

This article was downloaded by: [Cornell University]

On: 04 April 2013, At: 14:57

Publisher: Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Journal of New Music Research

Publication details, including instructions for authors and subscription information:
<http://www.tandfonline.com/loi/nnmr20>

The Cognition of Tonality - as We Know it Today

Carol L. Krumhansl^a

^a Cornell University, New York, USA

Version of record first published: 02 Feb 2010.

To cite this article: Carol L. Krumhansl (2004): The Cognition of Tonality - as We Know it Today, Journal of New Music Research, 33:3, 253-268

To link to this article: <http://dx.doi.org/10.1080/0929821042000317831>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.tandfonline.com/page/terms-and-conditions>

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae, and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand, or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

The Cognition of Tonality – as We Know it Today

Carol L. Krumhansl

Cornell University, New York, USA

Abstract

This article begins with a brief summary of how tonality is defined in musicology and some of the tensions that exist between different theoretical traditions. Then, the basic cognitive studies examining the elements of tonality are reviewed as background to some of the more recent literature. The introduction considers the relationship between some of these results and the acoustics of musical tones. The next section reviews studies that take a cross-cultural approach. This is followed by a summary of computational models that attempt to characterize the dynamics of tonality perception, particularly how listeners abstract a sense of key and how it changes over time. This is followed by a description of recent work at the intersection of music theory and music cognition. The final section considers the prospects for studying tonality with brain imaging techniques.

1. Introduction

The term tonality takes on diverse meanings within different musical periods, cultures, and theoretical traditions. Its meaning also depends on the disciplinary approach taken to understanding tonality. Empirical research in cognitive science has identified particular aspects amenable to experimental methodology, particularly concerning the perception and cognition of tonal organization. In addition, computational models of tonality have been developed to analyze music and model human cognition. The convergence with musicological descriptions is sufficiently strong to encourage the view that these empirical observations apply to musical patterns considered significant in music theory, analysis, and ethnomusicology. In other words, music cognition, and particularly the cognition of tonality, is an active, interdisciplinary field of research and scholarship that spans the arts and sciences.

Before reviewing the empirical findings, it would be well to indicate the musical phenomena included under the heading of tonality in musicological accounts. Hyer's (2003) Grove Dictionary entry on tonality will be used as a guide. To begin with its definition, "One of the main conceptual categories in Western musical thought, the term most often refers to the orientation of melodies and harmonies toward a referential (or tonic) pitch class. In the broadest possible sense, however, it refers to systematic arrangements of pitch phenomena and relations between them". To anticipate efforts to extend cognitive accounts beyond Western tonal-harmonic music, we will take the more general definition as the starting point, although the majority of empirical research to date has been concerned with tonality in Western music.

Apart from the tension just alluded to concerning whether or not tonality should be restricted to pitch organization in Western music, Hyer notes a number of other tensions. One is whether tonality can be discussed as separate and distinct from the way it is materially realized in the music itself. Viewing tonality as a separate entity leads to abstractions such as scale, harmonic function, cadence, mode, and key. In a parallel way, cognitive research has sometimes based stimulus materials on such abstractions, for example, presenting scales and chord cadences; other research has employed actual musical pieces or excerpts with greater complexity.

Musicological discussions of Western tonal-harmonic music emphasize harmony as fundamental to its organization, but different theoretical traditions exist. One tradition identifies harmonic functions, such as tonic, dominant, and subdominant, allowing multiple chords (not just the basic triads) to represent these functions. Another tradition identifies a chord by the scale degree on which it is built, analyzing the music for constituent tones and assigning Roman numeral notation. Another tension is whether these harmonies, however theorized, control melodic processes, or

whether they are the result of melodic processes. As required by concrete operationalization of musical questions in experiments, cognitive research has investigated melodic phenomenon with monophonic tone sequences, although these may well have harmonic implications. Other research utilizes chord sequences without an explicit melody, although a melody may arise from the chords. Despite the difficulty of separating melody from harmony in practice, different phenomena are discussed in reference to tones and chords depending on the materials used in the experiment or model. Harmony will be de-emphasized in the present context as it is particularly characteristic of Western music.

A final tension is whether tonality refers to objective properties of the music, or whether it is some aspect of the cognitive experience of listeners. Cognitive questions arise that cannot be determined by considering the music alone: whether sounded musical events are perceived in the categories of tonal theory (e.g., as the dominant function), the role of tonality in guiding listeners' expectations, the perception of dissonance/consonance and tension/release in music, and the possibility of using spatial metaphors for describing not only musical properties (e.g., remote modulation) but also mental representations. These themes will be revisited as six different approaches to studying tonality are considered: cognition, acoustics, cross-cultural comparisons, computational modeling, music theory, and brain science.

2. Tonality and cognition

Music cognition emerged as a topic of scientific investigation at approximately the same time that cognitive science began to be conceived as an interdisciplinary research area. Cognitive science refers to a rather loose affiliation among a number of established disciplines, including cognitive psychology, philosophy of mind, linguistics, artificial intelligence, and neuroscience. Because of this development, the study of music cognition has from its beginning drawn on vocabularies, theories, and methods from a variety of disciplines.

An important early influence was the suggestion that musical structure could be described by analogy with language. That is, structural representations of natural language could be applied to music. In a seminal paper, Lindblom and Sundberg (1970) argued, "... the research problem can in our opinion be formulated in the following fashion: Given a certain well-defined class of melodies what are the principles and laws by means of which the metric, harmonic, and tonal facts of these melodies can be derived?" In answer to this, they proposed a new kind of theory of music that accounts for musical behaviors (composition, performance, perception) in the same way that Chomsky approached language.

The initial test of this idea was to apply rules of linguistic prosody to a set of Swedish nursery tunes (Lindblom & Sundberg, 1970). They noted that the procedures for deriving stress or prominence contours were similar in the two

cases. The system of prominence rules provided a framework for formally analyzing the metric, harmonic, and melodic aspects of the music. In an application of the analysis-by-synthesis principle, they then used the rules to generate novel but stylistically similar melodies. Sundberg and Lindblom (1976) subsequently produced stylistically convincing variations of a lullaby folk melody with a similar set of grammatical principles.

This was followed by an extensive theory applying linguistic concepts to music, the *Generative Theory of Tonal Music* (Lerdahl & Jackendoff, 1983). It proposed four kinds of structural representations: grouping (partitioning the music into embedded temporal segments), meter (periodic patterns of stress), and two kinds of reduction: the time-span reduction (which assigns a tree structure specifying the relative dominance of each surface event) and the prolongation reduction (also a tree, describing the patterns of harmonic tension and relaxation). An extension of this theory is Lerdahl's (2001) *Tonal Pitch Space* will be described in more detail later in connection with empirical tests of the theory.

At the same time, a number of cognitive psychologists were beginning to conduct experiments on how tonal organization affects the way music is encoded and remembered. Among studies concerned with scale structure, Dowling (1978) found that listeners had difficulty distinguishing between exact transpositions of melodies (a constant shift in pitch, keeping the intervals constant but making some tones nondiatonic) and tonal answers (a shift in pitch that maintains contour but alters some intervals to keep all the tones diatonic). Similarly, Dewar et al. (1977) and Cuddy et al. (1979) found that listeners have better memory for tones in sequences with a well-defined tonality. These studies suggest that diatonic scale structure is used by listeners as a framework for encoding and remembering pitch sequences.

In an effort to quantify the hierarchy of tonal stability, Krumhansl and Shepard (1979) introduced the probe tone technique. In this study, an incomplete C major scale (without the final tonic, C) was sounded in either ascending or descending form. This context, intended to establish the key of C major, was followed on successive trials by each of the chromatic scale tones in the next octave (the probe tones). Listeners rated how well each tone completed the scale. The more musically trained listeners produced a pattern that would be expected from music theory: the tonic was most expected, followed by the fifth, third, the remaining scale tones, and finally the nondiatonic tones. Krumhansl and Kessler (1982) extended this method with a variety of key-defining contexts, including chord cadences, and both major and minor scales. The results for C major and A minor contexts are shown in Figure 1.

In a related paper, Krumhansl (1979) demonstrated influences that tonal structure has on pitch perception and memory. One experiment presented pairs of tones following various key-defining contexts (scale or triad chord). Subjects were asked to judge how similar the first tone is to the second

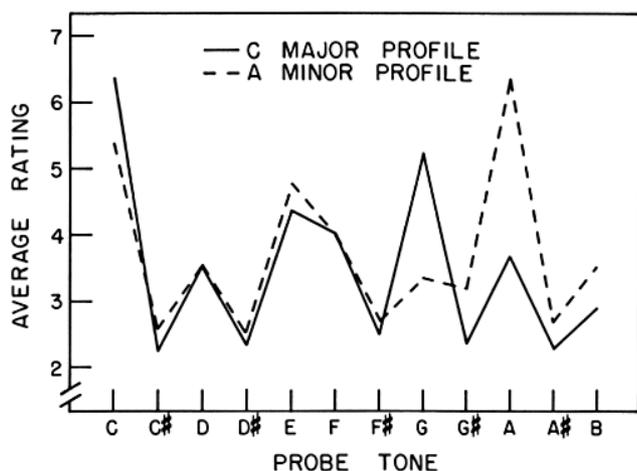


Fig. 1. The probe tone judgments for C major and A minor contexts (from Krumhansl & Kessler, 1982). Listeners rated how well the probe tone fit with key-defining contexts (scales, tonic chord, and chord cadences).

in the context provided. Listeners perceived a complex pattern of interrelationships among the individual tones. Diatonic tones were judged as more closely related to each other than they were to nondiatonic tones. The tones comprising the major triad chord formed a set of particularly closely related tones. The data also contained a regular pattern of asymmetries. Tones less closely related to the tonality (less stable tones) were judged as more similar to tones central to the tonality (more stable tones), than the reverse. Thus, the empirical results supported music-theoretic descriptions of the tendency of less structurally stable tones to move toward the tonic and its closely related tones. Similar asymmetries were also demonstrated in memory performance: nondiatonic tones were more often confused with diatonic tones than the reverse. These temporal-order asymmetries reflect the dynamic character of musical tones, a theme that will be developed later.

Krumhansl and Kessler (1982) used the probe-tone data (or profiles) to obtain a geometric map of musical keys. This analysis was based on the assumption that closely related keys have similar tonal hierarchies. In the first step, the profiles for all pairs of major and minor keys were correlated to give a quantitative measure of key distance. These correlations were then analyzed using multidimensional scaling. This produced a solution in four dimensions in which the 24 major and minor keys fell on the surface of a torus. Neighboring keys were related by circle of fifths and parallel and relative major/minor relationships. This representation corresponded well with music-theoretic description of key distances. Krumhansl and Kessler (1982) then used this map to trace how the sense of key develops and changes over time. Trained musicians rated probe tones presented after each successive chord in nine-chord sequences, some of which contained modulations between keys. These results were projected onto the key map. The sense of key developed

rapidly, and shifts to closely related keys were assimilated more readily than shifts to distantly related keys.

These particular results are described here because they will be relevant to what follows. But, it should be emphasized that from the 1980s onward music cognition research expanded rapidly to cover a wide range of musical phenomena. Quite a few studies were done on the perception of harmony, and its effects on musical memory and processing. Other experiments tested whether the results obtained with these abstract elements also applied to excerpts from actual musical pieces. Analogs of the results for tonal-harmonic music were considered in other styles, including twentieth-century polytonality and atonality. Theoretical proposals, particularly those of the Lerdahl and Jackendoff (1983) model, were tested. The literature expanded on the acquisition of musical skills and effects of musical training. An important empirical focus was on how performers' interpretations of a musical piece influenced their performance, particularly timing parameters. And, simultaneous with the research on the perception of pitch organization, a rich literature emerged on musical meter and rhythm. Other developments will be mentioned in the sections that follow that focus on tonality in particular.

3. Tonality and acoustics

As rapidly as music cognition has grown as an area of empirical research, it is young and undeveloped compared with music acoustics. Theorizing about musical pitch, and its possible natural basis in acoustics, has a very long and distinguished history. At the core of the theorizing is the belief that numbers provide a suitable vehicle to move between the realms of physics, psychology, and music. This belief is often attributed to Pythagoras (ca. 570–500 BCE), who associated harmonious-sounding musical intervals with vibrating strings whose lengths form simple numerical ratios (e.g., 2:1, 3:2), although the essence of the theory may have been developed earlier by the Chinese.

In his nineteenth-century treatise, *On the Sensations of Tone as a Physiological Basis for the Theory of Music* (1885/1954), Helmholtz set out a program of research that sought correspondences between physical acoustics of vibrating bodies, human auditory physiology and perception, and musical structures such as scales and harmonies. As shown in Figure 2, the tones used in Western music bear a close relationship to the harmonics of complex tones. The tones of the diatonic scale appear, except the seventh scale degree is mistuned. The intervals making major and minor triads (fifths and major and minor thirds) have small numerical ratios of frequencies. And, the circle of fifths – describing the relationship between keys – is based on the interval of a fifth, with a ratio of 3:2.

In addition to these elements of music theory, the perceptual dimension of consonance and dissonance is related to ratios of frequencies. In perceptual judgments, the unison

Harmonic Series

Harmonic	Frequency	Nearest Tone	Interval Formed
9	2358	D ₇	Major Second M2, 9:8
8	2096	C ₇	
7	1834	-Bb ₆	Minor Sixth m6, 8:5
6	1572	G ₆	
5	1310	E ₆	Major Sixth M6, 5:3
4	1048	C ₆	
3	786	G ₅	Perfect Fifth P5, 3:2
2	524	C ₅	
1	262	C ₄	Octave 2:1

Fig. 2. The first nine harmonics of middle C. The frequencies and nearest tones are indicated, as well as intervals described as elements of Western tonal-harmonic music.

(1 : 1) and octave (2 : 1) are the most consonant, followed by the perfect intervals, the perfect fifth (3 : 2) and the perfect fourth (4 : 3). The major third (5 : 4), minor third (6 : 5), major sixth (5 : 3), and minor sixth (8 : 5) are next most consonant. The least consonant are the minor second (16 : 15), the major seventh (15 : 8) and the tritone (45 : 32). The ordering along the dimension of consonance corresponds quite closely to the size of the integers in the ratios (Vos & van Vianen, 1984). An acoustic basis exists for this relationship. When the ratios are simple, the harmonics of the two tones either coincide or are sufficiently distant not to interfere. Interference between tones occurs when the ratio of the frequencies is less than 1.2 (a minor third).

The next question, then, is whether the tonal hierarchy judgments (such as those of Krumhansl & Kessler, 1982) can be accounted for by acoustics. Krumhansl (1985; 1990, Table 3.1 and Figure 3.1) compiled six studies of consonance that used both perceptual and computational measures of consonance. For the major key data, there was a good match with the consonance of the tones with the tonic. For minor keys, however, the fit was less good. The minor third was rated higher than the consonance values would predict, and the major third was rated lower. In addition, the major sixth was rated lower than the consonance prediction. These discrepancies suggest more is involved.

Another effort to relate the tonal hierarchy to music acoustics was made by Parncutt (1988, 1989, 1994; Huron & Parncutt, 1993). He extended the algorithm of Terhardt et al. (1982a,b), which predicts pitch salience for complex tonal stimuli. Pitch salience is assumed to be a combination of spectral and virtual pitches. As outlined in Huron and Parncutt (1993), Terhardt's model has three stages. The first stage analyzes the signal and determines the audibility of the spectral components (the pure tone components of the spectrum). The second stage determines the fundamentals that correspond to these harmonics. The spectrum is scanned for the presence of frequencies that approximate the harmonic

series of a tone. These tones are called virtual pitches. The virtual pitches are weighted according to how well the spectral components match its harmonic series or harmonic "template". Stage three combines the results of the first and second stages, allowing for the possibility that the relative weights of spectral and virtual pitches may vary depending on the mode of listening and other factors.

In one application, Parncutt (1994) found that this approach could account quite well for the results of Krumhansl and Kessler (1982). He applied the model to the IV-V-I, II-V-I, and VI-V-I cadences used in that study. The correlations with the listeners' judgments were reasonably high. Importantly, they were higher than the correlations between the tone distributions (coding simply the presence or absence of the tones) and listeners' judgments. This finding argues against the idea that the probe-tone data simply reflect the tone distributions in the contexts, and suggests that virtual pitches contribute to the judgments. Huron and Parncutt (1993) combined the pitch salience approach with a sensory memory decay. This model matched quite closely the probe tone judgments of the modulating chord sequences of Krumhansl and Kessler (1982).

Recently, Leman (2000) took an approach that is rather similar in that it sought to account for the Krumhansl and Kessler (1982) data with a model that incorporates virtual pitches and short-term memory. The input to the model is acoustic data, which is processed by a peripheral auditory model that simulates the filtering of the ear and then analyzes for periodicity pitches. The summed "completion image" is similar to the virtual pitch model just described. This pitch module is then entered into an echoic memory module, which incorporates both integration and decay over time.

Applied to the chord cadence contexts of Krumhansl and Kessler (1982), the model produced tonal hierarchies that are similar to the probe-tone profiles. From this, Leman concluded that the probe tone data could be accounted for by short-term memory for the perceptually immediate context. It should be noted, however, that the model also extracts virtual pitches which, together with spectral pitches, Parncutt (1994) had already shown match major and minor tonal hierarchies quite well. Thus, the conclusion concerning the role of short-term memory can be questioned. Other limitations of this approach include poor results for the major and minor scales and less consistent results across the different contexts than holds in the empirical data. An important strength of this model, however, is that it takes as input acoustic data rather than some symbolically coded representation of the sound.

Another possibility is that the probe tone judgments reflect musical experience. As a way to summarize something about this experience in a simple quantitative form, frequency distributions of tones in note-count studies by Youngblood (1958) and Knopoff and Hutchinson (1983) were used (Krumhansl, 1985, 1990). A good match was found between the probe tone data and how frequently each degree of the scale is sounded in the wide range of eighteenth- and nineteenth-century compositions analyzed

in these studies. Unlike consonance, the fit is also fairly good for compositions in minor keys. Thus, although the probe tone ratings bear some resemblance to consonance, they appear to be more closely related to how tones are distributed in music.

A general problem with all the approaches using values for single tones is that they do not take into account the resolution tendencies of tones. As noted earlier, these tendencies are found experimentally in both similarity judgments and memory confusions. One computational model that captures these tendencies was proposed by Tillmann et al. (2000). The neural network model contains three layers, corresponding to tones (pitch classes), chords, and keys (only major keys are represented). Different training sets were used: simple (idealized) harmonic sequences and more realistic harmonic sequences. They were either “sparsely coded” (just the presence of tones was coded) or “richly coded” (including Parncutt’s, 1988, measure of pitch salience). The resulting model reproduced a hierarchy of tones similar to that of Krumhansl and Kessler (1982), and was also able to match both the direction and magnitude of temporal-order asymmetries found in Krumhansl (1979). Other research on tonal implications will be described later.

The approaches discussed so far are also limited in that they do not address the wide range of tonal phenomena found across cultures. If tonal organization is rooted in acoustics and auditory physiology, then it should be largely invariant across the musics of the world. The diversity of tonal structures that exists suggests that experience with particular musical styles makes significant contributions to its cognitive representation. As noted by Helmholtz:

The construction of scales and of harmonic tissue is a product of artistic invention, and by no means furnished by the formation or natural function of our ear, as it has been hitherto most generally asserted. Of course the laws of the natural function of our ear play a great and influential part in this result; these laws are, as it were, the building stones with which the edifice of our musical system has been erected, and the necessity of accurately understanding the nature of these materials in order to understand the construction of the edifice itself, has been shewn [sic] by the course of our investigation upon this very subject. But just as people with differently directed tastes can erect extremely different kinds of buildings with the same stones, so also the history of music shews [sic] us that the same properties of the human ear could serve as the function of very different musical systems. (Helmholtz, 1885/1954, pp. 365–6)

4. Tonality and culture

One approach to studying how experience influences the cognition of tonal structure is to make cross-cultural comparisons. This approach may also be informative about how musical knowledge is acquired by determining, as a kind of base-line, the responses of naïve listeners. In this approach, the typical design uses a style that would be familiar to some listeners but not others. This has the advantage that it is pos-

sible to use listeners with the same amount of musical training, but different degrees of exposure to the musical style. Work in this area is also important from the point of view of extending our understanding of tonal cognition beyond Western tonal-harmonic music. However, conducting research of this kind requires understanding the various musical styles and their cultural context sufficiently well to construct appropriate materials and experimental methods.

In an early experiment, Castellano et al. (1984) used the classical music of North India, with Indian-trained musicians and Western musicians not familiar with the style. They first heard an extended passage from each of 10 *rāgā*, which the Indian listeners were generally able to identify. Then they heard short themes in a probe tone study. The *rāgā* were based on different underlying scales and this was reflected in the probe tone ratings, as was the stability of the *Sa* and *Pa* (the first and fifth scale degrees) and the *vādi* tone which is particular to each *rāgā*. What was surprising was that the two groups of listeners produced remarkably similar results. What could account for this similarity? Analysis of the music showed that the ratings of both groups of listeners reflected the relative emphasis given the tones in the contexts. That is, tones rated higher tended to be sounded more frequently and with longer durations. This was sufficient for the Western listeners to give responses consistent with music-theoretic descriptions of North Indian classical music.

These findings suggest that listeners have strategies for abstracting pitch structures from unfamiliar musical styles. This is in contrast to the general assumption that musical knowledge develops gradually over years. To what extent is the cognitive system flexible and adaptable? Oram and Cuddy (1995) addressed this question by presenting sequences based on either a diatonic tone set or a non-diatonic tone set (which did not conform to any major or minor scale). Within each sequence, the frequency of occurrence was controlled according to the following ratios: one tone of the tone set occurred eight times, two tones occurred four times each, and the remaining four tones occurred just once. Probe tones presented following the sequences were systematically rated according to the frequency with which the tone had appeared in the sequence. An interesting result of this study is that the effect of frequency was more pronounced for musically trained listeners than untrained listeners. Thus, extensive musical experience does not appear to restrict listeners to familiar tonal hierarchies. Rather, training appears to enhance sensitivity to the tone distributions.

In subsequent experiments reported by Cuddy (1997), composers wrote melodies that were constrained as to the total durations that were to be allotted to each scale tone. Again, novel tone sets were used. The durational emphasis influenced probe tone ratings, as would be expected from the previous results; tones sounded for longer durations were rated higher. In addition, a double-probe technique was used with a pair of probe tones following the melody. In this, listeners were asked to judge how related the first tone is to the second tone of the pair with respect to the melody. These

Table 1. Judgments of melodic continuations correlated with statistics of musical corpus.

	Spiritual folk hymns		North Sami <i>yoiks</i> (<i>yoiks</i> familiar to Sami)		
	Finnish musicians	Beseecher Choir	Western musicians	Finnish musicians	Sami <i>yoikers</i>
Tone distributions	0.64	0.72	0.65	0.78	0.67
Two-tone transitions	0.50	0.56	0.69	0.60	0.54
Three-tone transitions	0.41	0.59	0.49	0.46	0.65

judgments of relatedness depended on order in a way that could be predicted by the tones' duration in the melody. Relatedness was higher when a tone of less total duration was followed by a tone of greater total duration than the reverse. This is as would be expected given the temporal-order asymmetries discussed earlier.

Recent studies have also investigated how melodic expectancies are influenced by statistical distributions of tones in distinctive musical styles: Finnish spiritual folk hymns (Krumhansl et al., 1999), and North Sami *yoiks* (Krumhansl et al., 2000). In these studies, listeners were presented with short excerpts followed by possible continuation tones. The continuation tones were rated for how well they fit with their expectations for continuation. The question was how well these judgments could be accounted for by the frequency of tone occurrence, and the frequency of two- and three-tone transitions.

The spiritual folk hymn study compared melodic continuation judgments for Finnish musicians and members of the Beseecher choir who maintain the hymns as part of their religious tradition. As seen at the top of Table 1, the two groups differ primarily in the extent to which their judgments were influenced by three-tone transitions typical of the style. For the study with *yoiks*, the listeners were Western musicians, Finnish musicians (who had studied a few *yoiks* in their ethnomusicology course), and Sami *yoikers*. Again, a difference appeared primarily for the three-tone transitions. This suggests that musicians unfamiliar with a style are sensitive to the lower-order statistical properties of the style, but their ability to use distributional information declines for the higher-order statistics. In a related finding, Krumhansl (2000) showed that the higher-order statistics were most informative in classifying excerpts from the two musical styles.

Another approach to studying the effects of culture is to look through the lens of a particular theoretical proposal. Narmour's (1990) implication–realization model of melodic expectancy has been used for this purpose in a number of studies. The model describes a core set of five principles governing melodic expectation. These are based on Gestalt principles, presumed to be innate and universal. They are expressed precisely in musical interval sizes, permitting quantification of predictions for experimental tests. Briefly, the principle of Registral Direction states that if the implicative interval (the interval engendering expectation) is small

(perfect fourth or smaller), the direction of the melody is expected to continue; if it is large (perfect fifth or larger) then it is expected to reverse direction. The principle of Intervallic Difference states that small implicative intervals imply a similarly sized realized interval (the next interval), whereas large implicative intervals imply smaller intervals. The principle of Registral Return is an expectation to return to the pitch region around the first tone of the implicative interval. The principle of Proximity is a general preference for small intervals. The principle of Closure says that closure is strongest when the implicative interval is large and the realized interval is smaller and reverses direction.

The first test of the model (Krumhansl, 1995) used British folk tunes (with musicians and nonmusicians), atonal songs (with musically trained and untrained listeners), and Chinese melodies (with Chinese and Western listeners). In support of the theory's claim for generality, similar results were found for expert and non-expert listeners for all three musical styles. Most of the principles made statistically significant contributions to the judgments of the continuation tones. In addition, the coding of the tonal hierarchy of the style (with local tonal implications used for the atonal songs) added to the predictions of the model. Finally, there was considerable consistency across the groups of listeners. This suggests that the implication–realization model is identifying general principles of music cognition that operate independently of the listener's musical experience. In other words, they are characterizing basic psychological responses that do not depend on specialized training, extensive experience with the particular musical style, or knowledge of technical concepts or vocabulary.

In the various studies taking this approach, one of the largest effects of cultural background was found in the study of North Sami *yoiks* (Krumhansl et al., 2000). As shown in Figure 3, the Sami and Finnish musicians (who were trained on a few *yoiks* in their ethnomusicology course), gave similar weights to the five principles of the implication–realization model. However, the Western listeners unfamiliar with the *yoiks* gave stronger weights to three of the principles previously shown to be relatively strong in Western music. Unlike most Western melodies, the *yoiks* have relatively frequent large consonant melodic intervals. This is reflected in the high weights given to consonance for the Sami and Finnish listeners. Even the Western listeners adapted to this to some degree. It would seem, then, that a preference for large con-

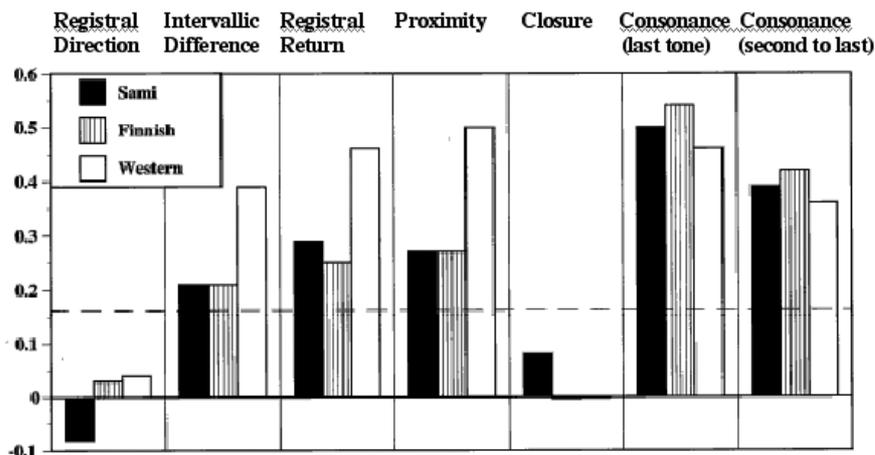


Fig. 3. The weights given to the five principles of Narmour's (1990) implication–realization model and consonance with the last and second to last tones (from Krumhansl et al., 2000). Three groups of listeners (Sami *yoikers*, Finnish musicology students, and Western musicians) made judgments of how well continuation tones fit with their expectations.

sonant intervals is latent in the Western listeners, but is not activated by Western melodies which tend to contain smaller intervals. When the style, however, emphasizes large consonant intervals, then this becomes a factor used by listeners in generating their expectations.

An even larger effect of cultural background was found in an extension of this work by Eerola (2003). The *yoiks* were presented to traditional healers from South Africa. The results again replicated the robustness of data-driven models (an auditory memory model and a frequency-based model) as well as the implication–realization model (Narmour, 1990). However, the South African listeners' responses did not conform to a model of Western schematic knowledge or a model developed specifically for the *yoiks*. In general, their responses were least like those of the Sami musicians.

In sum, quite a few studies have now found more versatility in responding to novel styles than would be expected given the more traditional view that cognitive representations require extensive experience for their internalization. A recent article, however, suggests some interesting limitations. Ayari and McAdams (2003) undertook an extensive perceptual analysis of Arabic improvised instrumental music (*taqsim*). Their listeners were of European and Arabic cultural origins, including some highly trained musicians in both groups. The listeners performed an intensive series of tasks: identifying musical elements, segmenting the musical piece, verbal description of significant elements, and melodic reductions of the identified elements. Although Western listeners were able to point to pivot or resting tones, only the Arab listeners were able to identify and produce reductions that reveal the nature of the underlying Arabic modes. They conclude

... it is the professional musicians who have developed, by way of explicit learning and implicit experience, the dimensions and perceptual skills necessary to appreciate what is incorporated in this music in terms of nuance and expression, at least as con-

cerns their ability to perform explicitly musical analytic tasks. (p. 213)

Before leaving the topic of how musical experience influences tonal cognition, brief mention should be made of studies considering the effects of development. This area has recently emphasized the predispositions infants have for music and music-like patterns (Papousek & Papousek, 1981; Trehub, 2000). For example, infants prefer consonant to dissonant music (Zentner & Kagan, 1996), and remember consonant intervals such as octaves and fifths better than dissonant intervals (Trainor, 1997). Infants attend longer to music in which pauses are inserted between phrases rather than in the middle of phrases (Krumhansl & Jusczyk, 1990; Jusczyk & Krumhansl, 1993), and may use octave simultaneities as one cue to phrase endings. Finally, most relevant to the above discussion of the effects of tone distributions, infants appear sensitive to sequential probabilities in tone sequences (Saffran et al., 1999). In this method, infants first listen to long sequences of tones in which some tones tended to co-occur more frequently than others. After this, they are able to discriminate sequences with more likely sequential probabilities from those with less likely sequential probabilities. That similar results have been found for sequences of phonemes suggests that early in development humans are sensitive to frequent successions of sounds, and this sensitivity may encompass both language and music.

5. Tonality and computational modeling

An area that has attracted considerable computational modeling is tonality induction or key-finding. This refers to the process through which a listener initially orients to the key of a piece of music and subsequently reorients to new keys or modulations (changes of key) that occur in the music. Various empirical results suggest that listeners can quite

accurately judge the key based on short excerpts (Cohen, 1991) and track changes of key when they occur (Krumhansl & Kessler, 1982; Thompson & Cuddy, 1989, 1997). The question considered here is whether a model can be devised that automatically analyzes the music and gives results that are consistent either with behavioral measures or the notated key signature. Key-finding algorithms take some musical input, which may be acoustic information or symbolically coded music. The input might be an entire piece of music, a short segment from the initial or final portion, or a sliding window of some length (e.g., one measure).

An early model of key-finding (Longuet-Higgins & Steedman, 1971) works by examining the musical sample for whether or not the tones are contained in the diatonic major and minor scales. It progresses through the musical sample, eliminating keys if the sample contains tones that are nondiatonic in the key. A tonic-dominant rule is invoked if either all keys are eliminated at some point, or if the end of the sample is reached and more than one candidate key remains.

Krumhansl and Schmuckler (Krumhansl, 1990) constructed an algorithm that matches (by correlation) the distribution of tone duration in the sample (the total duration of each tone in the chromatic scale) to the tonal hierarchies (Krumhansl & Kessler, 1982). The algorithm outperformed the Longuet-Higgins and Steedman (1971) algorithm when applied to Bach's *Well-Tempered Clavier*. It also conformed quite well to experts' judgments of modulations in one Bach prelude. For comparison, these results were projected onto the toroidal representation of musical keys (Krumhansl & Kessler, 1982). The algorithm performed less well on short initial excerpts of Chopin's preludes. One feature of this model is that it allows for the possibility that a number of keys might be quite strongly suggested simultaneously (each having relatively a relatively high correlation), or that no key is suggested (no key has a high correlation).

Two variants of this algorithm have been proposed recently. Temperley (1999) suggested a number of modifications. One modification increases the weight of the seventh scale degree in major and the raised seventh degree in minor. The algorithm ignores duration, and a matching formula other than correlation is used. A penalty is imposed for changing key from one input segment to the next. The purpose of the penalty was to reduce the key instability sometimes observed during modulations or for short excerpts. Finally, a retrospective reevaluation of key is permitted. The algorithm performed well in the most extensive test to date using a large corpus of key-annotated pieces that were analyzed on a measure-by-measure basis.

Shmulevich and Yli-Harja (2000) proposed an algorithm directed at smoothing local oscillations in key assignments, which were also a concern addressed by Temperley (1999). This algorithm did this by using neighboring key assignments. The input was a set of overlapping sliding windows of a fixed number of tones. There was one window for each tone of the music, so that the model traces changes of key in detail. The key assignments were then smoothed using a

graph-based method. The graph contains 24 nodes, corresponding to the major and minor keys. The edges were assigned lengths equal to the distances between keys found by Krumhansl and Kessler (1982). The final key assignment is the norm of the keys, that is, the key that is closest to all possible key assignments.

Another influential approach to key-finding uses neural networks. Leman's (1995) neural network model takes acoustic information as input, coded as described earlier in the section on acoustics. The patterns formed by the input self-organize into a stable perceptual schema. The resulting information then enters a self-organizing network that consists of cognitive schemas for pitch structures. These include a schema for tone-center perception (or key-finding) that has resulted from long-term learning. It yields a measure of key distances that is similar to that of Krumhansl and Kessler (1982; see also Leman & Carreras, 1997). Thus, the model derives a sense of the key based on a combination of the current information and previous tone-center perception. The model also allows for retrospective (retroactive) re-evaluation of tone centers. Another model also already described in the section on acoustics is the neural network model of Tillmann et al. (2000). Their model, which only represents major keys, was applied to the three sequences in major keys used by Krumhansl and Kessler (1982). The results approximate the listeners' sense of key found in the perceptual study.

Toivainen and Krumhansl (2003; see also Krumhansl & Toivainen, 2001) used a self-organizing map, or SOM, to represent how the sense of key develops and changes over time. The SOM consists of a two-dimensional map of 24×36 units, specified in advance to be a torus based on the earlier results of Krumhansl and Kessler (1982). Each unit has a vector of adaptable connection weights to 12 input units (corresponding to the tones of the chromatic scale). The SOM was trained with the 24 major and minor probe tone profiles (Krumhansl & Kessler, 1982). First, one of the 24 keys was selected randomly. The unit whose connection weights best matched that key's profile was determined. Its connection weights were adjusted to better match the input profile, as were those of its neighbors on the map – but to a lesser degree. Then this process is repeated. After training, the SOM had the configuration shown at the top of Figure 4 where the letters label the units with the best match to each key's probe tone profile. The positions of the keys are easily interpretable: each major key is flanked by its neighbors on the circle of fifths and its relative and parallel minor keys. The SOM gives essentially the same result as the map derived from multidimensional scaling (Krumhansl & Kessler, 1982).

The technique used in the behavioral experiment is called the concurrent probe tone task. The music is sounded in the right ear, in this case, a Bach Organ Duetto, *BWV 805*. Simultaneously, in the left ear, a probe tone is sounded. It is one of the 12 tones of the chromatic scale sounded in multiple octaves. It was pulsed slightly every half-measure to keep it from blending in with the music. The listeners judged how

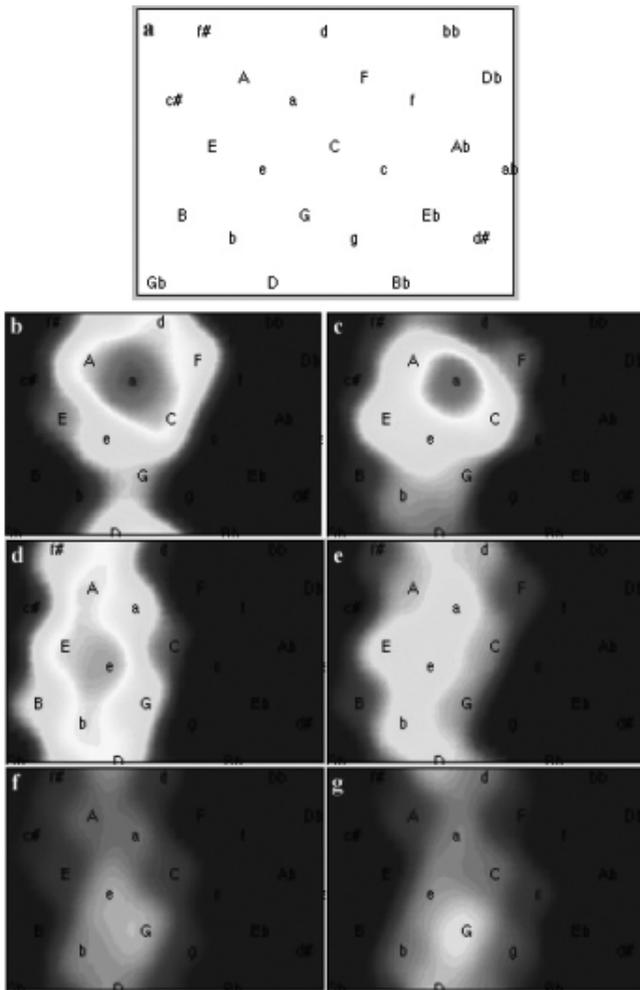


Fig. 4. Above: the self-organizing map (SOM) that was trained with the probe tone data of Krumhansl and Kessler (1982). The continuous probe tone data (left) and model predictions (right) from Toiviainen and Krumhansl (2003) are shown mapped onto the representation at three points in time during Bach organ duetto *BWV 805*.

well the probe tone fits with the context in a musical sense. They make this judgment by moving a slider as frequently as they wished to show how the degree of fit changes over time. Then, the listener repeats this with each of the other probe tones until all twelve probe tones have been presented on successive hearings of the piece.

We were interested in testing how well the perceptual judgments could be accounted for by a computational model that automatically analyzes the music. The model consists of a perceptual component, p . When a tone is sounded, the value for that tone reaches a maximum, and then decays until tone offset. The memory component, m , integrates the perceptual component over time, with a decay function that corresponds approximately to empirical estimates of the duration of the short-term acoustic store. These values could be compared with the listeners' judgments directly, or both could be projected onto the SOM as described next.

The human data are the ratings of each of the 12 probe tones at each point in time. The degree to which the human data match each unit on the SOM can be represented by shading, as shown on the left. The model predictions are the memory strength of each of the 12 tones at each point in time. These can also be mapped onto the SOM as described above and shown on the right. The top pair of panels in Figure 4 shows a point in time when both humans and model find a strong tonality of A minor, the key of the piece. The second pair of panels shows a point in time when the sense of key is more diffuse, especially for the model. The last pair of panels shows a point in time when neither the human data nor the model finds any key to be strong. A dynamic representation of the results, accompanied by sound, can be found at <http://www.perceptionweb.com/misc/p3312>.

The strength of each key over time for the human data and the computational model are remarkably close, suggesting that the process of tonality induction is driven very much by the tones sounded in time coupled with a short-term memory process. However, the dynamic representation also shows that listeners (who were all well-trained musicians) sometimes find the key more rapidly and with more certainty. This indicates that their internalized representations of compositional principles in the musical style are leading them to anticipate and consider salient certain of the sounded events. The listeners are apparently using some information that is not coded by this algorithm.

A second algorithm was devised to consider whether listeners might be using the temporal order of tones to refine their sense of key. The way in which tone transitions were coded was based on principles of auditory stream segregation (Bregman, 1990). The strength of each transition in the model depends on three factors: pitch proximity, temporal proximity, and the duration of the tones. More specifically, a transition between two tones has the highest strength when the tones are proximal in both pitch and time and have long duration. In addition, this model also contains dynamic perceptual and memory components that decay as described above. In this way, a dynamic pitch transition matrix is created. This is compared with Krumhansl's (1990, p. 125) data on judged two-tone sequences in tonal contexts (similar to Krumhansl, 1979, but with the addition of minor keys). Comparison with these empirical data gives a measure of key strength. This temporal-order model performed no better than the model described above that simply compared the dynamic tone distributions with the probe tone data of Krumhansl and Kessler (1982). However, the two models together performed slightly better than either model alone, which suggests that they were capturing somewhat different and complementary aspects of tonality induction.

Before leaving this topic, one other model should be mentioned. Chew (2000) proposed a geometrical model of tonality. It used a "spiral array" to represent the relations among pitches, intervals, chords, and keys. This allows comparisons among elements from different hierarchical levels. In one application, a key-finding algorithm is developed, called the

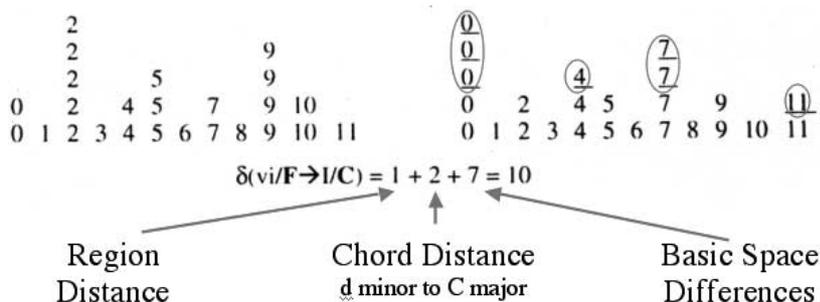


Fig. 5. How distance is computed between a d minor chord in the key of F major and a C major chord in the key of C major in Lerdahl's (2001) *Tonal Pitch Space* model.

“center of effect generator” (CEG). The algorithm takes pitch and duration information from the musical segment being analyzed. The center of effect is the (weighted) composite of the positions of the tones in the configuration. This point is then compared with points for the keys in the spiral array, with the closest key taken to be the result. When tested with the 24 fugue subjects of Book I of Bach's *Well-Tempered Clavier*, it outperformed the algorithms of both Longuet-Higgins and Steedman (1971) and Krumhansl and Schmuckler (Krumhansl, 1990).

The success of these models in various applications suggests the utility of this general approach. Other potential applications have not been explored extensively, but may prove interesting. These include models that examine how meter and rhythm contribute to tonality perception, the relation between harmony and melody in producing a sense of key, and models that explicitly attempt to characterize melodic and harmonic expectations.

6. Tonality and music theory

We turn now to some related developments in music theory, and some challenges and new directions that follow from those developments. As noted, one of the early efforts to combine music with linguistic formalisms was Lerdahl and Jackendoff's (1983) *Generative Theory of Tonal Music*. A number of empirical programs supported aspects of the theory (e.g., Deliège, 1987; Palmer & Krumhansl, 1987a,b; Bigand, 1990), although its extent prohibits exhaustive experimental tests. One component that was not fully explicit in the original model is the wealth of intuitions about tonal organization that was implicit in the model and that influenced the analyses. Lerdahl's (1988, 2001) *Tonal Pitch Space* model was developed to make explicit these aspects. As indicated below, it draws on research in music cognition. In addition, it proposes that the building blocks of the theory might be modified for other musical styles based on different scales and harmonic relations.

The basic pitch space in the theory consists of five levels: the highest level is the tonic (the first scale degree), the next adds the dominant (the fifth scale degree), the next adds the

mediant (the third scale degree), the next adds the other diatonic scale tones, and the final, and lowest level, includes all chromatic scale tones. This structure strongly resembles empirical measures of the Western tonal hierarchy (e.g., Krumhansl & Kessler, 1982). A second component of the model expresses the distances between chords in a key by the circle of fifths for chords. The third component is a map of keys (called regions) similar to the torus derived from the probe tone ratings by Krumhansl and Kessler (1982).

These three components are used to compute the distance from a chord in any region to any other chord in any other region. Consider the computation for a vi (d minor) chord in F major to an I (C major) chord in C major shown in Figure 5. The distance is the sum of the region distance (the i -distance in the map of regions, in this case, $i = 1$ for F major to C major), the chord distance (the j -distance on the circle of fifths for chords, in this case, $j = 2$, the distance between d minor and C major chords), and the number of differences in the basic space (the k -distance, the number of new entries in the basic space for the second chord compared to the first, in this case, $k = 7$ as outlined in the figure).

The predictions for a musical piece, however, depend on a hierarchical representation of how the events relate to one another in that piece. This is sometimes referred to as an event hierarchy, and is different from the hierarchies that apply to tones and chords in the style generally. The event hierarchy is shown as a tree, where the more stable events are those connected more directly to the root of the tree. Subordinate events branch off these. These embeddings are taken into account in computing distances between events in the music. An event's distance is computed from the event that is superordinate to it and, in addition, inherits the i -, j -, and k -distances from all events superordinate to it. The model also describes three different aspects of what Lerdahl calls surface dissonance that are also quantified. These sum together to give the global total value that can be compared to empirical results.

Lerdahl (1996) proposed that the global total, computed in this way, is a measure of tension. An early test of the idea used the initial eight measures of a Mozart piano sonata, K. 282 (Krumhansl, 1996). These two papers contain a fuller description of the model and its predictions for this particu-

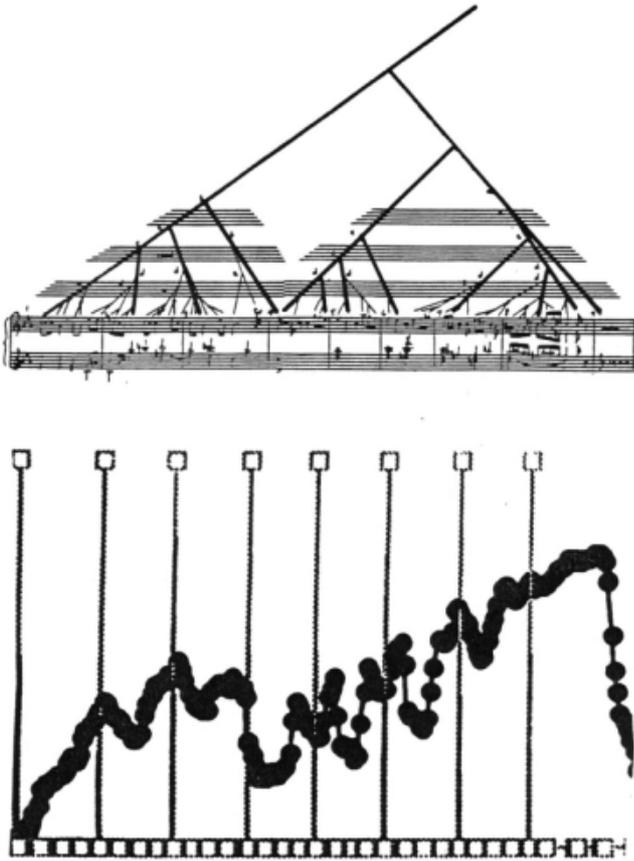


Fig. 6. The event hierarchy (top) and the tension judgments (bottom) for the first eight measures of Mozart's *Piano Sonata K. 282* (from Krumhansl, 1996). The most structurally significant events are shown as dark lines in the event hierarchy; these correspond to drops in tension judgments.

lar excerpt. In the experiment, listeners adjusted the position of a slider to indicate how the amount of tension changed over time. The listeners' judgments were fairly close to the predictions. A qualitative impression of the match can be seen in Figure 6. The tree structure is shown at the top. As can be seen, the branches connecting more directly to the top of the tree (those closest to the superordinate event in the excerpt, shown with dark lines) correspond to drops in tension in the listeners' judgments (shown at the bottom). Additional support for the model in short chord sequences can be found in the paper by Bigand et al. (1996).

In subsequent studies, we (Lerdahl et al., 2000; Lerdahl & Krumhansl, 2001a,b) have applied this approach to a number of different musical excerpts. There are four stages in the methodology. First, predictions are made for the excerpt from the theoretical model. Second, listeners' judgments of the rise and fall of tension are recorded and analyzed. Third, the predictions are compared with the data, noting discrepancies. Fourth, principled ways are sought within the theory to achieve a higher correlation between the model predictions and the data. In this iterative process, a fuller theoretical account can be made of the empirical results.

The excerpts that we have used in these studies are a diatonic Bach chorale, diatonic and hexatonic passages from Wagner's *Parsifal* (Lerdahl et al., 2000), Chopin's highly chromatic *E major Prelude* (Lerdahl & Krumhansl, 2001a), and an octatonic passage from Messiaen's *Quartet for the End of Time* (Lerdahl & Krumhansl, 2001b). In addition to the tension predictions described above, calculations were also made of attraction, another component of the *Tonal Pitch Space* model (Lerdahl, 2001). It is based on Bharucha's (1984, 1996) conception of the psychological factors underlying melodic attraction, which he called "anchoring". These factors are pitch proximity and the tendency of unstable pitches to resolve to more stable pitches. In Lerdahl's (2001) formulation, the degree of attraction is a function of both the position of the tones in the basic pitch space and their distance in semitones. A strong attraction is predicted when a tone that is low in the basic space is followed by a tone that is high in the basic space, when it is close in pitch.

Complete results and analyses will be presented in another context, but the case of the octatonic excerpt by Messiaen will be noted briefly here because of it illustrates how the theory can be applied outside diatonic music. The theory is modified for the octatonic case as follows. The scale level in the basic pitch space consists of the tones of the octatonic scale. There are only three different octatonic scales, so the region representation is simply a triangle. However, each octatonic scale has four major and four minor chords that can be formed from scale tones. They are arranged in a circle, with neighboring chords alternating between major and minor triads and roots separated by an interval of a minor third. (Thus, one circle consists of the chords: C-c-Eb-eb-F#-f#-A-a-C.) We tested this adaptation of the diatonic model using the beginning of the fifth movement from Messiaen's *Quartet for the End of Time*. The first 30 events are analyzed in E octatonic, followed by a shift to E major diatonic.

In the perceptual experiment, listeners made tension judgments in real-time as the music was sounded. The results conformed quite well to the model's predictions for tension and attraction. The listeners also performed the concurrent probe tone task as described in the last section. These judgments can be compared with the predictions of the basic pitch space, coded numerically as the number of levels represented for each tone. The results conformed well with the basic space predictors. In the section analyzed as octatonic, the octatonic basic space accounted better for the empirical data than did the diatonic basic space. Just how far this theoretical apparatus can be extended beyond traditional tonal-harmonic music is an open question, but these initial results are encouraging.

6. Tonality and brain science

Finally, let us turn to the question of what brain science might tell us about tonality, specifically, how it is processed by

neural mechanisms. There is now a growing literature investigating brain responses to music using a variety of techniques. For example, Besson and Faïta (1995) found a late positive component in event-related potentials (ERPs) when the final tone of a melody was changed to a tone out of the key of the melody. This was found whether or not a semantically incongruous word was sung to that tone, suggesting a dissociation between the processing of tone and word. Using both ERPs and EMFs (event-related magnetic fields), Koelsch and collaborators (Koelsch et al., 2000; Maess et al., 2001) found brain responses elicited by a tonally unexpected chord in a chord sequence. This was manifested as an early right-anterior negativity (ERAN).

Two fMRI studies have recently shown brain areas with different responses to expected and unexpected musical events. The first study by Tillmann et al. (2002) looked at brain activations in a harmonic priming paradigm. In this paradigm, a harmonic sequence of seven chords served as the priming context. The eighth and final chord was the target, which was either expected or unexpected given the tonality of the priming sequence. Half of the target chords of each type were made dissonant by adding a raised fifth. The task was to judge whether the final chord was consonant (a major chord) or dissonant (a major chord with a raised fifth added). Thus, the task was independent of the tonality manipulation (expected versus unexpected target chord). Scanning occurred continuously during the runs, which also included periods of rest.

The main interest in this paradigm is the effect that the related versus unrelated prime-target relationship has on the processing of consonant targets. For these targets, the blood oxygen level-dependent (BOLD) response was greater for unexpected targets than expected targets in a number of regions, notably in a region of interest around the frontal operculum (including inferior frontal gyrus, pars opercularis, and anterior insula). Activations in this region were generally greater in the right hemisphere than the left hemisphere, and for dissonant targets than consonant targets.

The second fMRI was reported by Koelsch et al. (2002). Sequences of five chords were presented in rapid succession and scanning occurred continuously during stimulus presentation. The key of the sequences changed approximately every three or four sequences. In-key sequences contained five chords in a single key. The study included two different stimulus manipulations that might be described as related to tonality. In the first, called modulations, the last three chords were in a key different from the key of the in-key sequences; this is most clearly a manipulation of tonality. In the second, called clusters, the final tonic chord of the five-chord sequence was replaced by a dissonant tone cluster that included out-of-scale tones. Although this manipulation introduced out-of-key tones, it might be considered more comparable to the dissonance manipulation of Tillmann et al. (2002). A third manipulation changed the timbre of one or two in-key chords from piano to an instrument other than piano. The subjects' task was to respond whenever they heard

the deviant instrument. However, they also had to detect the clusters, because the appearance of a cluster signaled that they should change the button used to indicate that a timbre deviant had occurred. Thus, as in the Tillmann et al. (2002) study, the task relevant variables were independent of the stimulus manipulation of tonality.

In the analysis of the fMRI data, the in-key sequences were used as controls for the other three conditions: modulations, clusters, and deviant instrument. Significant activations were found in superior temporal gyrus and inferior frontal gyrus, including Brodmann's areas 44/6 and insula, as well as thalamus and other areas. Activations in these conditions (compared to in-key sequences) were stronger for deviant instruments (directly response relevant) and clusters (indirectly response relevant) than modulations. For the modulations, the activation in inferior frontal gyrus was stronger in the right hemisphere than in the left hemisphere. At a general level, the pattern of activations was quite similar to that of Tillmann et al. (2002).

Krumhansl and Zatorre (2003a,b) compared the processing of tonality and rhythm. We were interested in this because tonal and rhythmic patterns have independent behavioral effects on judgments of phrasing and melodic similarity, and are treated as independent aspects of structure in music theory. Studies of patients also show dissociations between rhythmic and pitch organization in a variety of musical skills, including singing, reading music, and remembering melodies (Peretz, 1990; Peretz, Kolinsky et al., 1994). This suggests that separate brain mechanisms may be involved in perceiving these two aspects of musical structure.

The neural mechanisms were studied in two ways. First, either tonal or rhythmic violations (or both) were introduced into short melodic sequences. Second, the task was manipulated, such that listeners judged either the tonal structure or the rhythmic structure (or passively listened without performing any task). Eight 6-sec long melodies were written by Cornell composition student, Diego Vega, who was asked to make them as tonally and metrically clear as possible. When violations occurred, there were always four of them, two just after 2 sec in the melody and two just after 4 sec. The melodies had to be short in order to present them without the simultaneous noise of the scanner, as will be described later.

The design of the experiment was as follows. There were four kinds of stimuli: no violations, tonal violations, rhythmic violations, and both tonal and rhythmic violations. These were randomly intermixed. There were two runs, each lasting approximately 20 minutes. For one-third of each run listeners judged the tonality, for one-third they judged the rhythm, and for one-third they passively listened to the melodies. The order of the tasks was counterbalanced across listeners. The scanning produces a relatively loud magnet noise. The timing of the scanning was based on prior results about when maximal brain responses occur in fMRI measures, typically in the range of 2–8 sec after an event is heard. Because the sequences were short, the noise of scanning did not co-occur

with the musical sequence, unlike the other two fMRI studies just described. The measurement was made when a maximal brain response to the tonal and rhythmic violations would be expected. Subjects were musically trained, and they performed with high accuracy on both tasks.

The fMRI data analysis was conducted according to the following hierarchy of stimulus types and tasks: silence, monotone controls, passive listening, active judgments (that is, the tonality and rhythm judgments). The monotone sequences had the same timbre and rhythms as the musical sequences and they were interspersed with them, as were periods of silence. Compared to silence, the monotone sequences produced strong activations in superior temporal cortex (bilaterally) in a broad area covering primary and secondary auditory cortex. When passive listening was compared with monotone controls, additional activations were found in superior temporal cortex (somewhat lateralized to the right hemisphere). Thus, pitch varying sequences appear to recruit processing resources in auditory cortex not engaged by monotonic sequences.

When the active judgments were compared with passive listening, no additional activations were found in superior temporal cortex. This suggests that so-called passive listening involves complete perceptual analysis of incoming stimuli, and is not affected by specific demands to attend to one or another feature of the melody. However, two additional areas of activation were found for the active judgments in frontal cortex. Right dorsolateral frontal areas were found; these are known to be involved in working memory mechanisms. The clear unilateral activation within the right hemisphere further supports the existence of hemispheric functional asymmetries for music. In addition, inferior frontal activations were found bilaterally. This area has been implicated in a variety of tasks, including speech perception, semantic decision, and tonal retrieval. Thus, this activation may reflect a quite general process of accessing knowledge about stimulus structure in a variety of domains.

Turning now to possible differences in brain activations for processing tonality and rhythm, this was tested in two ways. In one test the tonality judgments were compared with passive listening, and the rhythm judgments were compared with passive listening. The activations were similar. In a second test, a direct comparison was made between the two active tasks; this test showed no significant differences. Thus, the manipulation of task did not reveal differences in brain activations, as might have been expected from some prior behavioral and patient results.

What was the effect of tonal and rhythmic violations? One analysis contrasted musical sequences containing both tonal and rhythmic violations with those without any violations. Even when the subjects were passively listening (not performing a task), superior temporal activations were found (especially on the right) as shown in Figure 7. Activations were stronger when violations were present than when they were not. The pattern was similar, but weaker with sequences that included only tonal or only rhythmic violations. This is

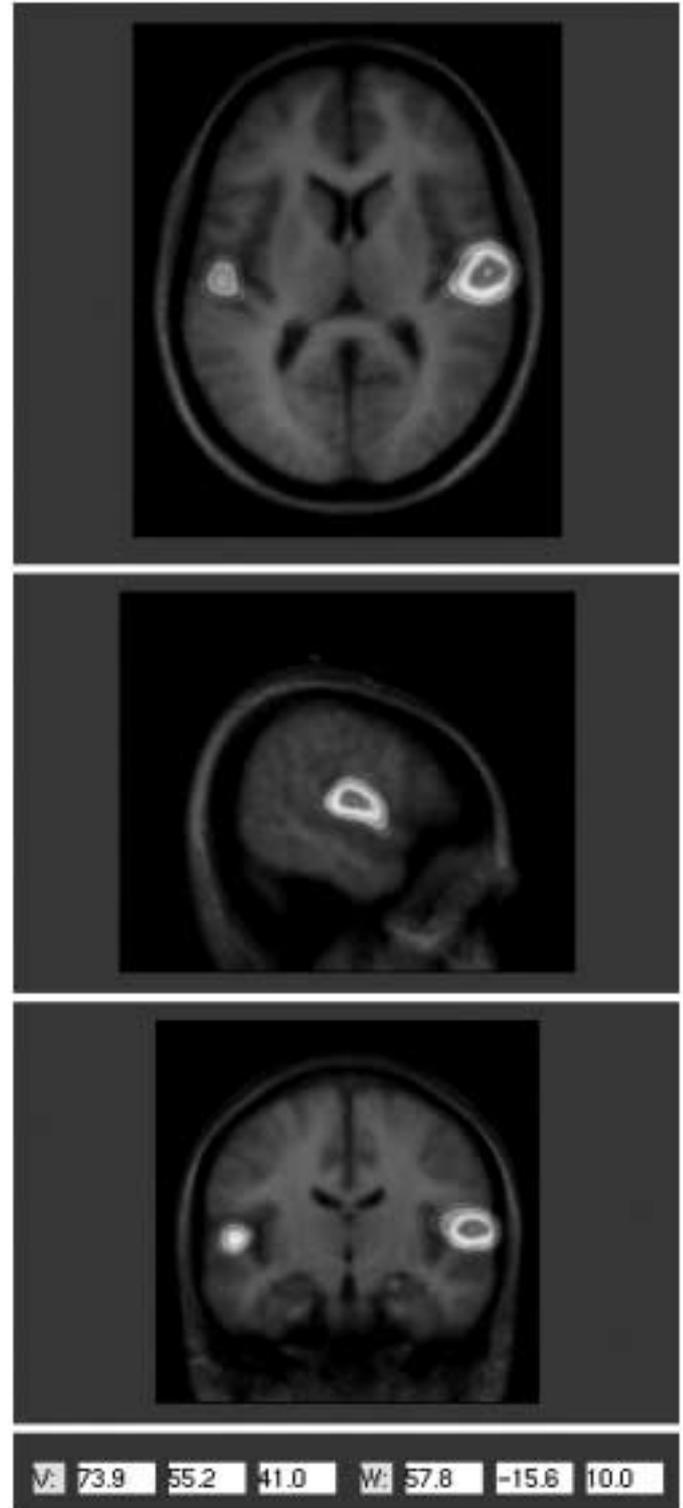


Fig. 7. The temporal lobe activation in response to violations of tonality and rhythm in an fMRI study (from Krumhansl & Zatorre, 2003a,b). The data are from the passive listening condition; additional frontal lobe activations were found when listeners made overt judgments of tonality and rhythm.

consistent with the idea mentioned above that right auditory cortex processes tonal pattern information automatically. More important, this finding indicates that sensory analysis is sensitive to the relatively abstract rules of tonality and rhythm.

When the same analysis was done for the active judgments, the same activation was found in superior temporal gyrus, again stronger on the right. Additional frontal activations were found for the tonal and rhythmic violations when subjects were performing active judgments. Thus, the frontal activations are interpreted as reflecting the memory and attentional task requirements, whereas the superior temporal activations appear to reflect the processing of the tonal and rhythmic violations themselves.

This interpretation of the frontal activations is somewhat different from that of the two fMRI studies described above. One reason for this may be the relationship between the task and the stimulus manipulation of tonality. In the present study, the task given listeners was directly relevant to the stimulus manipulation: they were to judge whether rhythmic or tonal violations occurred. In the other two studies, the overt task was independent of the tonality manipulation in the experimental design. In the Tillmann et al. (2002) experiment, it was to judge the consonance or dissonance of the last chord. The advantage for the related targets (the priming effect) shows, however, that this judgment is not independent of the tonality manipulation (whether the target chord is or is not related to the priming context). In the Koelsch et al. (2002) experiment, the primary task was to judge whether notes in a deviant timbre had occurred, with a secondary task of detecting dissonant tone clusters. Again, it is possible that the tasks interacted with the tonality manipulation, which was whether or not a modulation had occurred. This raises the possibility that in both studies the frontal activations reflect the difficulty of filtering out a tonality manipulation that is irrelevant to the task, rather than an effect of processing tonality itself.

A final study that should be mentioned before summarizing is that of Janata et al. (2002). They analyzed the brain activity of musicians as they performed two music perception tasks requiring them to detect a deviant timbre and notes that violated the local tonality. In this study, scanning occurred during the presentation of the music. They found brain activations in auditory cortex as well as a number of other brain areas, including temporal, parietal, frontal, and limbic lobes as well as thalamus and cerebellum. Thus, processing of music is complex, involving a great number of brain areas. These authors identified the frontal lobe activity in the orbital-frontal areas as related to the modulations that occurred within the sequence. This was based on a model that tracked the modulations in the toroidal map of musical keys, and which correlated with changes in brain activity. A number of questions might be raised about this procedure, including the very large number of parameters used to fit tonality. At a more general level, an important question is whether this activation reflects tonality processing in partic-

ular, or the more general functions of remembering and comparing tones as required for the tasks.

7. Summary

This article surveyed a range of approaches to understanding tonality. The concept, as treated in music theory, appears versatile and amenable to examination with a wide range of methods drawn from different disciplines. However, each approach suggests further questions. Basic experimental results establish that the elements of tonality have psychological reality in that they have demonstrable effects on how musical pitch is encoded and remembered. The way in which these effects play out with more complex and realistic musical materials needs more analysis. Limits in acoustic explanations for Western scales and harmony raise questions about the role of learning and cognition. Results of psychological studies indicate that Western listeners, even those without formal instruction, have extensive knowledge of typical tonal and harmonic patterns. However, contrary to traditional assumptions, at least some aspects of this knowledge are acquired without extensive experience and training. Cross-cultural studies suggest listeners possess a relatively flexible cognitive system. Short-term statistical learning is one mechanism that may enable this flexibility. In addition, listeners appear to build on a set of basic perceptual principles that may adapt to different styles. Computational modeling has captured some of the dynamic aspects of tonality perception, but the models generally fall short of human sensitivity and ignore other dimensions of music, particularly rhythm and meter. Another issue is whether tonality, as formally elaborated within music theory, can account for stylistic variations across historical periods, particularly compositional innovations in the last century. Preliminary results from a study with octatonic music are encouraging, but the limits of theoretical extensions beyond tonal-harmonic music are far from understood. A final approach investigated the neural systems involved in processing tonality with techniques of brain imaging. Localized cortical areas are identified, but these appear to be activated during a variety of cognitive tasks and suggest the functioning of rather general processing mechanisms. Thus, despite an intense multidisciplinary effort, tonality remains remarkably elusive.

References

- Ayiri, M., & McAdams, S. (2003). Aural analysis of Arabic improvised instrumental music (taqsim). *Music Perception*, 21, 159–216.
- Besson, M., & Fäita, F. (1995). An event-related potential (ERP) study of musical expectancy: Comparison of musicians with nonmusicians. *Journal of Experimental Psychology: Human Perception and Performance*, 21, 1278–1296.
- Bharucha, J.J. (1984). Anchoring effects in music: The resolution of dissonance. *Cognitive Psychology*, 16, 485–518.

- Bharucha, J.J. (1996). Melodic anchoring. *Music Perception*, 13, 383–400.
- Bigand, E. (1990). Abstraction of two forms of underlying structures in a tonal melody. *Psychology of Music*, 18, 45–59.
- Bigand, E., Parncutt, R., & Lerdahl, F. (1996). Perception of musical tension in short chord sequences: The influence of harmonic function, sensory dissonance, horizontal motion, and musical training. *Perception & Psychophysics*, 46, 9–17.
- Bregman, A.S. (1990). *Auditory scene analysis: The perceptual organization of sound*. Cambridge, MA: MIT Press.
- Castellano, M.A., Bharucha, J.J., & Krumhansl, C.L. (1984). Tonal hierarchies in the music of North India. *Journal of Experimental Psychology: General*, 113, 394–412.
- Chew, E. (2000). Towards a mathematical model of tonality. Unpublished doctoral dissertation, Cambridge, MA: MIT.
- Cohen, A.J. (1991). Tonality and perception: Musical scales primed by excerpts from the Well-Tempered Clavier of J.S. Bach. *Psychological Research*, 53, 305–314.
- Cuddy, L.L. (1997). Tonal relations. In: I. Deliège, & J. Sloboda (Eds.), *Perception and Cognition of Music* (pp. 329–352). Hove, Sussex: Taylor & Francis.
- Cuddy, L.L., Cohen, A.J., & Miller, J. (1979). Melody recognition: The experimental application of rules. *Canadian Journal of Psychology*, 33, 148–157.
- Deliège, I. (1987). Grouping conditions in listening to music: An approach to Lerdahl & Jackendoff's grouping preference rules. *Music Perception*, 4, 325–360.
- Dewar, K.M., Cuddy, L.L., & Mewhort, D.J.K. (1977). Recognition memory for single tones with and without context. *Journal of Experimental Psychology: Human Learning and Memory*, 3, 60–67.
- Dowling, W.J. (1978). Scale and contour: Two components of a theory of memory for melodies. *Psychological Review*, 85, 341–354.
- Eerola, T. (2003). *The dynamics of musical expectancy: Cross-cultural and statistical approaches to melodic expectations*. Jyväskylä, Finland: Jyväskylä University Printing House.
- Helmholtz, H.L.F. (1954). *On the sensations of tone as a physiological basis for the theory of music* (A.J. Ellis, Ed., & trans.). New York: Dover. (Revised edition originally published, 1885).
- Huron, D., & Parncutt, R. (1993). An improved model of tonality perception incorporating pitch salience and echoic memory. *Psychomusicology*, 12, 154–171.
- Hyer, B. (2003). Tonality. *Grove Music Online* (L. Macy, Ed.) accessed 23 October 2003, <http://ezproxy.library.cornell.edu:2691>.
- Janata, P., Birk, J.L., Van Horn, J.D., Leman, M., Tillmann, B., & Bharucha, J.J. (2002). The cortical topography of tonal structures underlying Western music. *Science*, 298, 2167–2170.
- Jusczyk, P.W., & Krumhansl, C.L. (1993). Pitch and rhythmic patterns affecting infants' sensitivity to musical phrase structure. *Journal of Experimental Psychology: Human Perception and Performance*, 19, 627–640.
- Knopoff, L., & Hutchinson, W. (1983). Entropy as a measure of style: The influence of sample length. *Journal of Music Theory*, 27, 75–97.
- Koelsch, S., Gunter, T., Friederici, A.D., & Schröger, E. (2000). Brain indices of music processing: "Nonmusicians" are musical. *Journal of Cognitive Neuroscience*, 12, 520–541.
- Koelsch, S., Gunter, T.C., v. Cramon, D.Y., Zysset, S., Lohmann, G., & Friederici, A.D. (2002). Bach speaks: A cortical "language-network" serves the processing of music. *NeuroImage*, 17, 956–966.
- Krumhansl, C.L. (1979). The psychological representation of musical pitch in a tonal context. *Cognitive Psychology*, 11, 445–463.
- Krumhansl, C.L. (1985). Perceiving tonal structure in music. *American Scientist*, 73, 371–378.
- Krumhansl, C.L. (1990). *Cognitive Foundations of Musical Pitch*. New York: Oxford University Press.
- Krumhansl, C.L. (1995). Music psychology and music theory: Problems and prospects. *Music Theory Spectrum*, 17, 53–80.
- Krumhansl, C.L. (1996). A perceptual analysis of Mozart's piano sonata K. 282: Segmentation, tension, and musical ideas. *Music Perception*, 13, 401–432.
- Krumhansl, C.L. (2000). Tonality induction: A statistical approach applied cross-culturally. *Music Perception*, 17, 499–517.
- Krumhansl, C.L., & Jusczyk, P.W. (1990). Infants' perception of phrase structure in music. *Psychological Science*, 1, 70–73.
- Krumhansl, C.L., & Kessler, E.J. (1982). Tracing the dynamic changes in perceived tonal organization in a spatial map of musical keys. *Psychological Review*, 89, 334–368.
- Krumhansl, C.L., Louhivuori, J., Toivianinen, P., Järvinen, T., & Eerola, T. (1999). Melodic expectancy in Finnish folk hymns: Convergence of behavioral, statistical, and computational approaches. *Music Perception*, 17, 151–196.
- Krumhansl, C.L., & Shepard, R.N. (1979). Quantification of the hierarchy of tonal functions within the diatonic context. *Journal of Experimental Psychology: Human Perception and Performance*, 5, 579–594.
- Krumhansl, C.L., Toivanen, P., Eerola, T., Toivianinen, P., Järvinen, T., & Louhivuori, J. (2000). Cross-cultural music cognition: Cognitive methodology applied to North Sami yoiks. *Cognition*, 75, 1–46.
- Krumhansl, C.L., & Toivianinen, P. (2001). Tonal cognition. In: R.J. Zatorre, & I. Peretz (Eds.), *The Biological Foundations of Music*. Annals of the New York Academy of Sciences, Volume 930, pp. 77–91.
- Krumhansl, C.L., & Zatorre, R.J. (2003a). Rhythmic and tonal violations in music: An fMRI study. Meeting of the Organization for Human Brain Mapping, New York City.
- Krumhansl, C.L., & Zatorre, R.J. (2003b). Tonal processing in music: The role of right frontal cortex. Meeting of the Cognitive Neuroscience Society, New York City.
- Leman, M. (1995). A model of retroactive tone-center perception. *Music Perception*, 12, 439–471.
- Leman, M. (2000). An auditory model of the role of short-term memory in probe-tone ratings. *Music Perception*, 17, 481–509.

- Leman, M., & Carreras, F. (1997). Schema and Gestalt: Test the hypothesis of psychoneural isomorphism by computer simulation. In: M. Leman (Ed.), *Music, Gestalt, and Computing – Studies in Cognitive and Systematic Musicology* (pp. 144–168). Heidelberg: Springer.
- Lerdahl, F. (1988). Tonal pitch space. *Music Perception*, 5, 315–349.
- Lerdahl, F. (1996). Calculating tonal tension. *Music Perception*, 13, 319–363.
- Lerdahl, F. (2001). *Tonal pitch space*. New York: Oxford University Press.
- Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music*. Cambridge, MA: MIT Press.
- Lerdahl, F., & Krumhansl, C.L. (2001a). *Modeling tonal tension and attraction in chromatic contexts: Chopin*. Kingston, Ontario: Society for Music Perception and Cognition.
- Lerdahl, F., & Krumhansl, C.L. (2001b). Modeling tonal tension and attraction in chromatic contexts: Wagner and Messiaen. *Symposium on Neo-Riemannian Theory*, University of Buffalo.
- Lerdahl, F., Krumhansl, C.L., Hannon, E., & Fineberg, J. (2000). Modeling tonal tension and attraction. Presentation given at Toronto: Musical Intersections.
- Lindblom B., & Sundberg, J. (1970). Towards a generative theory of melody. *Swedish Journal of Music*, 52, 71–88.
- Longuet-Higgins, H.C., & Steedman, M.J. (1971). On interpreting Bach. *Machine Intelligence*, 6, 221–241.
- Maess, B., Koelsch, S., Gunter, T.C., & Friederici, A.D. (2001). Musical syntax is processed in Broca's area: an MEG study. *Nature Neuroscience*, 4, 540–545.
- Narmour, E. (1990). *The Analysis and Cognition of Basic Melodic Structures: The Implication-Realization Model*. Chicago, IL: University of Chicago Press.
- Oram, N., & Cuddy, L.L. (1995). Responsiveness of Western adults to pitch-distributional information in melodic sequences. *Psychological Research*, 57, 103–118.
- Palmer, C., & Krumhansl, C.L. (1987a). Independent temporal and pitch structures in determination of musical phrases. *Journal of Experimental Psychology: Human Perception and Performance*, 13, 116–126.
- Palmer, C., & Krumhansl, C.L. (1987b). Pitch and temporal contributions to musical phrase perception: Effects of harmony, performance timing, and familiarity. *Perception & Psychophysics*, 41, 505–518.
- Papousek, M., & Papousek, H. (1981). Musical elements in the infant's vocalization: Their significance for communication, cognition, and creativity. *Advances in Infancy Research*, 1, 163–224.
- Parncutt, R. (1988). Revision of Terhardt's psychoacoustical model of the root(s) of a musical chord. *Music Perception*, 6, 65–94.
- Parncutt, R. (1989). *Harmony: A Psychoacoustical Approach*. Berlin: Springer.
- Parncutt, R. (1994). Template-matching models of musical pitch and rhythm perception. *Journal of New Music Research*, 23, 145–167.
- Peretz, I. (1990). Processing of local and global musical information by unilateral brain-damaged patients. *Brain*, 113, 1185–1205.
- Peretz, I., Kolinsky, R., Tramo, M., Labrecque, R., Hublet, C., Demeurisse, G., & Belleville, S. (1994). Functional dissociations following bilateral lesions of auditory cortex. *Brain*, 117, 1238–1301.
- Saffran, J.R., Johnson, E.K., Aslin, R.N., & Newport, E.L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, 70, 27–52.
- Shmulevich, I., & Yli-Harja, O. (2000). Localized key finding: Algorithms and applications. *Music Perception*, 17, 531–544.
- Sundberg, J., & Lindblom, B. (1976). Generative theories in language and music descriptions. *Cognition*, 4, 99–122.
- Temperley, D. (1999). What's key for key? The Krumhansl-Schmuckler key-finding algorithm reconsidered. *Music Perception*, 17, 65–100.
- Terhardt, E., Stoll, G., & Seewan, M. (1982a). Algorithm for extraction of pitch and pitch salience from complex tonal signals. *Journal of the Acoustical Society of America*, 71, 671–678.
- Terhardt, E., Stoll, G., & Seewan, M. (1982b). Algorithm for extraction of pitch and pitch salience from complex tonal signals. *Journal of the Acoustical Society of America*, 7, 679–688.
- Thompson, W.F., & Cuddy, L.L. (1989). Sensitivity to key change in chorale sequences: A comparison of single voices and four-voice harmony. *Music Perception*, 7, 151–168.
- Thompson, W.F., & Cuddy, L.L. (1997). Music performance and the perception of key. *Journal of Experimental Psychology: Human Perception and Performance*, 23, 116–135.
- Tillmann, B., Bharucha, J.J., & Bigand, E. (2000). Implicit learning of tonality: A self-organizing approach. *Psychological Review*, 107, 885–913.
- Tillmann, B., Janata, P., & Bharucha, J.J. (2002). Activation of the inferior frontal cortex in music priming. *Cognitive Brain Research*, 16, 145–161.
- Toiviainen, P., & Krumhansl, C.L. (2003). Measuring and modeling real-time responses to music: The dynamics of tonality induction. *Perception*, 32, 741–766.
- Trainor, L.J. (1997). Effect of frequency ratio on infants' and adults' discrimination of simultaneous intervals. *Journal of Experimental Psychology: Human Perception and Performance*, 23, 1427–1438.
- Trehub, S.E. (2000). Human processing predispositions and musical universals. In: N.L. Wallin, B. Merker, & S. Brown (Eds.), *The Origins of Music*. Cambridge, MA: MIT Press.
- Vos, P., & van Vianen, B.G. (1984). Thresholds for discrimination between pure and tempered intervals: The relevance of nearly coinciding harmonics. *Journal of the Acoustical Society of America*, 77, 176–187.
- Youngblood, J.E. (1958). Style as information. *Journal of Music Theory*, 2, 24–35.
- Zentner, M.R., & Kagan, J. (1996). Perception of music by infants. *Nature*, 383, 29.