestalt law of good continuation). Larson's algorithm invokes an analogy with classical mechanics as strong as that assumed by Todd (1995) to determine timing and dynamics in musical performance.

Numerous studies have addressed hypotheses formulated by Lerdahl and Jackendoff (1983). Bigand (1990) confirmed procedures for the abstraction of reduced and prolongational structures in a specific set of melodies; and Bigand (1993a) investigated factors underlying musical tension-relaxation schemas. Bigand, Parncutt, and Lerdahl (1996) pitted Lerdahl's, Krumhansl's, and Parncutt's theories against each other in an analysis of tension ratings of chords in short tonal contexts. An overview of psychological research related to tonal and event hierarchies, segmentation, and musical tension was provided by Bigand (1993b). Dibben (1994) demonstrated the psychological reality of Lerdahl-Jackendoff style reductions in tonal music (keyboard music by Handel and Brahms) but was unable to do so for atonal music (piano music by Schoenberg), raising the question of the nature of prolongation in atonal music – a question that had been addressed in more detail by Lerdahl (1989). London (1994) questioned Lerdahl and Jackendoff's "strong reduction hypothesis", replacing it by a "weak reduction hypothesis" that allows for considerable ambiguity and uncertainty in the structural description of musical events.

Järvinen (1993) replicated Krumhansl and Kessler's (1982) key profiles by analyzing frequency-of-occurrence distributions of pitches in jazz improvisations. Profiles were also obtained for melody tones played against individual chords.[5] West and Fryer (1990) obtained Krumhansl-style profiles even when the tones in a melody were presented in random order. Thus, contributions to tone profiles from frequency-of-occurrence distributions may be separable from the effect of temporal order. Lamont and Cross (1994) applied a similar paradigm to children aged 6 to 11 years; their results suggest that children first learn diatonic scalar patterns that are largely independent of serial order, and then gradually internalize hierarchies of tonal stability among scale members that are increasingly dependent on order .

Annabel Cohen's (1991) finding that musicians can recognize the tonic of a Bach Prelude within the first four notes seemed to conflict with intuitions regarding musical form. It would appear that several measures of music are needed to establish a key — consider, for example, the typical length of the first subject of a sonata movement. If no key has been established, then it would be more correct to say that Cohen's listeners recognized a triad, and then assumed that the triad was the tonic (which, of course, is not always the case).[6]

Thompson and Cuddy (1992) studied the perception of interkey distance in Bach Chorale excerpts, and in individual voices extracted from those excerpts. Results for individual voices differed from results for four-part textures. The authors' earlier finding that sharp-side keys were perceived as more closely related to the tonic than flat-side keys was replicated for harmonic progressions, but not for single voices.[7] In a related study, Thompson (1993) advanced a partially hierarchical model of melody, harmony, and key structure in which the contributions of melody and harmony to key need not necessarily be congruent, and in which each layer gives rise to partially independent expectations regarding the future course of the music (see also Povel & Egmond, 1993).

The perception of closure or finality in musical cadences has received considerable recent attention. Rosner and Narmour (1992) asked non-musicians to rate the finality of various textbook cadences in different voicings; they found a strong effect of root progression (notably, V-I was consistently rated more final than IV-I), while other factors (soprano position, inversion, common-toneness) played a subsidiary role. Voicing played a more prominent role in the experiments of Eberlein and Fricke (1992) – largely because their experimental stimuli included progressions that diverged considerably from textbook norms of "good" voice leading. Eberlein and Fricke's experimental stimuli also included cadences from the Middle Ages and Renaissance. Their experimental data suggested that the perception of cadential closure depends primarily on familiarity with existing music. It follows that an understanding of the historical emergence and development of tonal-harmonic syntax is an important prerequisite for an understanding of cadential closure.

The musical advantages and disadvantages of a cognitive approach to harmony and tonality were discussed by Eytan Agmon (1990) and Nicholas

Cook (1994). Agmon stressed the importance of a clear distinction in music-perceptual discourse between physical, perceptual, and cognitive domains, corresponding to the three worlds of Popper and Eccles (1977). He pointed out that enharmonic equivalence, for example, belongs to the Popper's third world, and so cannot be explained in physical or perceptual terms.[8] Agmon also distinguished between theories that attempt to describe psychological reality ("what" theories) from those that describe the processes by which that reality may be constructed on the basis of some external stimulus ("how" theories), claiming that one generally cannot answer the question "how" before first answering the question "what". Cook (1994) provocatively questioned the validity of psychological experimentation for music theory, referring to examples from a range of contemporary sources including Lerdahl and Krumhansl. Like Agmon, Cook also argued for a kind of conceptual separation: "music psychology loses the ability to pose fundamental questions through too ready an acceptance of music-theoretic concepts and categories ... and conversely, music theory risks losing its identity in a welter of incompatible claims regarding both procedures and objectives" (p. 92).

An encouraging trend in recent research on the perception of harmony and tonality has been an increasing willingness on the part of both musicians and scientists to engage in cross-disciplinary collaboration, in spite of the intellectual isolationism, or simply resistance to change, that remains prevalent in certain quarters. It remains to be seen whether the diversity of the current theories will eventually crystallize into a coherent, unified picture worthy of complementing or superseding the music-theoretical edifices of Rameau, Riemann, Schenker.

Summary

Recent research on the perception of harmony and tonality has tended either to consolidate and expand on established approaches such as those of Krumhansl, Lerdahl, and Terhardt, or to add smaller but nevertheless interesting and important details to an increasingly complex picture. Issues of interest have included the perception of consonance and of counterpoint (Huron); harmonic roots (Platt, Pritschet, Thomson); relationships between successive chords (Tekman, Parncutt) and keys (Thompson); the tonality of melodies (Auhagen, Brown, Butler, Cohen, Krumhansl, Vos); the finality of cadences (Rosner) and its dependence on the history of tonal-harmonic syntax (Eberlein); neural-net models of pitch and harmony (Bharucha, Eberlein, Griffith, Laden, Leman); melodic expectation following harmonic cadences (Larson, Povel); the psychological reality of Lerdahl's cognitive structures in music perception (Bigand, Dibben, London); and Krumhansl's key profiles in jazz improvisation (Järvinen), in the minds of children (Lamont), and in the absence of serial order information (West). The creative application of psychological methods and approaches to music theory, and vice-versa, continues to invite careful discussion, and to arouse controversy (Agmon, Cook, Cross).

Footnotes

1 The summary of my own work by Cross et al. (1991, p. 208) is unfortunately somewhat unclear and misleading. For example, they confuse the terms pitch category (applied in Parncutt, 1989 to any degree of the chromatic scale in any register) and pitch class (used in Parncutt, 1988 in its regular music-theoretical sense as an octave-generalized chromatic pitch).

2 Pritchet overlooked two convincing experimental verifications of the psychological existence of the root. The first was the tone profile experiment of Krumhansl and Kessler (1982) following individual major and minor triads constructed from octave-complex tones (replicated by Parncutt, 1993 for a wider range of chords). In each case, the peak of the resultant 12-element profile corresponded to the root. A second verification, using voiced (non-octave-generalized) chords, was described in Parncutt (1989, section 5.3).

3 According to the pitch model of Terhardt, Stoll, and Seewann (1982), the spectral pitches corresponding to the individual pure tones in the sonorities of Platt and Racine (1990) were much more salient than the virtual pitches corresponding to the triads' music-theoretical roots. (The predicted salience of the root pitches would have been greater had the chords been constructed from complex rather than pure tones.) In that light, Platt and Racine were incorrect to consider their experiment as a test of the theory of virtual pitch; in the end, they correctly attributed their data primarily to musical conditioning.

4 This may be regarded as an example of pitch commonality (Parncutt, 1989). The priming chord creates a tonal reference frame (salience versus pitch of each tone sensation within the chord), and the test chord is perceived relative to that frame.

5 Järvinen's chord profiles resemble linear combinations of the chord profiles in isolation (cf. Parncutt, 1988) and key profiles.

6 A comparison of the key profiles of Krumhansl and Kessler (1982) and the pitch-class salience profiles of Parncutt (1988), tested experimentally by Parncutt (1993), suggests that the process of key recognition is not necessarily fundamentally different from the recognition of a triad: The correlation coefficient between tone profile of a key and the pitch-class salience profile of its tonic triad exceeds 0.9 in both major and minor.

7 A possible explanation is that, in harmonic progressions, flats relative to a prevailing key are more salient than sharps (Parncutt, 1989). Flats tend to lie at major-third intervals below diatonic scale-steps, whereas sharps lie at major-third intervals above diatonic scale steps. The lower tone of a major-third interval is more likely to function as the root of a chord, and so is more perceptually salient. This effect is based on Terhardt's theory of virtual pitch, so only applies to tones sounding within sonorities – not to isolated melodic lines.

8 I have doubts about Agmon's description of enharmonic equivalence as "cognitive". It may be argued that enharmonic equivalence is a byproduct of notating chromatically-based music in a

diatonically-based system – a kind of notational artifact. Enharmonic equivalence may only becognitively relevant to musicians who are familiar with the score of the music under consideration.

References

Agmon, Eytan (1990). Music theory as cognitive science: Some conceptual and methodological issues. Music Perception, 7, 285–308.

Agmon, Eytan (1993). Tonicity and the tritone: Beyond the rarity issue. In J. Laaksamo & J. Louhivuori (Eds.), Proceedings of the First International Conference on Cognitive Musicology (pp. 74+87). Jyräskylä, Finland: University of Jyräskylä.

Auhagen, Wolfgang (1994). Experimentelle Untersuchungen zur auditiven Tonalitätsbestimmung in Melodien. Kassel: Bosse.

Bharucha, Jamshed J. (1992). Tonality and learnability. In M. R. Jones & S. Holleran (Eds.), Cognitive Bases of Musical Communication (pp. 213–223). Washington, DC: American Psychological Association.

Bharucha, Jamshed J. (1994). Tonality and expectation. In R. Aiello (Ed.), Musical Perceptions (pp. 213–239). London: Oxford University Press.

Bharucha, Jamshed J. (1992). The emergence of auditory and musical cognition from neural nets exposed to environmental constraints. Paper presented at the 2nd International Conference on Music Perception and Cognition, Los Angeles.

Bigand, Emmanuel (1990). Abstraction of two forms of underlying structure in a tonal melody. Psychology of Music, 18, 45+59.

Bigand, Emmanuel (1993a). The influence of implicit harmony, rhythm and musical training on the abstraction of "tension-relaxation schemas" in tonal musical phrases. Contemporary Music Review, 9, 123–138.

Bigand, Emmanuel (1993b). Contributions of music to research on human auditory cognition. In S. McAdams & E. Bigand (Eds.), Thinking in Sound (pp. 231–277). London: Oxford University Press.

Bigand, Emmanuel, Richard Parncutt, & Fred Lerdahl (1996). Perception of musical tension in short chord sequences: The influence of harmonic function, sensory dissonance, horizontal motion, and musical training. Perception & Psychophysics, 58, 125–141.

Brown, Helen (1992). Cognitive interpretations of functional chromaticism in tonal music. In: M.R. Jones & S. Holleran (Eds.), Cognitive Bases of Musical Communication (pp. 139–160). Washington, DC: American Psychological Association.

Butler, David (1990). Response to Carol Krumhansl. Music Perception, 7, 325–338.

Butler, David (1992). On pitch-set properties and perceptual attributes of the minor mode. In: M.R. Jones & S. Holleran (Eds.), Cognitive Bases of Musical Communication (pp. 161–169). Washington, DC: American Psychological Association.

Cohen, Annabel J. (1991). Tonality and perception: Musical scales primed by excerpts from The Well-Tempered Clavier of J. S. Bach. Psychologische Forschung, 53, 305–314.

Cook, Nicholas (1994). Perception: A perspective from music theory. In R. Aiello (Ed.), Musical Perceptions (pp. 64–95). London: Oxford University Press.

Cross, Ian, Robert West, & Peter Howell (1991). Cognitive correlates of tonality. In: P. Howell, R. West, & I. Cross (Eds.), Representing Musical Structure (pp. 201–243). New York: Academic.

Dibben, Nicola (1994). The cognitive reality of hierarchic structure in tonal and atonal music. Music Perception, 12, 1–26.

Eberlein, Roland (1990). Theorien und Experimente zur Wahrnehmung musikalischer Klänge. Frankfurt am Main: Peter Lang.

Eberlein, Roland (1994). Die Entstehung der tonalen Klangsyntax. Frankfurt am Main: Peter Lang.

Eberlein, Roland, & Jobst P. Fricke (1992). Kadenzwahrnehmung und Kadenzgeschichte – ein Beitrag zu einer Grammatik der Musik. Frankfurt am Main: Peter Lang.

Griffith, Niall (1993). Representing the tonality of musical sequences using neural nets. In: J. Laaksamo & J. Louhivuori (Eds.), Proceedings of the First International Conference on Cognitive Musicology (pp. 109–132). Jyräskylä, Finland: University of Jyräskylä.

Huron, David (1991). Tonal consonance versus tonal fusion in polyphonic sonorities. Music Perception, 9, 135–154.

Huron, David (in preparation). A derivation of the rules of voice-leading from perceptual principles.

Huron, David, & Peter Sellmer (1992). Critical bands and the spelling of vertical sonorities. Music Perception, 10, 129–150.

Huron, David, & Richard Parncutt (in press). An improved theory of tonality perception incorporating pitch salience and echoic memory. Psychomusicology.

Järvinen, Topi (1993). Key profiles in jazz improvisations. In: J. Laaksamo & J. Louhivuori (Eds.), Proceedings of the First International Conference on Cognitive Musicology (pp. 88–103). Jyräskylä, Finland: University of Jyräskylä.

Krumhansl, Carol L. (1990a). Cognitive Foundations of Musical Pitch. London: Oxford University Press.

Krumhansl, Carol L. (1990b). Tonal hierarchies and rare intervals in music cognition. Music Perception, 7, 309–324.

Laden, Bernice (1994). A parallel learning model of musical pitch perception. Journal of New Music Research, 23, 133–144.

Laden, Bernice, & D.H. Keefe (1991). The representation of pitch in a neural network model of musical chord classification. In: P. Todd & G. Loy (Eds.), Musical Applications of Neural Networks. Cambridge, MA: MIT Press.

Lamont, Alexander, & Ian Cross (1994). Children's cognitive representations of musical pitch. Music Perception, 12, 27–56.

Larson, Steve (1994). Musical forces, step collections, tonal pitch space, and melodic expectation. In: I. Deliège (Ed.), Proceedings of the 3rd International Conference for Music Perception and Cognition, Liège, Belgium (pp. 227–229).

Leman, Marc (1990). Emergent properties of tonality functions by self-organization. Interface, 19, 85–106.

Leman, Marc (1994). Schema-based tone center recognition of musical signals. Journal of New Music Research, 23, 169–204.

Leman, Marc (1995). Music and Schemata Theory — Cognitive Foundations of Systematic Musicology. Berlin: Springer.

Lerdahl, Fred (1989). Atonal prolongational structure. Contemporary Music Review, 4, 65–87.

Lerdahl, Fred, & Ray Jackendoff (1983). A Generative Theory of Tonal Music. Cambridge, MA: MIT Press.

London, Justin (1994). The "strong reduction hypothesis" of GTTM. In I. Deliège (Ed.), Proceedings of the 3rd International Conference for Music Perception and Cognition, Liège, Belgium (pp. 261–262).

Meddis, Roy, & M. J. Hewitt (1991). Virtual pitch and phase sensitivity of a computer model of the auditory periphery. I: Pitch identification. Journal of the Acoustical Society of America, 89, 2866–2894.

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

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

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

Platt, John R., & Ronald J. Racine (1990). Perceived pitch class of isolated musical triads. Journal of Experimental Psychology: Human Perception and Performance, 16, 415–428.

Platt, John R., & Ronald J. Racine (1994). Detection of implied harmony changes in triadic melodies. Music Perception, 11, 243–264.

Popper, Karl R., & J.C. Eccles (1977). The Self and its Brain. Berlin: Springer-Verlag.

Povel, Dirk-Jan (1994). Exploring the fundamental harmonic forces in the tonal system. In: I. Deliège (Ed.), Proceedings of the 3rd International Conference for Music Perception and Cognition, Liège, Belgium (pp. 223–224).

Povel, Dirk-Jan, & R. van Egmond (1993). The function of accompanying chords in the recognition of melodic fragments. Music Perception, 11, 101–105.
Pritschet, Hans (1992). Wahrnehmung des Grundtons von Dreiklängen. Regensburg: Roderer.
Rosner, Burton S., & Eugene Narmour (1992). Harmonic closure: Music theory and perception. Music Perception, 9, 383–412.
Tekman, Hasan Gürkan, & Jamshed J. Bharucha (1992). Time course of chord priming. Perception & Psychophysics, 51, 33–39.
Terhardt, Ernst (1976). Ein psychoakustisch begründetes Konzept der musikalischen Konsonanz. Acustica, 36, 121–137.
Terhardt, Ernst, Gerhard Stoll, & Manfred Seewann (1982). Algorithm for extraction of pitch and pitch salience from complex tonal signals. Journal of the Acoustical Society of America, 71, 679–688.
Thompson, William F. (1993). Modeling perceived relationships between melody, harmony, and key. Perception & Psychophysics, 53, 13–24.
Thompson, William F., & Lola L. Cuddy (1992). Perceived key movement in four-voice harmony and single voices. Music Perception, 9, 427–438.
Thomson, William (1993). The harmonic root: A fragile marriage of concept and percept. Music Perception, 10, 385–416.
Todd, Neil P. (1995). The kinematics of musical expression. Journal of the Acoustical Society of America, 97, 1940–1949.
Vos, Piet G., Erwin W. van Geenen, & Frans M. Gremmen (1994). A dual search key-finding algorithm. In: I. Deliège (Ed.), Proceedings of the 3rd International Conference for Music Perception and Cognition, Liège, Belgium (pp. 221–222).
West, Robert J., & Roz Fryer (1990). Ratings of suitability of probe tones as tones after random orderings of notes of the diatonic scale. Music Perception, 7, 253–258.