BOOK REVIEW

The Cognition of Basic Musical Structures
By DAVID TEMPERLEY
Cambridge, MA: The MIT Press, 2001
7 × 9, 360 pp., 211 illus
ISBN 0-262-20134-8, $50.00/£32.95 (CLOTH)
ISBN 0-262-70105-7, $27.00/£17.95 (PAPER)
Reviewed by ROBERT ROWE
New York University

DAVID TEMPERLEY'S book The Cognition of Basic Musical Structures is a lucid and compelling essay on techniques for simulating the human comprehension of music with a computer program. As stated in the preface, the book addresses the question “How do we extract basic kinds of musical information—meter, phrase structure, counterpoint, pitch spelling, harmony, and key—from music as we hear it?” In this endeavor the author has produced a valuable contribution to music cognition and music theory. He presents a group of elegantly designed, working algorithms and examines them through tests using common practice music—tests that show the processes working quite effectively.

The question arises, in what way is this book about cognition? It is clearly an important contribution to music theory, in that it provides well-designed and effective computational approaches to identifying some of music theory’s most basic constructs. As the algorithms are primarily causal, analyzing and categorizing music as it is performed cognitively in the same way. However, if a model does not perform a process successfully, then one knows that the process is not performed cognitively in that way.” Perhaps the most important way in which the book is about cognition is in the way that it measures success: whether or not the constructs identified by the computer programs correspond to what human listeners hear. Further, it is about music cognition because of the ways in which its questions are framed and because of the literature within which the work is grounded. Here the author is clearly engaging and contributing to some of the central concerns in the field.

Temperley’s algorithms are presented as a set of preference rules. "Preference rules are criteria for forming some kind of analysis of input. Many possible interpretations are considered; each rule expresses an opinion as to how well it is satisfied by a given interpretation, and these opinions are combined together to yield the preferred analysis.” The proximate model for preference rule systems is of course Lerdahl and Jackendoff’s Generative Theory of Tonal Music (1983) (hereafter referred to as GTTM), and David Temperley worked closely with Fred Lerdahl in the development of the work reported here.

The signature characteristic of David Temperley’s volume is that it presents working implementations of preference rule systems. Producing working implementations requires a number of choices, among them the computational approach and representational systems. In collaboration with Daniel Sleator, the author settled on dynamic programming as the computational approach. In dynamic programming, partial solutions to the traversal of a search space are maintained that are extended incrementally by considering progressively longer paths through the space. At any point, the leading solution to the problem under consideration is that path which incurs the least cost in its traversal of the space. The dynamic programming approach is well-chosen for Temperley and Sleator’s work, in that it affords a method for considering a piece of music from left to right (the search space) while computing an optimal (or close-to-optimal) traversal of that space that satisfies a number of conditions, in this case the preference rules (thereby minimizing the cost).

In terms of representation, music is stored and manipulated in what is called a "piano-roll" format. This is basically the Musical Instrument Digital Interface (MIDI) representation, in which notes are stored as pitch numbers together with their onset and offset times. (The MIDI representation also encodes the dynamic, or loudness, of each onset, but this information is not used in the rule sets here.) Some of the rule sets require additional information, such as a metrical framework, to be supplied as well.

Each of the preference-rule-set chapters of the book (chapters 2-7) follows a common format. First, the author establishes the importance of an element of musical infrastructure. These elements include metrical structure, melodic phrase structure, contrapuntal structure, pitch spelling and tonal-pitch-class representation, harmonic structure, and key structure. Next, a context of prior experimental and theoretical research around these structural elements is established. Much of the
context discussed comes from the music cognition literature, and it is this context that forms the foundation for the ways in which this book is understood as being about cognition, rather than about music theory more narrowly. With a basic framework established, preference rules are introduced and illustrated with examples from the common practice repertoire. The discussion for successive rules often indicates ways in which they are insufficient, which motivates introduction of the next rule. A general methodology emerges, then, that consists of trying out preference rule sets on real-world examples and looking for how they break. It is a tribute to the author’s good musical sense that he comes up with sets of preferences that are relatively compact, yet work to explain passages that are beyond the scope of many other comparable systems. Finally, tests are conducted and evaluated that show how well the complete rule sets work, often with reference to earlier systems that were described in the research overview.

This way of presenting the material makes for a clear exposition and interesting discussion of the motivation behind each preference rule. The surprise is how engrossing it becomes to read. The description of computer implementations can be deadly; beat-by-beat analyses of musical passages are often no better. But we may rejoice that the prose here is in no way lethal: actually the motivation for and the thinking that went into each rule of a set reads more like a detective story. One becomes curious to know how ambiguous or contradictory phenomena in metric or harmonic structures can be adequately addressed using an approach like this. That this text both stimulates such curiosity and does a fine job of suggesting plausible and testable answers should make it of interest to a wide audience of readers about cognition, rather than about music theory more narrowly.

As an example, consider the Phrase Structure Preference Rules (PSPRs). These refer to melodic phrases. The first rule is “PSPR1 (Gap Rule): Prefer to locate phrase boundaries at (a) large inter-onset intervals and (b) large offset-to-offset intervals.” This much looks similar to a number of prior systems, including Lerdahl and Jackendoff (1983) and Tenney and Polansky (1980), both of which are discussed in the text. Looking at the application of this rule to musical examples, however, leads to the conclusion that it tends to produce too many phrase boundaries. Hence the second rule, derived from a statistical analysis of phrases in Ottman’s *Music for Sight Singing* (1986): “PSPR2 (Phrase Length Rule): Prefer phrases to have roughly 8 notes.” Finally, a third rule is adapted from observations concerning the impact of meter on phrase perception: “PSPR3 (Metrical Parallelism Rule). Prefer to begin successive groups at parallel points in the metrical structure.” The rule set was then tested on the Essen folksong collection. After correcting the second rule to match the average phrase length of the Essen database (preferring 10-note phrases rather than 8 notes), the program correctly identified 75.5% of phrases relative to the phrase notations in the music itself.

This account does much to establish the right questions and provide a convincing approach to proposing and evaluating answers. As the author notes, it also raises further questions that are as complex as the ones considered initially: for example, the third preference rule (PSPR3) requires a representation of the metric structure of the input to work. Could this metric structure come from the metrical structure preference rule set? Certainly it could, but doesn’t here, for the basic methodological reason that each rule set should be evaluated for its own effects. An intriguing area of further research would be to evaluate how the rule sets fare when they are used in combination, and whether there is anything to be gained from having them interact. Should metric preferences be keyed to phrase preferences, and vice versa? How could we weight and arbitrate between them?

A recent paper by Hamanaka, Hira, and Tojo (2005) demonstrates a technique for attaching weights to the metric preference rules in GTTM. Performance of an algorithmic analyzer was improved from a baseline 84% correct to 90% correct by hand-tuning the relative importance of preference rules in the set. A similar modification could certainly be adapted to Temperley’s rule sets and might help to illuminate ways to combine the rules for several parameters simultaneously. I appreciate David Temperley’s book both for the questions it raises and for its integrity in systematically searching for (and finding) well-documented answers. He has established an original and important line of inquiry and laid out interesting paths for its continuation. The results he has already achieved should encourage him and many others to engage in an exploration of those paths for a long time to come.

**Author Note**

Address correspondence to: Robert Rowe, NYU Dept. of Music and Performing Arts Professions, 35 W. Fourth St., Room 777, New York, NY 10012. E-MAIL: robert.rowe@nyu.edu
References


