

**Listening to Music in the Real World? A Critical
Discussion of Marc Leman's (1995) *Music and Schema
Theory: Cognitive Foundations of Systematic Musicology****

Richard Parncutt

ABSTRACT

Leman's recent book is examined from both musical and scientific perspectives. The book describes an important new development in the context-sensitive modeling of musical pitch perception, and a promising basis for the modeling of other musical parameters such as timbre, and even emotion. The main advantage of the approach is its ecological validity: as far as current limitations of knowledge and technology allow, the model realistically simulates, or attempts to simulate, various peripheral and central parts of human auditory physiology; and the input to the model is not the artificial constructs of music theory or even notated music, but real, sounding music. Leman describes how schema of musical pitch perception can develop by listening to music, and how they may consequently be used to monitor the tonality of real pieces. A major unresolved problem is that, according to the various simulations presented in the book, the more complex models do not necessarily perform better than the simpler models when predictions are compared against experimental data. Thus, the book does not always present a strong case for the realism and usefulness of the more complex models.

INTRODUCTION

The past few decades have seen major advances in our understanding of the physiology of auditory perception at all levels of the auditory system. At the same time, there have been associated advances in auditory psychophysics.

*I am grateful to Marc Leman, Alan Marsden, and Georg Meyer for helpful comments.

Address for correspondence: Richard Parncutt, Dept. of Musicology, University of Graz, Mozartgasse 3, 8010 Graz, Austria. Tél.: +43-316-380-2405. Fax: +43-316-380-9755. E-mail: richard.parncutt@Kfunigraz.ac.at

Manuscript submitted: March 31, 1998

Most interestingly for the study of perception, there have been improvements in our understanding of the complex relationships that obtain between auditory physiology on the one hand, and auditory psychophysics on the other.

This book represents an important stage of that development. It describes and applies a number of more or less physiologically based models of auditory perception. In the most complex model (by which I mean the model requiring the greatest amount of information processing in the form of computing time and memory for its implementation, and including a considerable number of stages with associated opportunities for "tweaking" parameters), pitch is extracted at the periphery from the signal by correlational analysis of periodicities in firing spikes along auditory nerve fibers (Van Immerseel, 1993). After that, artificial neural networks are used to model (i) the emergence of perceptual schemata and tonal centers from passive exposure to pitch information in musical contexts, and (ii) the processes by which perceptual schemata monitor pitch perception during music perception. The study incorporates recently developed physiologically based models of the human auditory system; by applying them to specifically musical problems, it allows the musical usefulness and appropriateness of these models to be assessed in a more systematic fashion than has previously been possible.

On the basis of his simulation of human audition, and as an example of what this kind of modeling can do, Leman attempts to explicate the perception of tonality in Western music. The models presented in the book could, in principle, be applied to several other aspects of music or auditory perception, which will have to await detailed treatment in the future. A particularly promising avenue, only briefly introduced in the book, is the perception of timbre (cf. Toiviainen, 1996). The approach may conceivably also be used for integration of pitch, rhythm, and timbre perception into one model that attempts to account for the interactions between such parameters that underlie the emergence of musical meaning — an idea that is also introduced, but only tantalizingly briefly, on page 169. Regarding experimental methodologies, the book's approach could allow the results of innumerable semantic differential, similarity, and multidimensional scaling studies in music perception to be modeled.

WHAT LEMAN MEANS BY 'MEANING'

The book is to a large extent about musical meaning, but not in the sense in which I, for one, am used to using the word. There is no reference to the emotion-based approach of Meyer's (1956) classic work; his theory of emotional responses based on fulfillment or violation of expectations is not addressed. In principle, Leman's approach would provide an excellent starting

point for the prediction of emotional states, but the book (perhaps wisely!) does not go this far.

Bharucha (1994) presented a neural-net approach to musical emotion, explaining how expectations may in principle be acquired by supervised learning in a back-propagation network. In Leman's approach, a different principle (self-organizing network) would be used to model this process. Although still in its infancy, this idea looks like a very promising avenue for future research on the nature of music.

Leman denies that 'meaning' has an existence of its own beyond the physical world or the world of information processing, or as part of our private experience when listening to music. He takes a more 'naturalistic' approach, defining and explaining musical meaning as a phenomenon that is ultimately physical or physiological in nature. For him, musical meaning is based on acoustics, but constrained by human information processing (perceptual learning and cognition) and musical pitch-time configurations which are culturally determined. Meaning emerges when relationships are perceived between musical events that are embedded in musical structures. His specific term for this kind of meaning is *tone semantics*, defined as "an example of meaning formation at a non-symbolic level" that depends on "context-sensitive similarity relationships" (p. 180).

But no matter how successful neurophysiologically realistic models become for explaining musical meaning, I take issue with an approach that *equates* meaning with neurophysiology. With Popper and Eccles (1977), I would suggest instead that it is both scientifically practical and intuitively obvious to regard meaning as something non-physical that can never be 'observed' by studying the brain. The aim of (quasi-) neuroscientific research into meaning, such as Leman's, must then be to clarify relationships between meaning as reported by human subjects and the brain states that are observed to accompany them, keeping in mind that a correlation does not necessarily imply a causal relationship. Neurophysiology can tell us about underlying signal representations and their limitations, but not about meaning itself. Leman and Carerras (1997) have adopted this more moderate approach, stating that "The hypothesis of a psychoneural isomorphism assumes a correspondence between mental representations of perceived phenomena and brain-states or cortical representations" (abstract, p. 144).

Another central concept in the book is the *schema*, which Leman understands (diverging slightly from Bregman, 1990) as 'a control structure in the human brain that is sensitive to some frequently occurring pattern, either in the environment, in ourselves, or in how the two interact' (p. 40). The schema is thus an integral part of Leman's concept of musical meaning. A clear, concise account of how schemata may develop in self-organizing neural networks

is given in Chapter 6. In the case of tonal music, the frequently occurring patterns may include (at least conceptually) the pitch patterns of the audible harmonics of individual musical tones, the pitch patterns of the tones of chords and melodies, the relative frequency of occurrence of different tones in the chromatic scale, and so on. From an empirical point of view, this is probably the most solid aspect of Leman's model.

REVIEW OF MUSIC-PSYCHOLOGICAL LITERATURE

The book contains an extended literature review that covers the relevant territory thoroughly and concisely. There is, however, an unfortunate tendency to accept existing theories without criticism or qualification, leaving open the possibility for misinterpretation. For example, Stumpf's fusion 'data' are presented in idealized schematic form (p. 9, Fig. 2.3), but the caption 'Tone profile based on psychological data (Stumpf)', gives the false impression that they are real experimental data. The reader needs to be reminded that Sundberg's *melodic charge* values (p. 17) were not determined in the same empirical fashion as the key profiles of Krumhansl and Kessler (1982), but instead by a process of trial and error in which Sundberg's model was adjusted until its output sounded optimal (although Leman does indicate at the foot of page 16 that these data are obtained from different domains — perception and performance). On page 39 we are presented with a highly speculative theoretical object, the *rhythmogram* (Todd, 1994) — but with no indication in the text that, however promising its potential might be for understanding the perception of rhythm, this representation has not yet been the subject of careful independent analysis and testing, and should therefore be approached with some caution.

Leman rightly and refreshingly points out on page 18 that Shepard's double helix model of musical pitch does not account for dynamic, context-dependent aspects of pitch relationships; Shepard's well-known and widely accepted theory is simply not compatible with, and hence is superseded by, Leman's more realistic approach. On page 23, the well-known two-component model of pitch perception that underlies the pitch helix is appropriately criticized by pointing out that the dimensions *chroma* and *height* are emerging properties of perception, and not measures in the everyday sense like length and weight.

PITCH PERCEPTION AND FALSIFIABILITY

Perhaps the book's most important advantage is its physiological realism, which gives it an undeniable edge over purely cognitive or rule-based models. Leman's claims for realism are overstated, however, when applied to Van Immerseel's (1993) model of pitch. There is actually no *direct* evidence that autocorrelation of the kind used in that (and several other) pitch models actually underlies pitch perception. In this sense, the model is not necessarily more 'realistic' than a model that neglects temporal representations altogether.

Perhaps the most visible flaw of the models presented in the book is their unfalsifiability — yet this flaw is really common to all pitch models, and ultimately, perhaps, impossible to avoid. A falsifiable model would be one that could be shown to be somehow essentially 'wrong' by comparing its predictions with experimental data. But several different pitch models exist, all of which do a reasonably good job of accounting for a sizable core of psychoacoustical data — especially considering the noise inherent in most such data due to individual differences and uncertainties. In this sense, none of the models is falsifiable.

Another approach would be to compare the predictions of one model with those of others. Leman's analysis includes an interesting comparison of three conceptually very different pitch-perceptual models. In Chapters 7 and 8 he attempts to predict Krumhansl's tone profiles in three different ways, using models which differ in the kind of pitch-perceptual model they incorporate: the simple, octave-generalized chord-root model of Parncutt (1988); the more complex, but still phase-independent, spectral-psychoacoustic pitch-salience model of Terhardt, Stoll, and Seewann (1982); and the AMPEX multi-channel autocorrelation model of Van Immerseel (1993) (derived from that of Meddis & Hewitt, 1991a, but optimized for fast computation).

Which of the models comes closest to simulating available experimental data? The good news is that all three models closely approximate the profiles. The bad news, if one can call it that, is that the comparison is unable to 'disprove' any of the models. As found by Houtsma (1979), the more important pitch models typically perform better than their competitors in some cases, and worse in other cases.

On page 123, Leman comments on the very high correlation coefficients obtained using the model of Van Immerseel when simulating Krumhansl's key profiles, but neglects to note that similarly high correlations were obtained by Parncutt (1989), using a more straightforward model. Later in the book (under the heading *tone center attraction*), Leman reports that his model does not make predictions that are interestingly different from the predictions of

existing cognitive or rule-based models (e.g., Huron & Parncutt, 1993). Such comparisons do not help us to decide which of the models is the best, or whether the increasing sophistication of Leman's models is of any tangible use for the understanding of tonality. On the contrary, they suggest that the additional sophistication of the more complex models may be redundant for the purposes of the tasks to which they are assigned in this book. If Leman's models are indeed superior, their superiority lies not in their predictive power but in their physiological realism, and hence their ability to explain the processes of musical audition. Unfortunately, their failure to quantitatively outperform equivalent but simpler models casts serious doubts on their alleged realism.

The unfalsifiability of pitch-perceptual models may feel a little disconcerting for the uninitiated, given that important conceptual and structural differences exist between them (for example, pitch extraction may be performed either by harmonicity or by periodicity analysis). Because of the complexity of these theories, it is often impossible to test the various parts of them while controlling for confounds produced by other parts. But convergent evidence does exist from various sources that at least partially separate different stages of audition from each other (psychoacoustical experiments with humans, and physiological experiments with animals). On the basis of these various sources, Leman's model is about as close to reality as our present knowledge base permits.

'TEMPORAL' MODELS AND PHASE SENSITIVITY

On page 48, Terhardt's pitch model is misleadingly described as a 'place model'. It is true that his model belongs to a long tradition (Ohm, Helmholtz) in which spectral pitch, or the pitch of a pure tone, was originally thought to be determined only by the position of maximum excitation on the basilar membrane. In more recent times, the parallel role of temporal coding in determining the pitch of pure tones has been well documented. No-one, least of all Terhardt, would dispute that the auditory system uses both temporal and spatial coding to determine spectral pitches, the proportion depending on frequency (cf. Moore, 1989, p. 164; Shamma, 1985). There is also a large body of research on ways in which temporal signal structures might contribute to the perception of pitch in complex tones (virtual pitch); for a recent overview, see Langner (1997).

What makes Terhardt's model different from 'temporal' models is that virtual pitches are assumed to be extracted from spectral pitches by a central pattern recogniser (cf. Houtsma & Goldstein, 1972; Zatorre, 1988). Terhardt's model

may thus have been described more appropriately as a *pattern-recognition model* — a label that emphasizes its suitability for neural-net simulation, and is broad enough to accommodate the possibility of both temporally and spectrally represented information contributing to the determination of pitch at various levels of the auditory system. The algorithm of Terhardt et al. (1982) is implemented in the frequency domain (or more precisely — in his terminology — in the spectral pitch domain), but appears to account (unintentionally?) for temporal effects as well, due to the mathematical equivalence of (spectral) harmonicity and (temporal) periodicity.

Leman chose not to implement Terhardt's approach. The reason is not entirely clear from the text, but a possibility is that Terhardt's model ignores effects of *phase* relations between components. By contrast, the model of Meddis and Hewitt (1991a) accounts for effects of phase within critical bands or auditory filters (whereas between filters, phases are cancelled out by auto-correlation). Following Meddis and Hewitt's well-known (1991b) demonstration of the potential of their model to account for experimental evidence on phase sensitivity, it is widely assumed that pitch models should account for phase relationships among partials. I will argue here that this assumption invites serious challenge, especially in the context of an ecologically motivated approach such as Leman's.

In everyday listening environments, phase relationships are typically jumbled unrecognizably when sound is reflected off environmental objects; that is, when reflected sounds of varying amplitudes (depending on the specific configuration and physical properties of the reflecting materials) are added onto sound traveling in a direct line from the source. Thus, phase does not generally carry information that can reliably aid a listener in identifying sound sources in a reverberant environment (Terhardt, 1988; see also Terhardt, 1991, 1992). This is a matter of particular concern in an ecological approach, as non-reverberant environments (e.g., anechoic rooms, mountain tops) are extremely rare in the real world. On the other hand, again in real acoustic environments, spectral frequencies (that is, the frequencies of isolated components of complex sounds, or clear peaks in a running spectrum, forming frequency trajectories in time-varying sounds) cannot be directly affected by reflection off, or transmission through, environmental obstacles. They might be indirectly affected as a byproduct of the effect that such manipulations can have on *amplitudes* (e.g., a weakly defined peak could be pushed sideways if amplitudes increased on one side and decreased on the other), but such phenomena could hardly affect *audible* sound spectra.

So, for the auditory system to reliably identify sound sources, it needs to *ignore* phase information, which is merely a constant distraction, and focus as far as possible on a signal's spectral frequencies (and to a lesser extent on

the relative amplitudes of individual components, keeping in mind that these, too, are affected by reflection and transmission). The ear's phase deafness with regard to pitch perception is thus a positive attribute. In fact, it may be regarded as an important phylogenetic achievement — the result of a long evolutionary process in which animals whose ears allowed phase relationships to interfere with the identification of dangerous or otherwise important sound sources died before they could reproduce. If this scenario is correct, then it is no surprise that we are highly sensitive to small changes in frequency, and highly *insensitive* to phase relationships within complex sounds.

Straightforward evidence of the ear's insensitivity to phase in the sounds of the real human environment has been provided by Heinbach (1988). He reduced natural sounds including speech (with or without background noise and multiple speakers) and music to their spectral contours, which he called the *part-tone-time-pattern*. In the process, he completely discarded all phase information. The length of the spectrum analysis window was carefully tuned to that of the ear, which depends on frequency. Finally, he resynthesized the original sounds, using random or arbitrary phase relationships. The resynthesized sounds were perceptually indistinguishable from the originals, even though their phase relationships had been shuffled.

It is nevertheless possible to create artificial stimuli for which clear, significant perceptual effects of phase relationships on perception can be demonstrated. Patterson (1973, 1987) demonstrated that listeners can discriminate two harmonic complex tones on the basis of phase relationships alone (presumably using timbre, not pitch). Various other experiments have demonstrated that phase shifts affect pitch salience, but not pitch. Moore (1977) demonstrated that the relative phase of unresolved harmonics affects the virtual pitch of 3-tone complexes; relative phase determined which of the various possible virtual pitches was the most salient. For resolved harmonics, however, Moore (1977) found that pitch discrimination is almost independent of phase relations for resolved harmonics; phase only affects partials falling within a single critical band or auditory filter (see also Houtsma & Smurzynski, 1990, Experiment IV). Kubovy and Jordan (1979) shifted the phase of one of 12 equal-amplitude harmonics and observed that this increased the salience of the harmonic, allowing it to segregate from the complex when the phase difference exceeded about 30 degrees; the pitch of the complex as a whole was not affected. Hartmann (1988) similarly showed that the audibility of a harmonic depends on its phase relationship with other harmonics. The model of Meddis and Hewitt (1991b) succeeded in modeling the more important of these various phase effects. For a recent review of effects of phase on timbre, and on the degree to which the partials of complex tones mask each other, see Terhardt (1998, p. 380–382).

In an ecological approach such as Leman's, the existence of phase sensitivity in such stimuli (or such comparisons between stimuli) might be explained as follows. These stimuli (or stimulus comparisons) do *not* normally occur in the human environment. So the auditory system has not had a chance to 'learn' (e.g., through natural selection) to ignore the phase effects. As hard as the ear might 'try' to be phase deaf in the above cases, some phase sensitivity will always remain, for unavoidable physiological reasons.

There could, however, be some survival value associated with the ability to use phase relationships to identify sound sources during the first few tens of milliseconds of a sound, before the arrival of interference from reflected waves in typical sound environments. For example, partials are more likely to be in phase if the source is close, so phase relationships could be a cue to distance, with obvious implications for survival (Alex Galembó, personal communication, 1998). On this basis, we might expect phase relationships at least to affect *timbre*, even in familiar sounds. Supporting evidence for this idea in the case of synthesized musical instrument sounds has recently been provided by Dubnov and Rodet (1997). In the case of speech sounds, Summerfield and Assmann (1990) found that pitch-period asynchrony aided in the separation of concurrent vowels; however, the effect was greater for less familiar sounds (specifically, it was observed at fundamental frequencies of 50 Hz but not 100 Hz). In both cases, phase relationships affected timbre but not pitch.

The model that Leman chose for most of the simulations in his book is based on that of Meddis & Hewitt (1991a), which is capable of accounting for known phase dependencies in pitch perception (Meddis & Hewitt, 1991b). This raises the question: why might it be necessary or worthwhile to model something that does not have demonstrable survival value for humans (whereas music apparently does have survival value, as evidenced by the universality of music in human culture)? As Bregman (1981) pointed out, we need to 'think about the problems that the whole person faces in using the information available to his or her sense organs in trying to understand an environment' (p. 99). From this point of view, the human ear might be better off without any phase sensitivity at all. Bregman goes on to say that:

Because intelligent machines are required actually to work and to achieve useful results, their designers have been forced to adopt an approach that always sees a smaller perceptual function in terms of its contribution to the overall achievement of forming a coherent and useful description of the environment.

So if one were building a hearing robot, there would be no point in incorporating effects of phase on pitch perception, if such effects did not help the robot to identify sound sources.

Even if phase dependencies in pitch perception do play some role in our everyday interaction with the world, they are most unlikely to play any role in our perception of music, least of all in the perception of tonality, which is the primary research object of Leman's book. Phase relationships play no known role in the perception or appreciation of musical pitch structures. In the Western tradition, music is invariably heard in confined spaces filled with surfaces off which sound bounces before reaching the listener's ears, effectively randomizing phase relationships. These reflections are not regarded as something to avoid, but are an integral part of the musical experience. If music perception is being modeled, then, to account for phase sensitivity might best be regarded as redundant and computationally inefficient.

SELF-ORGANIZING MAPS

In Chapter 6, Leman introduces a self-organizing map based on Kohonen (1984), and argues cogently that such a map must play an important role in any neurophysiologically realistic model of pattern recognition — in particular, the perception of tonality. In an independent review of Leman's book, Large (1997) has pointed out that, according to Leman's own presentation, the inclusion of this map in the model does not add any predictive power to the final model when the model is used to explore tonal relationships in music. The question mark that this observation places against the role of self-organizing maps in the perception of tonality demands subsequent analysis. This is a matter to which further modeling and theoretical attention could have been devoted.

Chapters 7 and 8 contain several figures in which the output of a self-organizing map is represented as a 2-D image called a *CN-map*. A characteristic neuron (CN) is a neuron that responds most strongly to the stimulus with which it is labeled. CN-maps show relationships between stimuli in a way that is reminiscent of multi-dimensional scaling (MDS) results, where the physical distance between two stimuli in the map is a measure of the psychological distance between them (their dissimilarity) — although the mathematical details are quite different in the two cases.

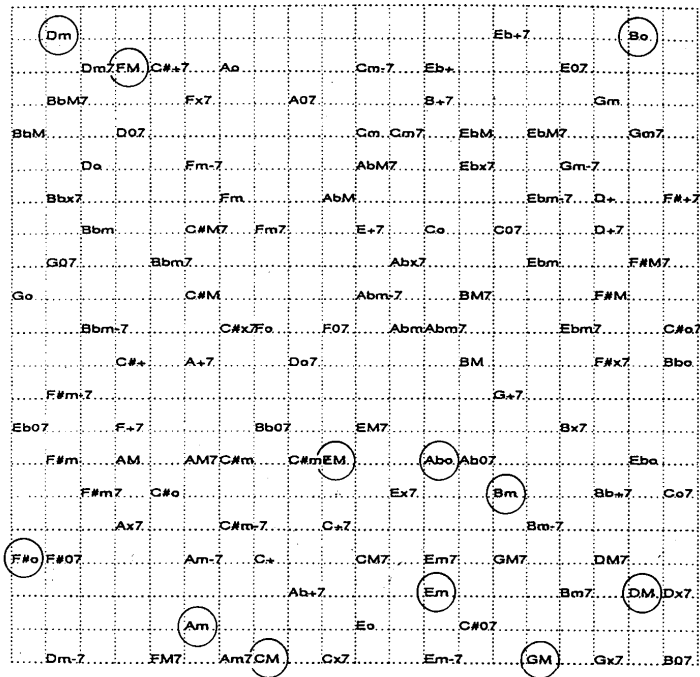


Fig. 1. Map of characteristic neurons (CN-map) corresponding to musical chords, after a network has been exposed 300 times to all chords shown (from Leman, 1995, Fig. 7.3d). ©Springer-Verlag Berlin Heidelberg 1995. Reproduced by permission. For ease of comparison, chords that also appear in Figure 2 are circled.

A question that immediately comes to mind is: is the relationship between CN maps and MDS solutions explicable in the light of the underlying neural architecture? I could not find a direct reply to this question in the book. However, ample quantitative evidence for such a connection is presented in Section 7.1 "SAMSON".

In that section, a total of 115 chords are fed through the model of Parncutt (1988) to produce chroma salience profiles, which in turn are used to train a self-organizing map (Kohonen, 1984) in the form of a two-dimensional grid of 20 × 20 neurons with a torus structure. The CN-map that results after a large number of cycles (Leman's Fig. 7d) is shown here as Figure 1.

To understand the figure, it is necessary to learn a set of somewhat unconventional chord symbols, including "Cx7" for C (major-minor or dominant) seventh (CEGBb, normally referred to simply as C7), "CO7" (without the familiar "/" through the "O") for C half-diminished seventh (CEbGbBb) by contrast to "Co7" for C diminished seventh (CEbGbBb), "C+7" for C augmented with major seventh (CEG#B, i.e. E/C – a somewhat surprisingly unusual member of the set), and "Cm-7" (*sic*) for C minor-major seventh (CEbGB, otherwise known as CmM7). The text and figure would benefit from the use of more familiar or Standard Chord Symbols (see e.g., Burbat, 1998, p. 172).

There is a general tendency in Figure 1 for chords that have tones in common, or that are regarded in music as related, to be close to each other. The CN-map thus resembles multidimensional scaling solutions of similarity data for musical chords such as those of Krumhansl, Bharucha, & Kessler (1982) shown in Figure 2. Such a comparison would not be entirely appropriate, how-

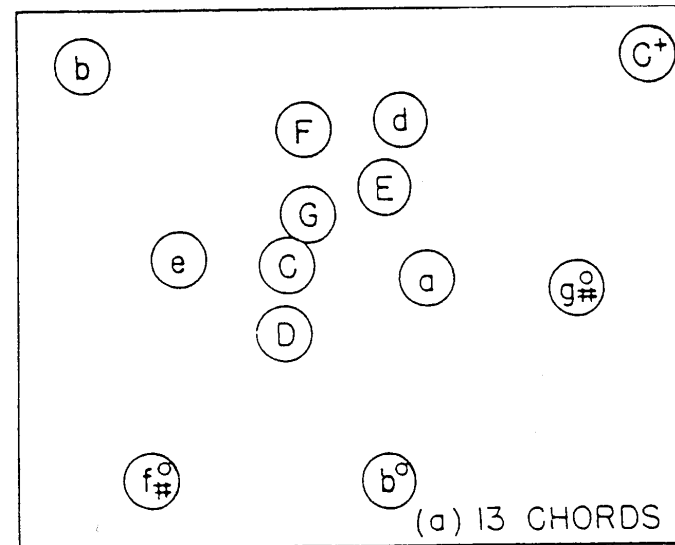


Fig. 2. Multidimensional scaling solution of similarity judgments of 12 chords when presented pairwise following an ascending scale in C major, G major, or A minor (from Krumhansl et al., 1982, p. 29, Fig. 2, top left panel). Similarity judgments were averaged over the three contexts. ©1982 American Psychological Association. Reprinted by permission.

ever, given that Krumhansl and colleagues obtained similarity judgements for chords in tonal contexts, while Leman's model applies to isolated pairs of chords. In Parncutt (1993), by contrast, listeners rated the similarity of 60 isolated chord pairs, and results were compared with correlation coefficients between 12-element chroma-salience profiles calculated according to a model based on Parncutt (1988) ("SAM") and additionally accounting for masking. Without actually doing the calculations, it is clear that a multidimensional scaling solution of the correlation coefficients (or pitch commonality) of all $(150 + 149 + \dots + 1)$ pairs of chroma salience profiles would resemble Leman's CN-map.

Summarizing these observations, Leman's treatment of harmonic relationships provides strong evidence in support of a self-organising neural basis, and paves the way for more detailed treatments and discussion of the relationship between his maps of characteristic neurons and MDS solutions, to address the broader implications of CN-maps for music perception and music theory. Some questions whose answers were not entirely clear from the book include the following. Is it always the case that the differences between perceived relationships and the configurations of the net gradually disappear as the net is exposed to more cycles of the chords? Or put another way, does the configuration into which the network stabilizes depend on the boundary conditions from which it begins? What implications might the answers to these questions have for the perception of chord relationships in music, and for the ability of someone steeped in one musical style to become fluent in another?

Returning for a moment to the earlier stage in the model that deals with pitch perception, self-organizing maps would have been ideal for modeling the harmonic pattern-recognition aspect of pitch perception as an emergent property of a neural network (cf. Bharucha & Mencl, 1996, in preparation). Leman chose instead to extract the pitch of complex tones using a periodicity model (cf. Langner & Schreiner, 1988; Langner has also shown how a neural net may be used to extract periodicity information). This again raises the question of whether such a complex model is appropriate or necessary in the context of this book, whose ultimate focus concerns musical tonality. It also raises an even more fundamental question in pitch perception: can periodicity 'cause' pitch, as Langner (1997) assumes, or is it merely associated with it? Leman points out that "the capacity of neurons to synchronize has been shown to be very important in the perception of low pitch and timbre" (p. 39), but this observation is only really causal at the level of physiology — the effect of one physiological parameter on another. Causation is inherently much more difficult to demonstrate between a neurophysiological parameter on the one hand (here, neural synchrony) and an experiential one on the other (perceived pitch or timbre). Given that pitch, like any experiential parameter, could (and pre-

sumably does) depend on a range of different neural parameters, and that periodicity and harmonicity are hard to disentangle in everyday sound signals, it is difficult to determine the extent to which pitch depends on periodicity, as opposed to other parameters such as harmonic pattern-recognition. The relative importance of these two processes is unknown; all that we can be sure of is that both play a role. Perhaps the real strength of Leman's approach is its future potential for combining both approaches, allowing information from each to be combined to reduce signal ambiguity (which is presumably something like what the auditory system actually does). Thus, a future, even more complex (and even more physiologically realistic) model could incorporate simulations of both periodicity and pattern-recognition models of pitch, and combine information from the two sources.

A paper that has appeared about the same time as Leman's book, and that could provide suitable inspiration for future improvements, is the elegant and parsimonious — but somewhat computationally expensive — model of Cohen, Grossberg, and Wyse (1995). In their model, as in Terhardt's, an input spectrum is transformed into a spatial representation of pitch strengths; here, harmonic pitch-pattern recognition is simulated by a harmonic sieve. Their model departs significantly from Terhardt's at the peripheral stage, which is more physiologically realistic, making pitch predictions slightly sensitive to phase. Predictions of the model compare favorably with a wide range of psychoacoustic data.

A flexible, physiologically based approach such as Leman's clearly has more potential than its cognitive counterparts to bring about significant future advances in understanding of pitch perception, especially as computers continue to become faster. It is already possible to vary the temporal resolution of parts of his model corresponding to different parts of the auditory system (e.g., there is higher temporal resolution in the inferior colliculus than in the auditory cortex — see page 39 — as the inferior colliculus transforms much of the peripheral 'temporal' representation into a more 'steady-state' representation). In many applications, physiological modeling may eventually supersede cognitive modeling, although the latter may still be valued for its simplicity, and hence its usefulness, in specific applications, just as the principles of classical or Newtonian mechanics are still used decades after they were 'superseded' by more general quantum mechanical and relativistic principles. Physiologically based models along the lines of Leman's may one day outperform their merely cognitive forebears, but the latter need not be discarded — they will still have practical and heuristic value.

TONALITY

A large part of the book is concerned with the application of Leman's neural-net based model to the understanding of tonality perception in western music. One of Leman's main starting points is the work of Krumhansl, and one of his achievements is to model her well-known major and minor key profiles in two distinct ways. First, on page 123, Krumhansl's key profiles are modeled in a way similar to my own in Parncutt (1989) — by inputting the cadences used by Krumhansl & Kessler (1982) in their experiments — and similarly high correlations are found. Second, Leman presents actual music to a neural net and allows the profiles to emerge during self-organization.

A problem with Krumhansl's key profiles as they stand is that they are not derived from first principles but rely for their derivation on the pre-existence of subdominant-dominant-tonic cadences in major and minor keys. This problem is solved by Leman's second method of deriving the profiles. The results are presented in his figure 7.6 (although the exact procedure is unfortunately not very clear from the accompanying text). Here, Leman's model transcends the work of Krumhansl, breaks free from the symbolic constraints of music theory texts, and addresses the perception of something closer to a real musical environment.

Parncutt (1989), and Huron and Parncutt (1993) showed that Krumhansl's key profiles can be modeled by adding ('smearing') pitch-class salience profiles of tonal sonorities over time. This method, it turns out, only works in the case of chord progressions. It does not work for unaccompanied melodies, or polyphonic textures where the voice leading is clearly audible. Not surprisingly, Leman's approach, which incorporates smearing, does not do a better job of predicting the tonality of melodies and chord progressions than its precedents. Predictions of key center could be improved in a future implementation in which voice-leading (i.e., the perception of familiar melodic fragments and their associated tonal and harmonic implications) were learned by a neural network in the way that Eberlein (1994) and Auhagen (1994) anticipate. The various analyses in the book of the changing tonality of real pieces music are thus somewhat premature. Connectionist modeling may eventually turn out to be the best way to predict the perceived tonality of melodies and contrapuntal passages, but for the moment the problem simply has not been solved.

On page 83, Leman shows that a self-organizing map, when fed the output of a chord-root model for a range of commonly occurring chords in different transpositions, eventually develops a global organization resembling the cycle of fifths — suggesting that this music-theoretic construct could have its origins in neuronal organization (an idea suggested also by Bharucha, 1992). This is one of the book's more interesting findings, and one that raises unresolved

questions about the 'true' or 'causal' origin of the circle. But the text does not address the possibility that the cycle of fifths could also have originated outside of people's heads, rather than in the brain. First, it may be regarded as an emergent property of the limited number of ways in which musical pitches can be organized (cf. Balzano, 1980). Second, it may have emerged originally from the simple physical constraints of musical instruments; as Eberlein (1994) has pointed out, string instruments have been tuned in fifths since antiquity, perhaps merely because the beats of a mistuned fifth are so clearly perceptible. In sum, it is not yet entirely clear where the *real* origin of the cycle of fifths lies; but Leman has certainly made an important contribution to the debate.

Chapter 9 of the book, *Schema and control*, presents the idea that, when we listen to music in major and minor keys, our perception of tonality tends to be "attracted" towards a particular key. In other words, we tend to be sure that we are in one key and not any other (a kind of categorical perception)? The idea of *attractor dynamics* is connected to the Gestalt notion of *Prägnanz*: the tendency to prefer 'good' forms in an ambiguous stimulus. This idea is intuitively appealing but I was surprised to see it presented without question; no empirical study is cited to justify it. My guess (and this would also require empirical testing) is that there is an expertise effect operating here, whereby music theorists listening to tonal music are more likely to show "tone center attraction" while other listeners are happier to accept ambiguities such that the music could effectively be in any of several different keys at once (or perhaps even the concept of perceived key is not particularly meaningful in the case of such listeners). In operational terms: you could play several different pitches following a given passage of music, ask listeners if they provide closure, and get the answer 'yes' significantly often to several of them (cf., e.g., Auhagen, 1994).

This raises the question of whether it is meaningful, in the context of a physiologically realistic model, to emulate a music-theoretical concept — key perception — when one is primarily concerned with the responses and experiences of an 'average' listener listening to tonal music. Perhaps it would be more appropriate to model something that listeners are more likely to be directly aware of, such as tension and relaxation (for which several papers have been published since Leman's book, e.g. Bigand, Parncutt, & Lerdahl, 1996; Krumhansl, 1996; Lerdahl, 1996), and only then relate this to more abstract qualities such as tonality.

At the end of the book, some of the models introduced earlier are applied to the analysis of real pieces of music. The input in every case is not the score, but the sound of a real performance. The output is a graph predicting how likely the music is to be (perceived to be) in any of the 24 major and minor

keys at each instant (Leman's Fig. 9.7) — much the same as output generated by Huron and Parncutt (1993). Leman's analysis includes an attempt to account for the effect of segmentation of the musical surface on the tonality of melodies and contrapuntal passages. In Chapter 10, before analyzing Bartók's *Through the Keys* (*Mikrokosmos* no. 104), he first segments the surface into phrases, assuming that they are perceived independently, and later shows that predictions of the model improve considerably when segmentation is taken into account. In reality, though, the situation is more complex than Leman's treatment would suggest. Segmentation is often unclear and ambiguous; it happens on different hierarchical levels; and the salience of segment boundaries varies continuously from very weak to very strong. Future implementations of Leman's model could try to incorporate the complexity of segmentation structures in real music perception. Again, his approach would seem an excellent starting point for this difficult and, as yet, barely addressed problem.

On page 127, an important concept is mentioned — the listener's delay in locking onto a given key in a modulating passage, and resistance to modulation once a key has been established. Leman appropriately describes this as a kind of *hysteresis* and models it by attractor dynamics with an elastic snail-like metaphor. But even with this sophisticated addition, the predicted key still seems to jump back and forth inappropriately. Recently, Vos and Van Geenen (1996) have successfully addressed this problem using a much simpler, but less ecologically valid, cognitive model. An incorporation of some features of Vos's model into that of Leman may allow this problem to be addressed more satisfactorily in the future.

On page 137, Leman presents an informal test of the models' predictions of tonal centers. Methodologically, the test is not very convincing. Musicologists are presented with the output of the model (which in this case consists of predicted major and minor keys instantiated by or touched upon in a passage of music), and asked to indicate whether this output does or does not correspond to the musicologist's analysis of the score. This tends to force the musicologist to look at keys that s/he would not otherwise have considered. It would have been preferable to (i) collect the tonal analyses independently of the model predictions, and then compare; (ii) have the musicians do the analyses by sound alone, for that is the process that the model is simulating; and (iii) test the predictions of the model against predictions of competing cognitive models such as Huron and Parncutt (1993), and Vos and Van Geenen (1996). That said, Leman does succeed in demonstrating that the model does a reasonable job, and in making clearer how it could be improved.

GENERAL PRESENTATION

The book benefits greatly from Leman's background as editor of the *Journal of New Music Research* and his encyclopedic acquaintance with relevant literature, both old and new, in several European languages. As research in systematic musicology becomes increasingly international, the ability to cross both national and linguistic divides is again becoming important, especially as Europe re-identifies itself as a multicultural unit in which no single language predominates (or I should say: in which efforts are being made to stop English taking over).

While on the topic of language, I should mention that Leman's English is a little odd at times, and I am surprised that the editors at Springer did not correct the style more carefully. Also, Leman uses the French/Italian *do re mi* (which Springer renders somewhat frighteningly in bold capitals) for absolute pitches, instead of C D E as is customary in English (and German). These minor quibbles aside, the writing style is refreshingly concise. Leman does not rave on, but gets quickly to the point. His sentences are often worth reading twice — and, not being an expert on connectionism, I have read some of them many more times than that.

I was disappointed by the preponderance of jargon in the book, much of which, it seems, could have been avoided and does not seem to make the material any easier to understand — although it could, admittedly, be quite helpful for programmers working on implementing the model and in need of convenient labels for functions and variables. Readers of this article who have the book to hand will notice, for example, that on page 21 the term *toneness* is introduced to replace good old *chroma*. Leman's justification: *chroma* suggests absolute properties, like colors. But his proposed alternative, *toneness*, is so vague and non-self-explanatory as to be worse! On page 30, Leman introduces Risset's terms *hauteur spectral* and *hauteur tonale*, without pointing out that they are essentially the same as Terhardt's spectral pitch and virtual pitch (presented later, on page 43).

While on the topic of jargon, I should mention that my model of musical chord-roots (Parncutt, 1988) is referred to repeatedly throughout the book as Simple Auditory Model or SAM, giving the impression that the model is of Leman's creation. The acronym SAM is easy to remember and contrast with other acronyms TAM (Terhardt et al., 1982) and VAM (Van Immerseel, 1993), but these abbreviations are not entirely appropriate, given that the models to which they refer (especially SAM and TAM) are not really "auditory" but primarily about *pitch* — not loudness, timbre, or any other auditory attribute. Might it have been better to call it Simple Octave-Generalized Pitch Model (SOGPIM?). Easier still, Parncutt (1988)? Similarly, the model of Terhardt

et al. (1982) is called TAM (“Terhardt Auditory Model” — but again, it is really a pitch model), but could more easily be referred to in the accepted way as “Terhardt et al. (1982)”.

The jargon continues — at times, quite thick and fast. On page 119, one encounters the abbreviations CN, CRR, and TCAD in a single paragraph. To find out what these mean (*characteristic neuron*, *characteristic response region*, and *tone center attraction dynamics*, respectively) one needs to hunt about on previous pages. Many of these abbreviations could have been avoided altogether for the benefit of the general reader.

Another aspect of the book that makes it difficult to understand is the somewhat sparse annotation of many of the figures. In particular, the figures that present output of the neural-net simulations could do with some extra axis labels and more detail in the captions for the benefit of the general reader. To find out what the vertical axis of Fig. 9.2 (reproduced here as Figure 3) meant, I eventually returned to page 57, where I found evidence suggesting that the axis in question must represent time lags in an auto-correlation. Further detective work revealed that each number on the vertical axis of the figure should be multiplied by 0.4 ms to produce real time values. Things would have been much easier had that been explained on the figure itself (as it is in a later but similar figure: Leman & Carreras, 1997, Fig. 2). And that’s not all! The figure in question has no labels at all on the horizontal axis (one assumes it is a time axis, but relative to what?). Another example: Fig. 9.4 includes both full and dotted lines, but there is no indication in the caption as to their meaning of the distinction between them. Again, detective work is required. After some searching I discovered that the full lines had already appeared in Fig. 2.8, seven chapters earlier, and concluded that the dotted lines in Fig. 9.4 must be new.

A final problem with presentation, in addition to the problems mentioned above, is simply that the text is often unnecessarily hard to understand. It seems that much of the argumentation could have been rendered much more intelligible by talking the author through the technical sections of the book, asking at each new paragraph what it really means, and writing down the reply in terms that everyday musicians and music psychologists will understand.

These quibbles about presentation may seem boring and unnecessary in the context of an article of this kind. I have included them for two main reasons. First, the specific points raised in this discussion will help readers of the book to understand it. Second, and more importantly, I wish to address the general issue of effective cross-disciplinary communication. This book uses scientific means to explain music structure, and as such could be of great interest to musicologists, music theorists, and music analysts. Yet few such peo-

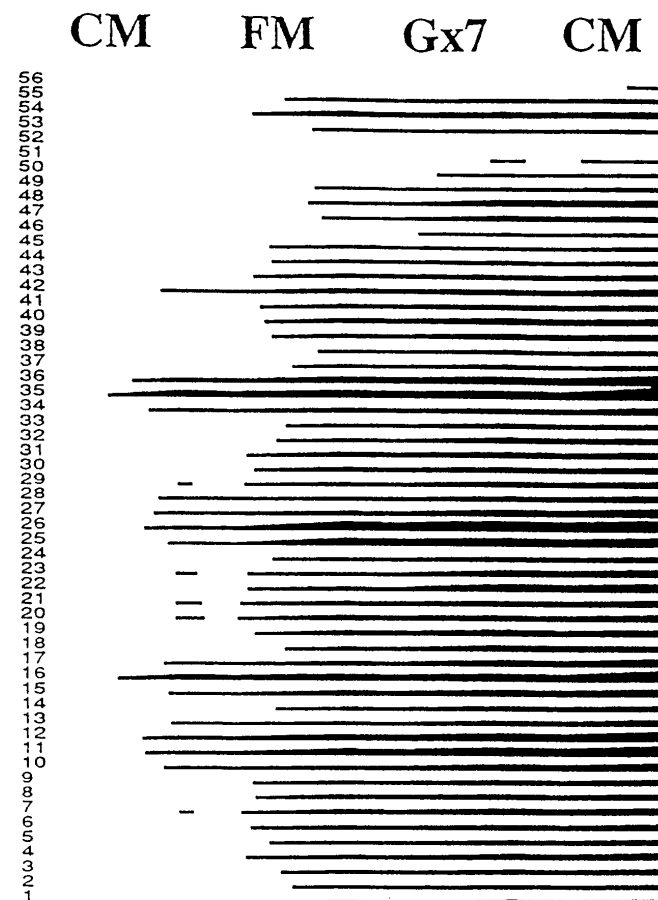


Fig. 3. Tone context image of the Shepard-chord sequence C-F-G7-C (from Leman, 1995, Fig. 9.2); note that Leman calls G7 (=GBDF) “Gx7”. Vertical axis: time lags in an auto-correlation function (0.4 ms per unit). Horizontal axis: time. ©Springer-Verlag Berlin Heidelberg 1995. Reproduced by permission.

ple will ever have the time, energy and/or background to wade through a dense, jargon-ridden text. There is a lesson to be learned here for the discipline of systematic musicology as a whole, a field that is struggling to be recognized by mainstream musicologists (Leman & Schneider, 1997; Parncutt, in press). In a very real sense, the overuse of jargon can cause the work of systematic musicologists to fall on deaf ears.

HISTORICAL AND ACADEMIC CONTEXT

During this century, the disciplines of systematic musicology and music psychology have fallen into disrepute in historical musicological circles. ‘Scientific’ approaches to musicology have especially been unable to account for important sociohistorical developments such as the 20th-century dismantling of major-minor tonality, and large cross-cultural variations in tonal systems. We have known for a long time that the way we respond to musical stimuli is largely determined by the music to which we have already been exposed, and subsequently the response to any given sound depends on the context in which it is embedded; but we have not been in a position to model the various stages of this complex process.

Leman’s approach shows a new way through. It heralds a significant paradigm shift in cognitive musicology, from cognitive models based on intuition and the results of psychoacoustic experiments to models based on physiological modeling of the auditory system. Although the new approach is still in its infancy, it is already starting to make its presence felt. Leman sums up his approach as follows:

The cognitive structures for tone perception are supposed to be learned, but the statement that listeners internalize the summary statistics of tones in music is far too general. Such structures can only be fully understood if we understand something about their development. Hence, learning mechanisms should be studied, as well as the musical conditions under which schemata emerge (p. 18).

Accordingly, Leman’s model learns gradually (“data-driven long-term dynamics”, p. 135) but can quickly apply the results of that learning (“schema-driven short-term dynamics”). In this sense it may be regarded as more realistic than previous, more ‘pre-wired’ models such as Bharucha’s (1987) model of spreading activation.

The second-last chapter of the book (Ch. 12, *Epistemological Foundations*) is one of the most satisfying. It brings together Leman’s state-of-the-art computational skills and his well-grounded philosophical insights. After that, the last chapter puts the book as a whole into its historical context, and optimistically predicts a revival of the disciplines of cognitive and systematic musicology based on a new convergence of various disciplines related to music — psychology, computing, musicology, philosophy and neuroscience (see Leman, 1997 for further remarks on the issue of convergence). The chapter reads like a manifesto for the future progress of music cognition research, and systematic musicologists could do worse than to put their collective weight behind it. This is a feel-good chapter that is honest about the limitations of the approach but at the same time realistically confident about future possibilities. An inspirational companion to this text is the first chapter of Leman’s more recent edited book (Leman & Schneider, 1997), which also looks at the past, present and future of systematic musicology and emphasizes the important role of gestalt psychology, cognition, neuroscience, and computing in modern systematic musicological research.

This last chapter also makes explicit one of the book’s most attractive and original features: its ecological approach. An important player in Leman’s ecological scenario is the computer modeler, who limits his/her role to an objective description of the conditions in which an interaction can take place between a physical world of musical sound and vibration (in principle, complete with social constraints) and a physiological model of information processing. The knowledge which is built up by the human information processing system is obtained by interaction with the environment by means of self-organisation. In principle, this helps Leman to sidestep, at least in part, one of the most important confounds that plague research of this kind, and indeed all research in music theory — the musical experience and biases of the theorist or scientist.

Leman presents a theory (in the scientific sense of the word) that is capable of explaining aspects of the perception of musical structure. Yet few music theorists would seriously consider this to be a book about music theory, let alone think about buying it — or any other book on music psychology, for that matter. Why? Two main reasons come to mind. First, books like this tend to contain lots of technical stuff that is hard to understand and is often irrelevant for mainstream music theory. Second, books like this tend to lack musical sophistication.

Imagine you are a music theorist (perhaps you are one!) standing in a bookstore, flicking through the book, and wondering whether to buy it. You open the book randomly at page 139, where you see that Bartók’s *Through the Keys* from the *Mikrokosmos* is chosen for analysis (reproduced here as Fig. 4). At last, here is a page that you can immediately understand — but there’s



Fig. 4. First 6 bars of Bartók: *Through the Keys* (also known as *Wandering Through the Keys*) (from Bartók, 1940, p. 20). ©1940 by Hawkes & Son (London) Ltd. Reproduced by kind permission of Boosey & Hawkes Music Publishers Ltd.

a problem. The piece is said to allude to the keys of E minor or G major in measure 3 — but from an everyday music-analytic viewpoint, the first four bars of the piece represent about the clearest imaginable example of the key of D major. The author of this book seems to have misunderstood Bartók's title as including key movement within the opening statement of D major! You put the book back on the shelf.

Leman's is certainly not the only book in the area of systematic musicology that is both too technical and insufficiently musically sophisticated for a typical musicological or music-theoretical audience. On the other hand, it is equally unrealistic to expect a computational approach of this kind to deal appropriately with the subtleties of music theory when scientists are still grappling with explanations of aspects of music theory that music theorists take for granted. At the very least, one could aim to avoid any statement that a mainstream musicologist would regard as erroneous or naïve.

These objections aside, the book is to be commended for taking a sizable step in the direction of coupling together of real musical sophistication with real scientific rigor — a goal that, I feel, should be close to the heart of any genuine music theorist. I consequently do not hesitate to recommend it, warts and all, not only to music psychologists but also to music theorists who have the time and energy to devote to the scientific aspects of their discipline.

SUMMARY

Taking a broad view, Leman's book has been instrumental in enabling the following two important steps to be taken in understanding music perception and cognition.

1. Complex music-perceptual phenomena are beginning to become accessible to exploration at the level of brain function. This idea has been around for a long time, but until now it could hardly be taken seriously, due to a combination of the brain's inherent complexity (are we intelligent enough to understand our own intelligence?) and the discouragingly serious practical and ethical problems associated with direct observation of brain function (most of the underlying physiological data comes from animal research). The book provides musically important and relevant examples of neurophysiological explanations of musical phenomena, especially tonality. Perhaps most importantly, the book demonstrates the potential for future development in the neurophysiologically based understanding not only of tonality but also of timbre, rhythm, emotional responses, and interactions between these.

2. It is no longer necessary for models of perception to operate in isolation from real soundworlds — although non-ecologically-valid cognitive models will always have a useful role to play in music psychology and music theory. As computer power has increased, it has been increasingly feasible to run feasible subsymbolic (neural-net) simulations of human perception. Perceptual models are increasingly able to realistically mimic the interaction between humans and their sound environment, and the developmental shaping of perception by that interaction. This new tool will become increasingly useful and pertinent for music historians and ethnomusicologists (most of whom appear at present to be blissfully unaware of what is around the corner!) because it will allow them to study, in a surprisingly concrete manner, how the perception of music by individuals is influenced by the musical patterns and styles to which they have been exposed.

As in any pioneering effort, the book has a number of problematic features, some of which are serious enough to prevent potential customers from investing in it (and it must be said — Springer books are not cheap). But without mistakes there can be no learning! The book has enabled the following three lessons to be learned:

1. Physiologically realistic models can only really be described as *interesting* if they not only formalize physiological knowledge, but also predict experimental data as accurately as equivalent heuristic (psychoacoustic/cognitive) models. From a musically pragmatic viewpoint, they can only be *useful* if they perform even better than their heuristic equivalents. According to these criteria, the models presented in Leman's book may be regarded as generally interesting, but not useful. None of the models presented reliably predicts quantitative

data on tonal relationships between musical sounds significantly more accurately than existing heuristic models such as Huron & Parncutt (1993), Krumhansl (1996), Parncutt (1988, 1989, 1993), and Terhardt et al. (1982) — even though the more complex models tend to provide more opportunities for ‘tweaking’ arbitrary constants, coefficients, and free parameters to improve correlations. But the science of applying neurophysiologically realistic models to music-perceptual data is still in its infancy, and although the problem of quantitative performance has been addressed again more recently but still not solved (see Leman & Carreras, 1997, Tables 1 and 2), it seems likely that future implementations will become both interesting and useful.

2. It is not generally possible to demonstrate the realism of a complex, neurophysiologically realistic model by merely showing that it works — that is, by demonstrating significant correlations between model predictions and a range of experimental data. Realism, in this case, can only be demonstrated by convergent evidence confirming the various stages of the model individually. Moreover, if certain stages of the model do not contribute to the predictive power of the model, then the predictive power of the model as a whole cannot be passed off as evidence for the realism of those individual stages. For the more complex models that Leman presents in his book, it is often the case that one or more stages could be omitted from the models without seriously affecting their predictive power. The book could have been much more honest about which parts of which models did not help to predict which results, to reassure the reader that the author is not only interested in creating models that do exciting things, but also in getting rid of models, or parts of models, that quantitatively simply do not produce the goods.

3. Jargon is a major interdisciplinary barrier. If we systematic musicologists want mainstream musicologists to take us seriously, then we will need to try harder to avoid jargon. We need to write in a language that musicians and musicologists can understand, while at the same time maintain an appropriate level of scientific rigor. This is not so difficult to do as it sounds, but it does require some practice. A useful strategy is to ask musical colleagues to comment on clarity of presentation, and to address the conceptual issues that they raise. Similarly, ‘scientific’ modes of presentation in articles, especially graphs and equations, require clear, thorough, but still concise explanation. If a book is not understood by its target audiences, then it has failed to achieve its primary goal, regardless of the inherent value of its content.

The overall flavor of this book is not one of sureness and completeness, but instead one of rapid evolution of new ideas. Much of this work has only really got off the ground in the past decade, and for that reason many of the results presented in the book are rather preliminary. In another ten years things will look rather different. In fact, in some ways they look rather different already, as increasingly large computers are being used for some of the modeling presented in the book (Leman & Carreras, 1997). That only goes to show that this book is having, and will continue to have, a significant impact on research in music perception and systematic musicology. Get it soon, it is almost out of date already!

REFERENCES

- Auhagen, W. (1994). *Experimentelle Untersuchungen zur auditiven Tonalitätsbestimmung in Melodien*. Kassel: Bosse.
- Balzano, G.J. (1980). The group-theoretical description of 12-fold and microtonal pitch systems. *Computer Music Journal*, 4, 66–84.
- Bartók, B. (1940, corrected 1987). *Mikrokosmos 4*. London: Boosey & Hawkes (UK no. 8004).
- Bharucha, J.J. (1987). Music cognition and perceptual facilitation: A connectionist framework. *Music Perception*, 5, 1–30.
- Bharucha, J.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, UCLA, CA (abstract).
- Bharucha, J.J. (1994). Tonality and expectation. In R. Aiello & J.A. Sloboda (Eds.), *Musical perceptions* (pp. 213–239). Oxford: Oxford University Press.
- Bharucha, J.J., & Mencl, W.E. (1996). Two issues in auditory cognition: Self-organization of octave categories and pitch-invariant pattern recognition. *Psychological Science*, 7, 142–149.
- Bharucha, J.J., & Mencl, W.E. (in preparation). A self-organizing model of pitch.
- Bigand, E., Parncutt, R., & Lerdahl, F. (1996). Perception of musical tension in short chord sequences: Influence of harmonic function, sensory dissonance, horizontal motion, and musical training. *Perception & Psychophysics*, 58, 125–141.
- Bregman, A.S. (1981). Asking the ‘What for?’ question in auditory perception. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization* (pp. 99–118). Hillsdale, N.J.: Lawrence Erlbaum.
- Bregman, A.S. (1990). *Auditory scene analysis*. Cambridge, MA: MIT Press.
- Burbat, W. (1988). *Die Harmonik des Jazz*. München: dtv/Bärenreiter.
- Cohen, M.A., Grossberg, S., & Wyse, L.L. (1995). A spectral network model of pitch perception. *Journal of the Acoustical Society of America*, 98, 862–879.
- Dubnov, S., & Rodet, X. (1997). Statistical modeling of sound aperiodicities. *Proceedings of the International Computer Music Conference, Thessaloniki, Greece*, (pp. 43–50). San Francisco: Computer Music Association.
- Eberlein, R. (1994). *Die Entstehung der tonalen Klangsyntax*. Frankfurt: Peter Lang.
- Hartmann, W. (1988). Pitch perception and the segregation and integration of auditory entities. In G.M. Edelman, W.E. Gall, & W.M. Cowan (Eds.), *Auditory function* (pp. 623–645). New York: Wiley.

- Heinbach, W. (1988). Aurally adequate signal representation: The Part-Tone-Time-Pattern. *Acustica*, 67, 113–121.
- Houtsma, A.J.M. (1979). Musical pitch of two-tone complexes and predictions by modern pitch theories. *Journal of the Acoustical Society of America*, 66, 87–99.
- Houtsma, A.J.M. & Goldstein, J.L. (1972). The central origin of the pitch of complex tones: Evidence from musical interval recognition. *Journal of the Acoustical Society of America*, 51, 520–529.
- Houtsma, A.J.M., & Smurzynski, J. (1990). Pitch identification and discrimination for complex tones with many harmonics. *Journal of the Acoustical Society of America*, 87, 304–310.
- Huron, D., & Parncutt, R. (1993). An improved model of tonality perception incorporating pitch salience and echoic memory. *Psychomusicology*, 12, 152–169.
- Kohonen, T. (1984). *Self-organisation and associative memory*. Berlin: Springer.
- 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., Bharucha, J.J., & Kessler, E.J. (1982). Perceived harmonic structure of chords in three related musical keys. *Journal of Experimental Psychology: Human Perception and Performance*, 8, 24–36.
- Krumhansl, C.L. & Kessler, E.J. (1982). Tracing the dynamic changes in perceived tonal organisation in a spatial representation of musical keys. *Psychological Review*, 89, 334–368.
- Kubovy, M., & Jordan, R. (1979). Tone-segregation by phase: On the phase sensitivity of the single ear. *Journal of the Acoustical Society of America*, 66, 100–106.
- Langner, G. (1997). Temporal processing of pitch in the auditory system. *Journal of New Music Research*, 26, 116–132.
- Langner, G., & Schreiner, C.E. (1988). Periodicity coding in the inferior colliculus of the cat (I) — Neuronal mechanisms. *Journal of Neurophysiology*, 60, 1799–1822.
- Large, E.W. (1997). Review of Leman (1995). *Music Perception*, 14, 319–325.
- Leman, M. (1995). *Music and schema theory: Cognitive foundations of systematic musicology*. Berlin, Heidelberg: Springer.
- Leman, M. (1997). The convergence paradigm, ecological modeling, and context-dependent pitch perception. *Journal of New Music Research*, 26, 133–153.
- Leman, M., & Carreras, F. (1997). Schema and gestalt: Testing 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). Berlin: Springer.
- Leman, M., & Schneider, A. (1997). Origin and nature of cognitive and systematic musicology: An introduction. In M. Leman (Ed.), *Music, gestalt, and computing: Studies in cognitive and systematic musicology* (pp. 13–29). Berlin: Springer.
- Lerdahl, F. (1996). Calculating tonal tension. *Music Perception*, 13, 319–364.
- Meddis, R., & Hewitt, M.J. (1991a). Virtual pitch and phase sensitivity of a computer model of the auditory periphery. I: Pitch identification. *Journal of the Acoustical Society of America*, 89, 2866–2882.
- Meddis, R., & Hewitt, M.J. (1991b). Virtual pitch and phase sensitivity of a computer model of the auditory periphery II: Phase sensitivity. *Journal of the Acoustical Society of America*, 89, 2883–2894.
- Meyer, L.B. (1956). *Emotion and meaning in music*. Chicago: University of Chicago Press.
- Moore, B.C.J. (1977). Effects of relative phase of the components on the pitch of three-component complex tones. In E.F. Evans, & J.P. Wilson (Eds.), *Psychophysics and physiology of hearing* (2nd ed. pp. 349–362). New York: Academic Press.
- Moore, B.C.J. (1989). *An introduction to the psychology of hearing* (3rd ed.). London: Academic Press.
- 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. (1993). Pitch properties of chords of octave-spaced tones. *Contemporary Music Review*, 9, 35–50.
- Parncutt, R. (in press). Implications for mainstream musicology of recent research in music cognition. *Proceedings of the International Musicological Society*.
- Patterson, R.D. (1973). The effects of relative phase and the number of components on residue pitch. *Journal of the Acoustical Society of America*, 53, 1565–1572.
- Patterson, R.D. (1987). A pulse ribbon model of monaural phase perception. *Journal of the Acoustical Society of America*, 82, 1560–1586.
- Popper, K.R., & Eccles, J.C. (1977). *The self and its brain*. Berlin: Springer.
- Shamma, S.A. (1985). Speech processing in the auditory system I: The representation of speech sounds in the responses of the auditory nerve. *Journal of the Acoustical Society of America*, 78, 1612–1621.
- Summerfeld, Q., & Assmann, P.F. (1990). Perception of concurrent vowels: Effects of harmonic misalignment and pitch-period asynchrony. *Journal of the Acoustical Society of America*, 89, 1364–1377.
- Terhardt, E. (1988). Psychoakustische Grundlagen der Beurteilung musikalischer Klänge. In J. Meyer (Ed.), *Qualitätsaspekte bei Musikinstrumenten* (pp. 9–22). Celle: Moeck.
- Terhardt, E. (1991). Music perception and sensory information acquisition: Relationships and low-level analogies. *Music Perception*, 8, 217–240.
- Terhardt, E. (1992). From speech to language: On auditory information processing. In M.E.H. Schouten (Ed.), *The auditory processing of speech* (pp. 363–380). Berlin: Mouton de Gruyter.
- Terhardt, E. (1998). *Akustische Kommunikation: Grundlagen mit Hörbeispielen*. Berlin: Springer.
- Terhardt, E., Stoll, G., & Seewann, M. (1982). Algorithm for extraction of pitch and pitch salience from complex tonal signals. *Journal of the Acoustical Society of America*, 71, 679–688.
- Todd, N. (1994). The auditory 'primal sketch' — a multiscale model of rhythmic grouping. *Journal of New Music Research*, 23, 25–70.
- Toiviainen, P. (1996). Optimizing auditory images and distance metrics for self-organizing timbre maps. *Journal of New Music Research*, 25, 1–30.
- Van Immerseel, L. (1993). *Een functioneel gehoormodel voor de analyse van spraak bij spraakherkenning*. PhD thesis, University of Ghent.
- Vos, P.G. & Van Geenen, E.W. (1996). A parallel-processing key-finding model. *Music Perception*, 14, 185–225.
- Zatorre, R.J. (1988). Pitch perception of complex tones and human temporal-lobe function. *Journal of the Acoustical Society of America*, 84, 566–572.



Richard Parncutt
Department of Musicology
University of Graz
Mozartgasse 3
8010 Graz
Austria
E-mail: richard.parncutt@kfunigraz.ac.at

A cognitive musicologist, music theorist, and pianist. Recently appointed Associate Professor of Systematic Musicology at the University of Graz. Previously, Lecturer in Psychology of Music and Psychoacoustics at Keele University. Before that, research fellow in Munich, Stockholm, Halifax (Nova Scotia), and Montreal. Taught undergraduate music theory in Australia, Canada and England, and music perception and psychoacoustics in Germany and Canada. Associate Editor of *Contemporary Music Review*, *Journal of New*

Music Research, and *Psychology of Music*; Editorial Board of *Systematische Musikwissenschaft*; Advisory Board of *Jahrbuch Musikpsychologie*. Author of *Harmony: A Psychoacoustical Approach* (Springer-Verlag, 1989) and research articles on the perception of harmony, tonality, microtonality, musical tension, rhythm, psychoacoustic and cognitive theories and models, computer-based composition and ear training, expression and fingering in piano performance, prenatal audition, and music notation.