Review of "Music and Probability" by David Temperley
(Cambridge, MA: MIT Press, 2007)
Richard Parncutt

Psychology of Music, in press 2007

NOTE TO TYPESETTER: PLEASE SEARCH FOR THE HAT SYMBOL ^ THROUGHOUT THE DOCUMENT AND SHIFT THE HATS TO THE RIGHT SO THAT THEY ARE DIRECTLY ABOVE THE FOLLOWING NUMBER. THIS IS A MUSIC-THEORETICAL CONVENTION FOR REFERRING TO THE STEPS OF A MUSICAL SCALE.

Temperley's book is the first to apply Bayes' theory of conditional probability, in depth and detail, to questions of music theory and music psychology. Since these two disciplines often address questions of probability – either directly or indirectly, quantitatively or qualitatively - the topic is, or should be, of central interest to researchers from both disciplines.

Temperley begins with a quote from Leonard Meyer, who observed that the style of a piece of music largely depends on the probability that given elements (such as timbres or harmonies) will occur, both in absolute and relative senses. By "absolute" I mean independent of other musical context: if one tenth of the chords in a chord progression are C-major triads, the probability that a randomly selected chord from that progression will be a C-major triad will be 10%. By "relative" I am referring to the probability that an event will occur given *some other condition*. For example, if we know in advance that a chord progression is in C major, we might expect that the proportion of C-major triads will be higher than 10%. Put another way, the probability that a chord will be a C-major triad is *conditional* on the key of the passage. Since such conditional probabilities depend on musical style, they can be used to describe style quantitatively.

Conditional probabilities are related to *subjective probabilities*: the degree to which a human observer expects that something will happen given a set of initial conditions. Since expections play such an important role in the perception of music (Huron, 2006), one might expect Bayes' theory to play a similarly important role in music perception and theory.

To understand Bayes' theory in more detail, consider the probability of an event *A* occurring given that event *B* also occurs. That probability is generally different from its inverse, i.e. the probability that *B* will occur given that event *A* also occurs. Thus, the probability that the tone C will happen in the context of C major is different from the probability that the passage surrounding any tone C will be in C major. But the two probabilities are evidently related. According to Bayes, if you take the probability that event *A* will occur given *B,* and multiply it by the absolute probability of *B*, you will get the same numerical result as if you take the probability that event *B* will occur given *A,* and multiply it by the absolute probability of *A*. If you understand that somewhat odd (but remarkably general) sentence, you have the essence of Bayes' idea. Another example: If you take the probability that tone C will happen in the context of C major (say, in music from the classical period) and multiply it by the absolute probability that C major will occur (in the same repertoire), the product will be the same as taking the probability that the passage surrounding a tone C will be in C major (in that repertoire) and multiplying that by the absolute probability that the tone C will occur (in that repertoire).

Reading Temperley's book made me realise that I had been thinking for many years in a Bayesian way without realising it. Perhaps other music psychologists reading this book (or this review) will have similar "aha" experiences. I had assumed in Parncutt (1988, 1989) and later related work that the most salient pitch in a chord corresponds to its root, and operationally defined salience as a probability of noticing (or consciously perceiving). In other words, I assumed that the root of a chord is the pitch that is most likely to be noticed in many different realisations (voicings, spacings, doublings, timbres, relative amplitudes etc.) of the chord. More generally, I assumed that the probability of perceiving a pitch as the root of a chord is proportional to the probability of perceiving that pitch *at all* within the chord. In Parncutt (1994), I made a similar assumption with respect to musical rhythm and meter: the downbeat of a repeating rhythmic pattern is the event that is most likely to be consciously perceived in different realisations of that pattern, and more generally, the probability of perceiving an event as the downbeat in a given pattern is proportional to the probability of perceiving that event at all in the same pattern. In both cases, I was assuming that the probability of perceiving a given musical structure (here, a simple harmonic or periodic pattern with corresponding root or pulse) beneath a given musical surface is proportional to the probability of perceiving that surface above that structure. Until reading Temperley's book I was unsure how I could justify these assumptions; his book has given me the answer.

Temperley applies Bayes' theory in a general way to the relationship between foreground and background in music. Schenker's use of these terms is of limited interest to music psychologists, since it would be far fetched to regard Schenker's concept of background as an accurate representation of the corresponding psychological reality. But Schenker's approach is psychologically useful in a more general sense. The foreground may be regarded as any events heard on the musical surface such as individual tones or noises, and the background as any musical structure that is perceived (or cognitively constructed) by the listener on the basis of musical experience. Defined in that way, a musical background can be a theme, scale, chord, tonality, rhythmic pattern or meter. Given that the meaning of a piece of music for an individual listener generally depends on which background structures are perceived in it and on their personal and cultural associations, this general concept of background is worthy of both music theoretical and psychological attention.

Bayes' theory enables us to study the probability of given foreground events occurring together with given backgrounds - an idea that has already been pursued in linguistics with considerable success. The trouble with music is that it is more complicated and abstract than speech. It is relatively easy to segment speech into a hierarchical structure of phonemes, words and phrases. Of course there are ambiguities, but they can generally be counted and kept track of. In music, there is an additional interaction between pitch and time. Gestalt theory and auditory scene analysis address the complex issue of which sound events group with which other events in these two competing dimensions. In general, there are many different possible groupings that compete with each other for the listener's attention. In music theoretic terms, we might say that there are generally different possible analyses of a given piece of music, even within the same theoretical system.

The problem becomes more tractable, both theoretically and empirically, when certain variables are selected for special attention. For example, tonality mainly involves the relative stability of the 12 chromatic scale steps; it also depends on variations among octave registers and tonal implications of specific voice-leading patterns, but in a first approximation these may be considered subsidiary. In considering the perception of tonality and the prediction of musical key, Temperley considers and compares several different "key profiles" (Table 6.1), which, in different ways, estimate the relative importance of the 12 chromatic scale degrees in the context of major and minor keys. His collection of profiles includes those of Krumhansl and Kessler (1982), which – 25 years on – are still recognized as empirical psychology's most important contribution to the theory of tonality. Temperley's collection of profiles also includes what he calls the *Kostka-Payne profiles*. These represent the frequency of occurrence of chromatic scale degrees in major and minor keys in the notated scores of a representative collection of common-practice tonal music (Kostka, 1995).

Temperley notes that models of tonality that use tone profiles to track modulations in multi-voiced music generally work better with the Kostka-Payne profiles than with Krumhansl's profiles. This is not surprising given that the Kostka-Payne profiles are derived directly from the frequency of occurrence of notes in scores. But Krumhansl's profiles do something else: they explain the relationship between the frequency of occurrence of notes in a score and the perceived degree of closure (or in music-theoretic terms, stability) of the corresponding pitches in terms of an underlying cognitive representation.

The interesting thing about the Kostka-Payne profiles for me is how they differ from Krumhansl's profiles at specific scale degrees. Temperley notes that Krumhansl's minor-key profiles emphasize the minor seventh scale degree, whereas the Kostka-Payne profiles emphasize the major seventh or leading tone. Fixing this problem by adjusting the weights in Krumhansl's profiles improves predictions of musical key (at least in "classical" music) but reduces the model's elegance (parsimony) and does not clearly explain the psychological function or the origin the leading tone.

Temperley fails to address other differences between the profiles (both major and minor) at the position of the tonic scale degree (^1). First, in Krumhansl's profiles, the value at ^1 is clearly greater than at all other scale degrees, while in the Kostka-Payne profiles, the values for ^1 and ^5 (the tonic and dominant scale degrees) are about equal in both major and minor. That is presumably because ^1 is more stable than ^5, because the root of the tonic triad is more perceptually salient than the fifth (Parncutt, 1988). Second, in Krumhansl's profiles, the second most stable pitch is ^5 in the major key and ^3 in the minor. In the Kostka-Payne profiles, the second most stable pitch is ^5 in both cases. Why did ^3 rate so highly in Krumhansl's experiments in the minor key? According to Parncutt (1988), the third is the second most salient pitch in a minor triad because it is reinforced by the major third interval between the chord's third and fifth.

My 1988 model was based on Terhardt's chord-root model (1982) and pitch algorithm (Terhardt et al., 1982 b). To my knowledge, no other theory successfully predicts the root (s) of all possible sonorities in the chromatic scale and accounts appropriately and quantitatively for their ambiguity (Parncutt, 2007). The theory is parsimonious (and hence falsifiable) and grounded in both widely accepted perceptual theory (Gestalt; auditory scene analysis) and empirical evidence (Parncutt, 1983; Terhardt et al., 1982 a; Thompson & Parncutt, 1997). Given this background and the *psychophysical* nature of the theory, it is hardly surprising that it can account for differences between a profile based on the *physical* frequency of occurrence of tones and a profile based on the *perception* of corresponding pitches.

One of the most promising aspects of Temperley's probability-based approach is its potential to unite theories of tonality, rhythm and expectancy. The problem is essentially to predict the probability of *any* given continuation to *any* given polyphonic sequence (or pitch-time pattern in the chromatic scale). Temperley makes considerable progress toward this goal. His approach could be enriched and made even more general by incorporating psychoacoustic theories of pitch and pulse perception (Parncutt, 1988, 1989, 1994). In my model of pulse and meter (Parncutt, 1994), I attempted to predict the probability that a listener will perceive an event (i.e., its salience) in a periodically repeating pattern as the downbeat. Combining pitch and time, the general problem is to predict the probability of perceiving each pitch in the chromatic scale at each note onset in a performance. Such a model would unite music analysis with music perception. But it would not be easy to develop it, because it is empirically difficult to evaluate the perceptual salience of individual pitches that are embedded in a contrapuntal context.

Both music theorists and music psychologists tend to theorize as if listeners hear all the notes in the score and no others. Perhaps this assumption holds for some expert listeners when the style of the music is familiar to them. Some musicians can transcribe entire scores in real time in their minds as they listen (a feat that incidentally has not yet been the subject of systematic empirical study). But that is a strange and unusual way to enjoy music. It is evidently more usual to hear some tones and not others, and to hear some tones as central and others as subsidiary. The probability that a given tone will be heard, or heard as a point of reference, depends not only on its loudness in some absolute sense (i.e. independent of context) but also on the musical context in which it appears.

Terhardt's theory and accompanying empirical work (cited above) suggests furthermore that it is commonplace for listeners to experience pitches that do not correspond to any note in the score at all. In the following I will call these *implied pitches,* to distinguish them from virtual pitches; recall that the pitch of *any* complex tone is virtual, whether the fundamental is present or not. A well-known example of implied pitch perception is the diminished triad that functions as a dominant seventh with missing root. Listeners evidently perceive an implied pitch at the missing root, consistent with the chord's perception as a dominant.

The cited literature suggests that implied pitches are heard more often that we might like to admit. In the case of untrained listeners, the presence of implied pitches can explain why the recognition of musical chords in ear training courses can be so difficult – at least by comparison to the recognition of analogous visual patterns, that seldom include implied lines (except in those well-known visual illusions). In the case of musically trained listeners, Seither-Preisler et al. (in press) observed that they are more likely than untrained listeners to hear virtual pitches at missing fundamentals.

The final chapter is about *communicative pressure*. Temperley's point is that a composer or performer who wishes to communicate a given musical structure to a listener must do so within given surface constraints. A composer who wants listeners to perceive a number of independent melodic lines must conform to perceptually based conventions of counterpoint (Huron, 2006). A performer who wishes to add emotional expression to a performance while at the same time communicating the music's metrical structure can add more rubato to metrically simple than to metrically complex music. The more complex the harmonies used by improvisers in bebop jazz, the less freedom they have to present chords in inversion without obscuring chord roots and root progressions.

Temperley's book is an excellent example of interdisciplinary balance. He has taken two disciplines, music theory and music perception, whose object of study is very similar but whose epistemologies and methods are very different. He has then identified a range of issues that are interesting to representatives of both disciplines and developed a promising new method of addressing them. It is rare to find an approximately equal balance between epistemologically distant disciplines and I hope that this book will contribute to a more productive interaction between music psychology and music theory in the future.

Literature

Huron, D. (2006). *Sweet anticipation: Music and the psychology of expectation*. Cambridge, MA: MIT.

Kostka, S. (1995). *Instructor's manual to accompany Tonal harmony.* New York: McGraw-Hill.

Krumhansl, C. L., & Kessler, E. J. (1982). Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychological Review, 89*, 334-368.

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 psychoacoustic approach.* Berlin: Springer.

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

Parncutt, R. (1994). A perceptual model of pulse salience and metrical accent in musical rhythms. *Music Perception*, *11*, 409-464.

Parncutt, R. (2007). Tonal implications of harmonic and melodic Tn-sets. Spoken presentation at Mathematics and Computation in Music (MCM2007), Berlin, Germany, 18-20 May.

Seither-Preisler, A., Johnson, L., Krumbholz, K., Nobbe, A., Patterson, R.D., Seither, S., & Lütkenhöner, B. (in press). Observation: Tone sequences with conflicting fundamental pitch and timbre changes are heard differently by musicians and non-musicians. *Journal of Experimental Psychology: Human Perception and Performance.*

Thompson, W. F., & Parncutt, R. (1997). Perceptual judgments of triads and dyads: Assessment of a psychoacoustic model. *Music Perception, 14* (3), 263-280.

Terhardt, E. (1982). Die psychoakustischen Grundlagen der musikalischen Akkordgrundtöne und deren algorithmische Bestimmung. In C. Dahlhaus & M. Krause (Eds.), *Tiefenstruktur der Musik* (pp. 23-50). Berlin: Technical University of Berlin.

Terhardt, E., Stoll, G., & Seewann, M. (1982 a). Pitch of complex signals according to virtual-pitch theory: Tests, examples, and predictions. *Journal of the Acoustical Society of America, 71,* 671-678.

Terhardt, E., Stoll, G., & Seewann, M. (1982 b). Algorithm for extraction of pitch and pitch salience from complex tonal signals. *Journal of the Acoustical Society of America, 71,* 679-688.