Music Psychology and Music Theory: Problems and Prospects
Author(s): Carol L. Krumhansl
Source: *Music Theory Spectrum*, Vol. 17, No. 1 (Spring, 1995), pp. 53-80
Published by: University of California Press on behalf of the Society for Music Theory
Stable URL: http://www.jstor.org/stable/745764
Accessed: 03/04/2013 14:47

Music Psychology and Music Theory: Problems and Prospects

Carol L. Krumhansl

"In the present work an attempt will be made to connect the boundaries of two sciences, which, although drawn towards each other by many natural affinities, have hitherto remained practically distinct—I mean the boundaries of *physical and physiological acoustics* on the one side, and of *musical science and esthetics* on the other. The class of readers addressed will, consequently, have had very different cultivation, and will be affected by very different interests . . . The horizons of physics, philosophy, and art have of late been too widely separated, and, as a consequence, the language, the methods, and the aims of any one of these studies present a certain amount of difficulty for the student of any other of them; and possibly this is the principal cause why the problem here undertaken has not been long ago more thoroughly considered and advanced towards its solution."[1]

With these words, Hermann Helmholtz opens his 1863 treatise, *Die Lehre von den Tonempfindungen als Physiologische Grundlage für die Theorie der Musik*. Despite Helmholtz's laudable example, the problems confronting interdisciplinary studies of music appear to be undiminished today. This article presents research using the methods of contemporary experimental psychology that examines a recent music-theoretic proposal, Eugene Narmour's implication-realization model of melodic expectancy.[2] This research is used as a vehicle for explicating some of the underlying assumptions and goals of experimental psychology. The article also provides a brief summary of related empirical studies, and discusses the implications of the empirical findings for

An earlier version of this article was presented in the Special Session entitled "Theory and Evidence: An Interface with Cognitive Psychology" organized by Dr. Janet Hander-Powers at the Annual Meeting of the Society for Music Theory, Kansas City, 1992. A preliminary report of the experimental results is contained in Carol L. Krumhansl, "Melodic Structure: Theoretical and Empirical Descriptions," in *Music, Language, Speech and Brain*, ed. J. Sundberg, L. Nord, and R. Carlson (London: MacMillan, 1991), 269–83. This article was completed while the author was a Fellow at the Center for Advanced Study in the Behavioral Sciences with support from the National Science Foundation (SES-9022192) and the James McKeen Cattell Fund. I am grateful to Robert O. Gjerdingen, Kathleen Much, Fred Lerdahl, and Leonard Meyer for comments on this article, and to Eugene Narmour for generous consultation at every stage of this research project.

[1]Hermann L. F. von Helmholtz, *On the Sensations of Tone as a Physiological Basis for the Theory of Music*, ed. and trans. A. J. Ellis (New York: Dover, 1954), 1; originally published as *Die Lehre von den Tonempfindungen als Physiologische Grundlage für die Theorie der Musik* (Braunschweig: F. Vieweg und Sohn, 1863).

[2]Eugene Narmour, *The Analysis and Cognition of Basic Melodic Structures: The Implication-Realization Model* (Chicago: University of Chicago Press, 1990).

psychological theory. In addition, the research illustrates one approach to operationalizing experimentally a music-theoretic proposal, and highlights those aspects of this proposal that makes it suited to such analysis.

MUSIC PSYCHOLOGY AND MUSIC THEORY

Helmholtz envisioned a science of music that consists of three interrelated types of investigation: musical acoustics, auditory physiology and perception, and music theory. The first section of the monograph describes the acoustics of complex tones and the perception of timbre or tone color. The second section concerns phenomena produced by multiple tones, such as beats, combination tones, and dissonance. The third and final section treats the formation of musical scales and harmonies. In this third section, Helmholtz argues that these elementary aspects of musical organization are the joint product of artistic choice and the sensory function of the auditory system. However, he found the connections between the physical, physiological, and perceptual observations, on the one hand, and music theory, on the other, so compelling as to conclude that such scientific results can serve as an explanation for the kinds of structures identified by music theorists.

Despite great advances in auditory psychology (particularly physiology and psychoacoustics), psychological research until recently intersected with music theory almost exclusively on a single topic, consonance. Even on that topic, the two disciplines maintained a rather distant relationship with one another, with minimal cross-referencing between the literatures. Psychological treatments tended to examine the perceived consonance of isolated intervals, whereas music-theoretic treatments emphasized influences of style and context. Then, in the mid-1950s, two books appeared showing that a broader and deeper exchange between psychology and music theory was possible. Leonard Meyer's *Emotion and Meaning in Music* drew on diverse psychological literatures, including Gestalt psychology, motivation and emotion, learning, and information theory.[3] Robert Francès's *La Perception de la musique* studied the perception of melody, harmony, tonality, and atonality, using a wide range of methods, musical materials, and levels of training.[4]

The idea that music theory could offer insights into musical behaviors began to take hold and influence the nature of experimentation in the 1970s. The logic of the approach was that theoretical descriptions of musical patterns might offer suggestions about listeners' knowledge of music (sometimes referred to as schemas or mental representations). This knowledge presumably affects how music is encoded in perception, interpreted, remembered, and performed. The shift toward a cognitive orientation that emphasizes learning and memory, seen perhaps most clearly in W. Jay Dowling's, Lola Cuddy's, and my research, is consistent with music theory's emphasis on culturally determined aspects of music.[5] Music-theoretic concepts played various roles in the research: to

[3]Leonard B. Meyer, *Emotion and Meaning in Music* (Chicago: University of Chicago Press, 1956).

[4]Robert Francès, *The Perception of Music*, trans. W. Jay Dowling (Hillsdale, N.J.: Erlbaum, 1988); originally published as *La Perception de la musique* (Paris: J. Vrin, 1958).

[5]W. Jay Dowling, "Recognition of Melodic Transformations: Inversion, Retrograde, and Retrograde Inversion," *Perception & Psychophysics* 12 (1972): 417–21; idem, "Scale and Contour: Two Components of a Theory of Memory for Melodies," *Psychological Review* 85 (1978): 341–54; Lola L. Cuddy and Annabel J. Cohen, "Recognition of Transposed Melodic Sequences," *Quarterly Journal of Experimental Psychology* 28 (1976): 255–70; Lola L. Cuddy, Annabel J. Cohen, and Janet Miller, "Melody Recognition: The Experimental Application of Musical Rules," *Canadian Journal of Psychology* 33 (1979): 148–57; Carol L. Krumhansl and Roger N. Shepard, "Quantification of the Hierarchy of Tonal Functions within a Diatonic Context," *Journal of Experimental Psychology: Human Perception and Performance* 5 (1979): 579–94; Carol L. Krumhansl, "The Psychological Representation of Musical Pitch in a Tonal Context," *Cognitive Psychology* 11 (1979): 346–74.

suggest musical variables of potential interest, to guide the construction of stimulus materials, and to interpret the results obtained.

The psychological literature now extensively documents the importance of scale, harmony, key, meter, and rhythm for understanding the experience of tonal-harmonic music.[6] By and large, the theoretical concepts invoked by the experimental studies are elementary, widely accepted, and unchallenged by the empirical results. Thus it might be argued that music theory has served primarily as a source of external validation for the experimental methods. This should not diminish the fact that the experiments have successfully demonstrated the psychological reality of certain music-theoretic concepts, however. Listeners in the experiments typically do not have training in music theory, and yet the experimental results show that their implicit knowledge about music contains some similar concepts. Although the emphasis has been on tonal-harmonic music, theoretical proposals about twentieth-century and non-Western musics have been examined in a few studies.[7] The recent literature also contains some experimental tests of more speculative and technical proposals, such as Fred Lerdahl and Ray Jackendoff's *Generative Theory of Tonal Music,*[8] Leonard Meyer's melodic archetypes,[9] Pieter Van den Toorn's partitionings of the

[6]For reviews, see John Sloboda, *The Musical Mind: The Cognitive Psychology of Music* (Oxford: Oxford University Press, 1985); W. Jay Dowling and Dane L. Harwood, *Music Cognition* (Orlando: Academic Press, 1986); Stephen Handel, *Listening: An Introduction to the Perception of Auditory Events* (Cambridge, Mass.: MIT Press, 1989); Carol L. Krumhansl, *Cognitive Foundations of Musical Pitch* (New York: Oxford University Press, 1990); Carol L. Krumhansl, "Music Psychology: Tonal Structures in Perception and Memory," *Annual Review of Psychology* 42 (1991): 277–303; S. McAdams and E. Bigand, eds., *Thinking in Sound: The Cognitive Psychology of Human Audition* (Oxford: Oxford University Press, 1993); Rita Aiello and John A. Sloboda, eds., *Musical Perceptions* (New York: Oxford University Press, 1994).

[7]See Francès, *The Perception of Music*; Dowling, "Recognition of Melodic Transformations"; Carol L. Krumhansl, Gregory S. Sandell, and Desmond C. Sergeant, "The Perception of Tone Hierarchies and Mirror Forms in Twelve-Tone Serial Music," *Music Perception* 5 (1987): 31–78; Carol L. Krumhansl "Memory for Musical Surface," *Memory & Cognition* 19 (1991): 401–11; Mary A. Castellano, Jamshed J. Bharucha, and Carol L. Krumhansl, "Tonal Hierarchies in the Music of North India," *Journal of Experimental Psychology: General* 113 (1984): 394–412; Edward J. Kessler, Christa Hansen, and Roger N. Shepard, "Tonal Schemata in the Perception of Music in Bali and the West," *Music Perception* 2 (1984): 131–65.

[8]Fred Lerdahl and Ray Jackendoff, *A Generative Theory of Tonal Music* (Cambridge, Mass.: MIT Press, 1983). Studies of the grouping component are Irène Deliège, "Grouping Conditions in Listening to Music: An Approach to Lerdahl & Jackendoff's Grouping Preference Rules," *Music Perception* 4 (1987): 325–60; Eric F. Clarke and Carol L. Krumhansl, "Perceiving Musical Time," *Music Perception* 7 (1990): 213–52. Studies of the metrical component are Caroline Palmer and Carol L. Krumhansl, "Independent Temporal and Pitch Structures in Determination of Musical Phrases," *Journal of Experimental Psychology: Human Perception & Performance* 13 (1987): 116–26; Caroline Palmer and Carol L. Krumhansl, "Pitch and Temporal Contributions to Musical Phrase Perception: Effects of Harmony, Performance Timing, and Familiarity," *Perception & Psychophysics* 41 (1987): 505–18; Caroline Palmer and Carol L. Krumhansl, "Mental Representations for Musical Meter," *Journal of Experimental Psychology: Human Perception and Performance* 16 (1990): 728–41. Studies of the time-span reduction component are Palmer and Krumhansl, "Independent Temporal and Pitch Structures" and "Pitch and Temporal Contributions." A study of the prolongation reduction component is Emmanuel Bigand, "Abstraction of Two Forms of Underlying Structure in a Tonal Melody," *Psychology of Music* 18 (1990): 45–59.

[9]Leonard B. Meyer, *Explaining Music* (Berkeley: University of California Press, 1973). Studies of the melodic archetypes are Burton S. Rosner and Leonard B. Meyer, "Melodic Processes and the Perception of Music," in *The Psychology of Music*, ed. Diana Deutsch (New York: Academic Press, 1982), 317–41; Burton S. Rosner and Leonard B. Meyer, "The Perceptual Roles of Melodic Process, Contour, and Form," *Music Perception* 4 (1986): 1–39; Mark A. Schmuckler, "Expectation in Music: Investigation of Melodic and Harmonic Processes," *Music Perception* 7 (1990): 122–47.

octatonic scale,[10] and Eugene Narmour's implication-realization model.[11]

METHODOLOGY IN EXPERIMENTAL PSYCHOLOGY

Some general remarks about experimental methodology may provide useful background to the experiments that are reported below. The focus is on those aspects of methodology that seem to distinguish psychology most sharply from music theory. The topics covered within scientific psychology range from emotion, personality, and social interactions to memory, perception, and neuropsychology. The methods used in any particular study depend on the topic of interest, but a common core of methodological norms has emerged. As a basic science, psychology is similar to the physical sciences in its concern with experimental control. In many cases the stimulus materials are constructed to vary only certain properties of interest. Sometimes more musically representative, "ecologically valid" materials (such as excerpts from compositions) are used. In the ideal case, both types of materials are employed and the results converge on the same conclusions.

Most experimental observations take a quantitative form, or can be translated into one.[12] Verbal reports are used less frequently because they are difficult to summarize and often add another layer of interpretation to understanding the results. Psychology seeks to identify general principles underlying complex behaviors. Differences between individuals may exist, but these are treated by systematic methods and understood with respect to relevant psychological theories of individual differences. Francès describes the approach taken in experimental studies of music as follows, "Experimental research operates in a region between the infinite variety of individual modes of thinking and feeling and the abstractions of (in principle) universal aesthetic experience. In that situation research turns more and more toward definitions that limit the diversity to types and stages that can be defined by objective criteria."[13]

Psychological methodology is rooted in the technical areas of experimental design and statistics. Experimental design considerations guide the choice of stimulus materials, participants, and experimental tasks. These choices affect how the data are treated statistically. Any set of data is subjected to a series of tests that evaluate whether a result is "statistically significant" or whether it might have occurred by chance. The underlying statistical theory specifies the range to which a result can be generalized. Once a result has passed the test of statistical significance, its importance for current conceptions about the psychological topic is considered. Judgments here are more subjective.

The relationship between theory and experiment is one in which the theory predicts that a certain effect should occur under particular conditions. The experiment then seeks to create these conditions. If the prediction is disconfirmed, then the theory is rejected (assuming that the experiment is well designed and that subsequent experiments replicate the find-

[10]Pieter C. Van den Toorn, *The Music of Igor Stravinsky* (New Haven: Yale University Press, 1983); Carol L. Krumhansl and Mark A. Schmuckler, "The *Petroushka* Chord: A Perceptual Investigation," *Music Perception* 4 (1986): 153–84.

[11]Narmour, *The Analysis and Cognition of Basic Melodic Structures*; experiments reported in this article; Krumhansl, "Melodic Structure: Theoretical and Empirical Descriptions"; and Carol L. Krumhansl and Eugene Narmour, "Melodic Expectancy: Effects of High-Level Implications and Similarity of Form," manuscript under review.

[12]Some of the technical issues involved in psychological measurement are reviewed in R. Duncan Luce and Carol L. Krumhansl, "Measurement, Scaling, and Psychophysics," in *Stevens' Handbook of Experimental Psychology*, 2nd edition, ed. R. C. Atkinson, R. J. Herrnstein, G. Lindzey, and R. D. Luce (New York: John Wiley & Sons, 1988), 3–74.

[13]Francès, *The Perception of Music*, 3.

ing). If confirmed, then the theory is supported—but not proven, because alternative theories might make the same prediction. This process of hypothesis testing describes research in an area that is sufficiently developed to generate specific predictions. Sometimes, however, the area is less well developed and the experiments are more observational or exploratory in nature. In either case, the experimental data are used to suggest how the psychological theory might be developed, extended, or refined; occasionally the results suggest that an established theory may need to be replaced. As an example of how a proposal originating in music theory might be tested experimentally, I report a series of studies testing predictions of Narmour's implication-realization model for tone-to-tone melodic expectancies.[14]

STUDIES OF MUSICAL EXPECTANCY

Expectancy plays an important function in a wide variety of behaviors, including perception, speech understanding and production, and skilled performance. Burton Rosner describes the advantages of relating music-theoretic treatments to empirically studied topics within psychology: "Psychology can subsume various aspects of music theory into broader categories. Insofar as psychology provides systematic treatments of those categories, the music theorist can avail himself of a wider analytic framework than he might otherwise possess."[15] A thorough review of the literature on expectancy is beyond the scope of this paper, but it would be well to summarize some of the studies on music. Two main objectives of this research can be identified. First, studies of musical expectancy uncover listeners' knowledge about musical patterns and the psychological processes they use to encode, organize, and remember music. Second, the studies offer insights into the dynamic processing of information over time, with continuously changing expectancies for subsequent events which may have implications for emotional responses.

A wide variety of methods have been developed to study musical expectancy, including production, memory, detection, priming, and structural judgments. This diversity lends richness to the results obtained, but it also reflects the difficulties of studying an ongoing psychological process. In various ways, the methods necessarily interrupt the musical experience, so generalizations to more typical listening conditions must be made with caution. The stimulus materials used in the studies range from isolated intervals to extended passages, and this brief review roughly follows this proression.

James Carlsen and collaborators presented subjects with two successive tones (ranging from a descending octave to an ascending octave) and instructed them to sing what they believed would be the continuation of the melody had it not been interrupted.[16] These vocal productions were transcribed, and the tones produced immediately after the two-tone contexts were compiled.[17] A number of findings are

[14]Narmour, *Analysis and Cognition of Basic Melodic Structures*.

[15]Burton S. Rosner, "Music Perception, Music Theory, and Psychology," in *Explorations in Music, the Arts, and Ideas*, ed. E. Narmour and R. A. Solie (Stuyvesant, N.Y.: Pendragon Press, 1988), 150.

[16]James C. Carlsen, Pierre L. Divenyi, and Jack A. Taylor, "A Preliminary Study of Perceptual Expectancy in Melodic Configurations," *Council for Research in Music Education Bulletin* 22 (1970): 4–12; James C. Carlsen, "Some Factors which Influence Melodic Expectancy," *Psychomusicology* 1 (1981): 12–29; Anna M. Unyk and James C. Carlsen, "The Influence of Expectancy on Melodic Perception," *Psychomusicology* 7 (1987): 3–23.

[17]Mark A. Schmuckler, "The Performance of Global Expectations," *Psychomusicology* 9 (1990): 122–47, notes a number of limitations of this method. First, the task requires a certain level of musical skill so that it is limited to participants who are musically trained. Second, producing the continuations requires conscious attention, although expectancies may well be largely unconscious under normal listening conditions. Finally, the two-tone contexts may be insufficient to generate strong or consistent expectancies.

interesting in light of certain aspects of the implication-realization model described below. First, each stimulus interval resulted in multiple responses, suggesting that expectancies select a set of tones rather than a single tone. Second, the tones produced tended to be proximate to the second of the two context tones. This was true especially for ascending and small descending intervals; large descending intervals resulted in a number of relatively large response intervals in the opposite direction. Third, response tones frequently returned to the starting pitch of the context interval or to the tone that would complete the octave.

Dowling took a very different approach to musical expectancy, using two interleaved melodies in which successive tones alternated back and forth between the two melodies.[18] Listeners found it extremely difficult to identify the two different melodies when played in the same register, but could detect the presence or absence of a target melody when it was specified in advance. This difference can be accounted for by the idea that expectancy operates by selecting only those elements that match the cued target melody. More generally, this research has led to the proposed existence of "expectancy windows" that aim perceptual processing to particular pitch regions at particular times, an idea that is similar in spirit to the notion of "dynamic attending" developed in the work of Mari Reiss Jones and Marilyn Boltz.[19]

Expectancies derived from knowledge of tonal-harmonic style have been documented in numerous experiments, only some of which will be mentioned here. Listeners remember tones and chords that are expected in a given context better than those that are unexpected. Stylistically unexpected tones and chords are poorly remembered. In addition, changes from an unexpected element to an expected element are difficult to detect, whereas changes from expected to unexpected elements are easily detected.[20] Caroline Palmer has found similar patterns in studies of errors made in musical performances.[21] Lucinda DeWitt and Arthur Samuel demonstrated effects of style knowledge in another task.[22] The task required listeners to distinguish between two types of items: those in which a target tone was excised from the recording and replaced with noise (a sound with broad spectral content), and those in which noise was merely added to the signal. Listeners were more accurate when the melody was either familiar or predictable, a result consistent with the proposals of Dowling and of Jones and Boltz that expectancy can aid processing by directing attentional resources.

[18]W. Jay Dowling, "The Perception of Interleaved Melodies," *Cognitive Psychology* 5 (1973): 322–37; W. Jay Dowling, Kitty Mei-Tak Lung, and Susan Herrbold, "Aiming Attention in Pitch and Time in the Perception of Interleaved Melodies," *Perception & Psychophysics* 41 (1987): 642–56; see W. Jay Dowling, "Expectancy and Attention in Melody Perception," *Psychomusicology* 9 (1990): 148–60, for a recent review.

[19]Mari R. Jones and Marilyn Boltz, "Dynamic Attending and Responses to Time," *Psychological Review* 96 (1989): 459–91.

[20]Cuddy, Cohen, and Miller, "Melody Recognition: The Experimental Application of Rules"; Krumhansl, "The Psychological Representation of Musical Pitch in a Tonal Context"; Jamshed J. Bharucha and Carol L. Krumhansl, "The Representation of Harmonic Structure in Music: Hierarchies of Stability as a Function of Context," *Cognition* 13 (1983): 63–102; Carol L. Krumhansl, Jamshed J. Bharucha, and Mary A. Castellano, "Key Distance Effects on Perceived Harmonic Structure in Music," *Perception & Psychophysics* 32 (1982): 96–108; for a review, see Krumhansl, *Cognitive Foundations of Musical Pitch*.

[21]Caroline Palmer, "The Role of Interpretive Preferences in Music Performance," in *Cognitive Bases of Musical Communication*, ed. M. R. Jones and S. Holleran (Washington, D.C.: American Psychological Association, 1992), 249–62; Caroline Palmer and Carla van de Sande, "Units of Knowledge in Music Performance," *Journal of Experimental Psychology: Learning, Memory, & Cognition* 19 (1993): 457–70.

[22]Lucinda A. DeWitt and Arthur G. Samuel, "The Role of Knowledge-based Expectations in Music Perception: Evidence from Musical Restoration," *Journal of Experimental Psychology: General* 119 (1990): 123–44.

The most temporally precise measures of expectancy come from the studies of Jamshed Bharucha and collaborators.[23] They used a priming paradigm in which a *target* chord was presented after another chord, called the *prime*. The task was to judge whether the target chord was or was not in tune, which is an indirect measure of the perceived relatedness between the prime and target chords. The time listeners took to judge whether the target chord was in tune was shorter when the prime and target chords belonged to the same diatonic collection than when they did not. This and other research has led to the development of models of musical expectancy that are implemented as neural networks.[24] These models contain "nodes" representing tones, chords, and keys. Activation spreads through the network in a way that simulates how learning alters neural connections. These models promise insights into understanding how musical patterns might be learned and how they govern expectancies. The models developed by Bharucha make the useful distinction between *schematic expectancies*, abstracted from large numbers of sequences, and *veridical expectancies*, instance-based expectancies for what will occur next in a particular sequence.

Mark Schmuckler's studies of musical expectancy used more extended musical excerpts, taken from an accompanied Schumann song.[25] Ten stopping points (*probe positions*) were identified at which interesting melodic or harmonic progressions occur. The excerpts were interrupted at these probe positions and then followed by different possible continuations. Listeners rated each continuation on a numerical scale for how well it fit with their expectancies. These judgments were influenced by knowledge of the tonal-harmonic style (tonal hierarchies and harmonic progressions). In addition, the data supported the idea that listeners base their responses in part on conformance to linear patterns (conjunct motion) and gap-fill patterns (disjunct motion) as proposed by Leonard Meyer.[26] Schmuckler found similar patterns in a performance study, in which participants improvised continuations on a keyboard.[27]

The majority of these studies have considered effects that learned musical patterns have on expectancy, focusing on such aspects as tonality, harmony, and melodic familiarity. This is consistent with the cognitive orientation of much of the recent research in the psychology of music. However, general principles of perceptual organization that do not require extensive learning, such as those studied by Albert S. Bregman and Diana Deutsch, may also influence melodic

[23]Jamshed J. Bharucha and Keiko Stoeckig, "Reaction Time and Musical Expectancy: Priming of Chords," *Journal of Experimental Psychology: Human Perception and Performance* 12 (1986): 403–10; Jamshed J. Bharucha and Keiko Stoeckig, "Priming of Chords: Spreading Activation or Overlapping Frequency Spectra?" *Perception & Psychophysics* 41 (1987): 519–24; Hasan Gurkan Tekman and Jamshed J. Bharucha, "Time Course of Chord Priming," *Perception & Psychophysics* 51 (1992): 33–39.

[24]Jamshed J. Bharucha, "Music Cognition and Perceptual Facilitation: A Connectionist Framework," *Music Perception* 5 (1987): 1–30; Jamshed J. Bharucha and Katherine L. Olney, "Tonal Cognition, Artificial Intelligence and Neural Nets," *Contemporary Music Review* 4 (1989): 341–56; Jamshed J. Bharucha, "Pitch, Harmony, and Neural Nets: A Psychological Perspective," in *Music and Connectionism*, ed. P. M. Todd and D. G. Loy (Cambridge, Mass.: MIT Press, 1991); Jamshed J. Bharucha and Peter M. Todd, "Modeling the Perception of Tonal Structure with Neural Nets," *Computer Music Journal* 13 (1989): 44–53. In related work, Robert O. Gjerdingen, "Categorization of Musical Patterns by Self-Organizing Neuronlike Networks," *Music Perception* 7 (1990): 339–70, and "Apparent Motion in Music," *Music Perception* 11 (1994): 335–70, has demonstrated that neural nets are able to abstract significant voice-leading combinations and melodic motion, both of which may influence expectancy.

[25]Schmuckler, "Expectation in Music: Investigation of Melodic and Harmonic Processes" and "The Performance of Global Expectations."

[26]Meyer, *Explaining Music*.

[27]Schmuckler, "The Performance of Global Expectations."

expectancy.[28] Narmour develops this idea in the implication-realization model in which central roles are played by the Gestalt principles of similarity, proximity, and good continuation. These principles, he claims, are innate, automatic, and largely unavailable to conscious introspection. If such a theory of melodic expectancy is supported, it would provide a useful complement to the existing psychological literature. Let us turn now to see how the implication-realization model was formulated as an experimentally testable proposal.

THE IMPLICATION-REALIZATION MODEL

The implication-realization model invites experimentation for a number of reasons.[29] First, Narmour explicitly claims that the model describes the listener's cognitive response to music, and he offers it as an object for experimental test. His contribution is primarily one of music analysis, rather than of psychological formulation, but implicit in the work is a set of testable hypotheses. Second, his hypotheses concern a phenomenon (tone-to-tone expectancies) that occurs over a short span of time. Experiments are generally easier to implement if shorter stimulus sequences are required, although some experiments have used extended excerpts or entire musical pieces.[30] Third, the model treats musical parameters separately. Melodic expectancy is described in terms of parametric scales of pitch independent of other musical parameters. Although rhythm and duration are important for determining *when* strong expectancies occur, the principles proposed to describe *what* tones are expected depend only on interval size and direction.[31] Fourth, the parametric scales can be quantified as shown below and, thus, compared with numerical data from the experiments. Fifth, the model's principles are presumed to be general, perhaps even universal and innate. Although this claim is untestable in principle, it does suggest that only minimal differences should be found with stimulus materials from different styles and with listeners varying in musical training and experience.

Only those aspects of the implication-realization model that are relevant to the experiments will be summarized here. The experiments test the model's proposal that principles of perceptual organization influence melodic expectancies. These principles do not depend on pre-existing knowledge of the piece (*intra-opus knowledge*) or knowledge of the style (*extra-opus knowledge*). Rather, the principles depend only on the perceptually immediate musical context. Narmour adopts the distinction made in psychology between *top-down* processes, which interpret incoming perceptual information by relating it to knowledge acquired through previous experience, and *bottom-up* processes, which do not involve such knowledge. The experiments reported here focus exclusively on the bottom-up component, although the model also proposes effects of intra- and extra-opus knowledge. In addition, the experiments focus exclusively on hypothesized expect-

[28]Comprehensive summaries of this research can be found in Albert S. Bregman, *Auditory Scene Analysis* (Cambridge, Mass.: MIT Press, 1991) and Diana Deutsch, "Auditory Pattern Recognition," in *Handbook of Perception and Human Performance*, Vol. 2, *Cognitive Processes and Performance*, ed. K. R. Boff, L. Kaufmann, and J. P. Thomas (New York: Wiley, 1986), Chapter 32.

[29]Narmour, *Analysis and Cognition of Basic Melodic Structures*.

[30]See, for example, the studies by Lucy Pollard-Gott, "Emergence of Thematic Concepts in Repeated Listening to Music," *Cognitive Psychology* 15 (1981): 66–94; Irène Deliège, "Recognition of Musical Forms through Listening," *Contemporary Music Review* 14 (1989): 325–60; Krumhansl, "Memory for Musical Surface," Clarke and Krumhansl, "Perceiving Musical Time."

[31]The practice of controlling all but one parameter of interest follows from experimental design considerations. This has tended to generate studies of pitch separate from studies of rhythm; see Krumhansl, "Music Psychology: Tonal Structures in Perception and Memory," for a summary of some of the issues and relevant experiments.

ancy effects on the tone-to-tone level, although the model also hypothesizes expectancy effects on higher levels between noncontiguous tones.[32]

According to the model, the cognition of melodies can be described as successive points of *closure*, *implication*, and *realization*. Closure and implication have opposite effects on expectancy for melodic continuation. When closure occurs, expectancy for melodic continuation is weak. When non-closure (or implication) occurs, expectancy for melodic continuation is strong. According to the model, six conditions lead to a sense of closure: 1) a rest, 2) a strong metrical position, 3) dissonance resolving to consonance, 4) a short tone followed by a long tone, 5) a large interval followed by a smaller interval, and 6) a change in registral direction (up-to-down, up-to-lateral, down-to-up, down-to-lateral, lateral-to-up, or lateral-to-down).[33] The strength of closure increases with the number of conditions that are present in the music. When none of these conditions holds, the pattern is unclosed and a point of implication is established that generates expectancies for how the melody will continue. The last interval that appears at a point of implication is called an *implicative* interval. The interval that follows is called a *realized* interval; it is formed by the second tone of the implicative interval and the following tone. Five principles, to be discussed shortly, underlie the bottom-up component of the model. These principles determine classes of tones that are expected to follow any given implicative interval, and they depend on the size and direction of the implicative interval.

The central assumption of the model is that a small implicative interval implies a *process*, which means that the realized interval will be in the same direction as the implicative interval and it will be similar in size, whereas a large implicative interval implies a *reversal*, which means that the realized interval will be in a different direction from the implicative interval and it will be smaller in size. Implicative intervals of a perfect fourth or smaller are considered *small*; implicative intervals of a perfect fifth or larger are considered *large*. The tritone (augmented fourth/diminished fifth) is a threshold where the implicative strength of process and reversal are equal; in tonal music this ambiguity is usually resolved by dissonance and harmonic rhythm.[34] As will be shown below, the model precisely defines other terms (such as similar-sized, smaller in size, larger in size, proximate, and non-proximate) on the parametric scale of interval size.

I introduce the grid representation shown in Figure 1 as a visual aid for summarizing the five principles underlying the model. (The principles are also summarized in Table 1.) In the grid, the vertical dimension represents the size of the implicative interval measured in semitones (0 semitones = unison, 1 semitone = minor second, 2 semitones = major second, etc.). As can be seen, the vertical dimension divides into small implicative intervals (up to 5 semitones) and large implicative intervals (7 semitones or larger). The horizontal dimension represents the size of the realized interval (again measured in semitones) and its direction (same or different) relative to the direction of the implicative interval.[35] Any square in the grid represents a combination of implicative and realized intervals. For example, the sequence of tones C_4–E_4–F_4 (where 4 denotes the octave beginning with middle C) is represented by the cell with an implicative interval equal to 4 semitones (ascending major third) and realized interval equal to 1 semitone (ascending minor second) going in the

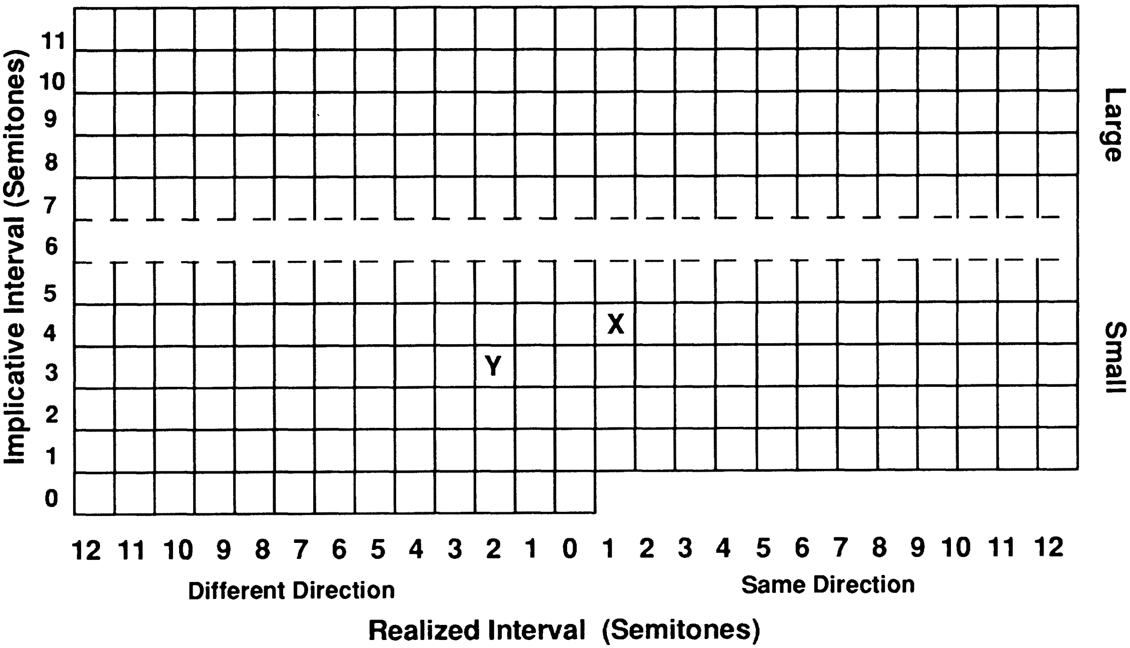
[32]An experimental test of expectancy effects on higher levels is reported by Carol L. Krumhansl and Eugene Narmour, "Melodic Expectancy: Effects of High-Level Implications and Similarity of Form."

[33]Narmour, *Analysis and Cognition of Basic Melodic Structures*, 11–12.

[34]Ibid., 79.

[35]In principle, the grids could be extended indefinitely in both horizontal and vertical directions, but most intervals in tonal music fall within the range shown.

Figure 1. The grid representation used to describe the principles underlying the bottom-up component of the implication-realization model. The vertical axis represents the size of the implicative interval in semitones. The horizontal axis represents the size of the realized interval in semitones and its direction relative to the implicative interval.



Examples:

$X = C_4 - E_4 - F_4$ (4 semitones, 1 semitone same direction)
$Y = C_4 - A_3 - B_3$ (3 semitones, 2 semitones different direction)

same direction (up–up). The "X" in Figure 1 marks this combination of implicative and realized intervals. To give another example, the sequence of tones C_4–A_3–B_3 is represented by the cell with an implicative interval equal to 3 semitones (descending minor third) and realized interval equal to 2 semitones (ascending major second) going in a different direction (down–up); the "Y" marks this combination.

Notice that both interval size and interval direction are used in the definitions of process (small interval followed by a similar-sized interval in the same direction) and reversal

Table 1. Five Principles of Melodic Expectancy Underlying the Bottom-up Component of the Implication-Realization Model

Registral Direction

Small implicative intervals imply realized intervals in the same direction.

Large implicative intervals imply realized intervals in a different direction.

Intervallic Difference

Small implicative intervals imply realized intervals that are similar-sized.[1]

Large implicative intervals imply realized intervals that are smaller in size.[2]

Registral Return

The interval formed by the first tone of the implicative interval and the second tone of the realized interval is no greater than a major second.

Proximity

Independent of the size and direction of the implicative interval, implied realized intervals are no larger than a perfect fourth.

Closure

Closure is strongest when 1) the implicative interval is large and the realized interval is smaller[2] and 2) registral direction of implicative and realized intervals are different.

[1] Within a minor third if registral direction of implicative and realized intervals is the same; within a major second if registral direction is different.

[2] Smaller by more than a minor third if registral direction is the same; smaller by more than a major second if registral direction is different.

(large interval followed by a smaller interval in a different direction). Thus, these melodic structures are the joint product of two independent underlying principles, one that concerns the relative directions of the implicative and realized intervals, and one that concerns the relative sizes of the implicative and realized intervals. The first of these, the principle of *registral direction*, states that the perfect fourth is the upper bound for registral continuation, and the perfect fifth is the lower bound for registral reversal. In other words, if the implicative interval is a small interval (perfect fourth or smaller), the direction is expected to continue (up–up, down–down, or lateral–lateral); if the implicative interval is a perfect fifth or larger, the direction is expected to be different (up–down, down–up, up–lateral, or down–lateral).

The shaded cells in Figure 2 indicate the combinations of implicative and realized intervals that satisfy the principle of registral direction (see also Table 1). For any given implicative interval on the vertical axis, the shaded cells going across the grid indicate the realized intervals that satisfy the principle. For example, the sequence C_4–E_4–F_4 (4 semitones, 1 semitone same direction) falls in one of the shaded cells in the lower right quadrant of the grid and, thus, satisfies the principle of registral direction. On the other hand, the sequence C_4–A_3–B_3 (3 semitones, 2 semitones different direction) falls in an unshaded cell in the lower left quadrant, and thus does not satisfy the principle of registral direction. The sequence C_4–A_4–G_4 (9 semitones, 2 semitones different direction) falls in a shaded cell in the upper left hand quadrant, and thus does satisfy the principle of registral direction, and so on. As specified in the implication-realization model, the principle of registral direction is treated here as *all-or-none*: for any given implicative interval, a realized interval either does or does not satisfy the principle. Graded variants of the principle, representing different degrees of implicative strength, can be formulated and tested.

Figure 2: The principle of registral direction. Small implicative intervals imply the direction will continue; large implicative intervals imply the direction will be different.

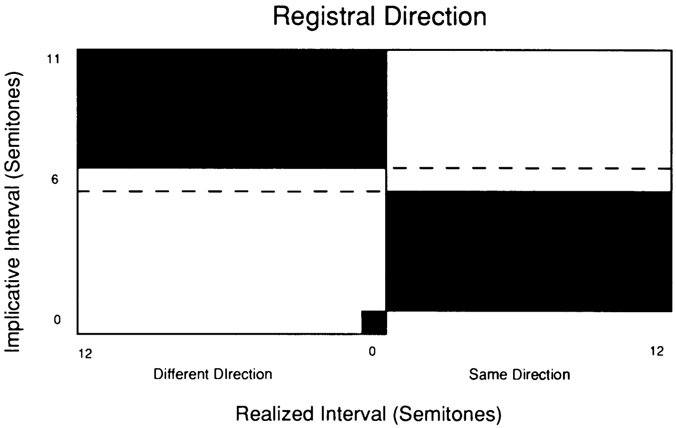
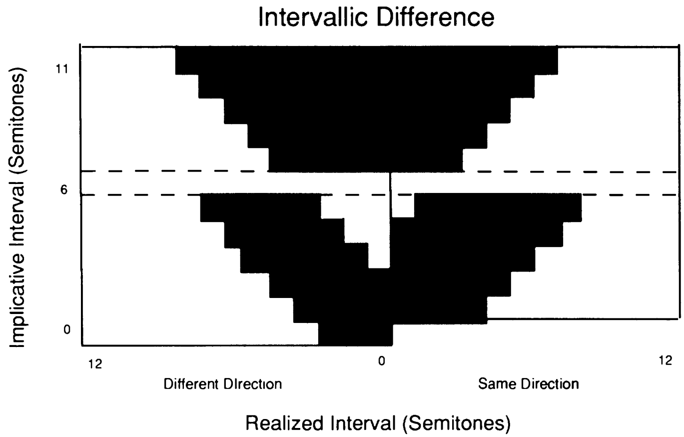


Figure 3: The principle of intervallic difference. Small implicative intervals imply similar-sized intervals; large implicative intervals imply smaller intervals.



The second principle concerns the interval-size aspect of the definitions of process and reversal. This principle of *intervallic difference* similarly distinguishes between implications for small and large implicative intervals. For small implicative intervals, the principle states that small implicative intervals imply similar-sized realized intervals. The definition of *similar-sized* depends on whether registral direction is the same or different. In the former case, similar-sized means the same size plus or minus a minor third. In the latter case, similar-sized means the same size plus or minus a major second. Figure 3 shows as the shaded cells on the bottom half for small implicative intervals those combinations of implicative and realized intervals that satisfy the principle of intervallic difference. The shaded region is slightly asymmetric because the definition of similar-sized depends on whether registral direction is same or different. The sequences C_4–E_4–F_4 (4 semitones, 1 semitone same direction) and C_4–A_3–B_3 (3 semitones, 2 semitones different direction) fall in shaded cells, but not the sequence C_4–D_4–A_4 (2 semitones, 7 semitones same direction).

For large implicative intervals, the principle of intervallic difference states that large implicative intervals imply realized intervals that are smaller in size. The definition of *smaller* depends on whether registral direction is same or different. In the former case, smaller means smaller by more than a minor third. In the latter case, smaller means smaller by a more than a major second. Figure 3 shows as the shaded cells on the top half for large implicative intervals those combinations of implicative and realized intervals that satisfy the principle of intervallic difference. For example, the sequences C_4–G_4–F_4 (7 semitones, 2 semitones different direction) and C_4–B_4–D_5 (11 semitones, 3 semitones same direction) fall in shaded cells, but not the sequence C_4–G_4–B_4 (7 semitones, 4 semitones same direction). Again, this principle is treated here, as in the implication-realization model, as all-or-none.

Before describing the three remaining principles underlying the bottom-up component of the implication-realization model, I will show that the two principles of registral direction and intervallic difference define not only process and reversal, but the complete set of basic melodic structures described by Narmour. These structures fall into two categories: *prospective* structures, which are denoted by initials without parentheses, and *retrospective* structures, which are denoted by initials with parentheses. In the theory, this distinction produces a set of correspondences between the melodic structures for small and large implicative intervals; possible psychological effects of the distinction are unclear. These melodic structures will be described with reference to both Figure 4 and Table 2.[36]

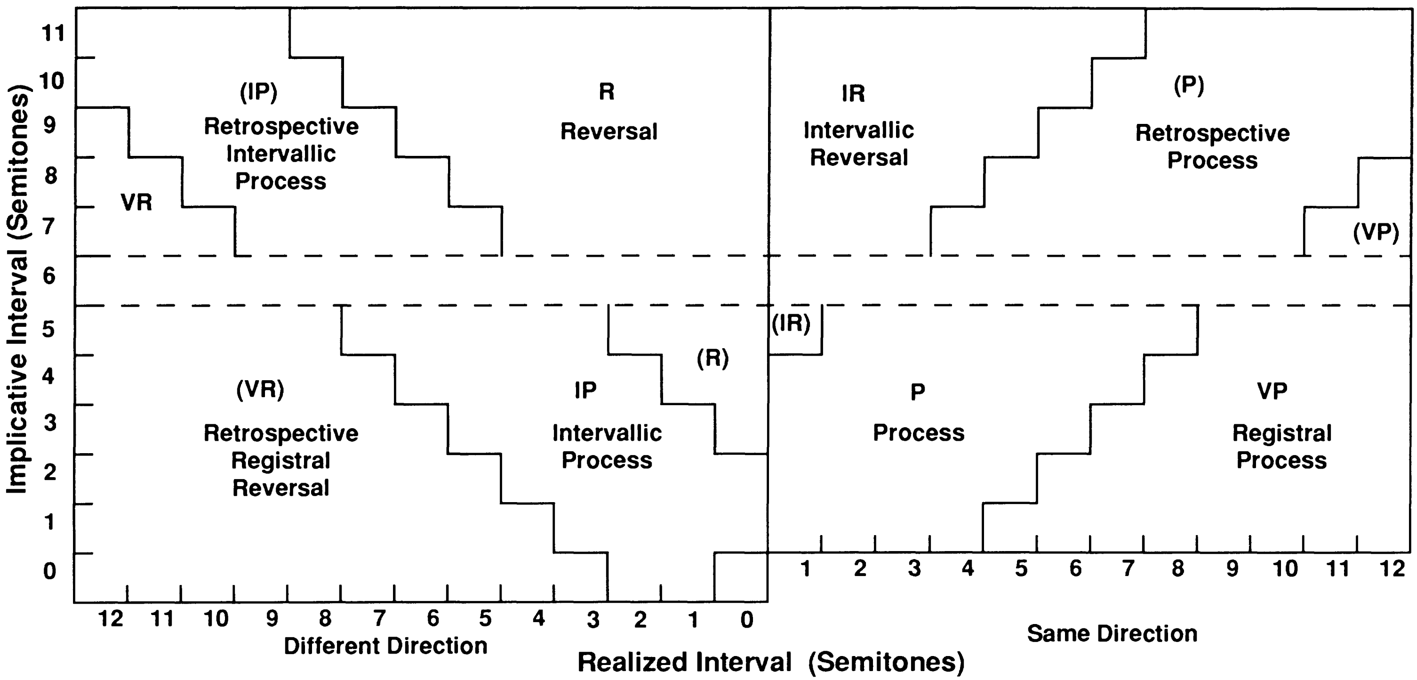
Consider first the melodic structures for small implicative intervals. *Process*, defined as above and denoted P, is the case in which both principles of registral direction and intervallic difference are satisfied; that is, the realized interval is in the same direction as the implicative interval and is similar in size. The region shown for process in the lower right quadrant of Figure 4 is the intersection of the regions for the principles of registral direction and intervallic difference shown in Figures 2 and 3. *Intervallic process*, denoted IP, is the case in which only the principle of intervallic difference is satisfied; the realized interval is similar in size but the realized interval is in a different direction from the implicative interval (and the region for intervallic process shown in the lower left quadrant is shaded in Figure 3 but not Figure 2). *Registral process*, denoted VP (where V = Vectorial), is the case in which only the principle of registral direction is satisfied; the realized interval is in the same direction as the implicative interval but is larger (and the region shown for registral process in the lower right quadrant is shaded in Figure 2 but not Figure 3).

The three remaining melodic structures for small implicative intervals are retrospective: *retrospective reversal*, denoted (R), *retrospective intervallic reversal*, denoted (IR), and *retrospective registral reversal*, denoted (VR). These are called *retrospective* melodic structures because they are heard in retrospect as variants of the corresponding prospective melodic structures for large implicative intervals which are defined next. The logic of the names given to these retrospective structures can be understood by comparing the retrospective structures for small implicative intervals to the corresponding prospective structures for large implicative intervals going out along the diagonals of Figure 4 (retrospective reversal, (R), to reversal, R; retrospective intervallic reversal, (IR), to intervallic reversal, IR; and retrospective registral reversal, (VR), to registral reversal, VR), and also by comparing the definitions of the corresponding structures in Table 2.

The melodic structures for large implicative intervals are defined in a similar way. *Reversal*, defined as above and denoted R, is the case in which both principles of registral direction and intervallic difference are satisfied; the realized interval is in a different direction and is smaller than the implicative interval. The region for reversal in the upper left quadrant of Figure 4 corresponds to the intersection of the regions for registral direction and intervallic difference in Figures 2 and 3. *Intervallic reversal*, denoted IR, is the case in which only the principle of intervallic difference is satisfied (and the region for intervallic reversal in the upper right quadrant is shaded in Figure 3 but not in Figure 2). *Registral reversal*, denoted VR, is the case in which only the principle of registral direction is satisfied (and the region for registral reversal in the upper left quadrant is shaded in Figure 2 but not Figure 3). Again, three retrospective structures are defined: *retrospective process*, (P), *retrospective intervallic*

[36]Duplication, D, and its variants, intervallic duplication, ID, and retrospective intervallic duplication, (ID), are not marked, as they can be treated as special cases of process, P, intervallic process, IP, and retrospective intervallic process, (IP), respectively.

Figure 4: The classification of pairs of implicative and realized intervals into a comprehensive set of basic melodic structures.



Notes: VR = Registral Reversal (R) = Retrospective Reversal (IR) = Retrospective Intervallic Reversal (VP) = Retrospective Registral Process

Table 2. Definition of Basic Melodic Structures

Basic Melodic Structure	Direction of Realized Interval Relative to Implicative Interval	Size of Realized Interval Relative to Implicative Interval[1]
For Small Implicative Intervals:		
Process, P	Same[2]	Similar[3]
Intervallic Process, IP	Different	Similar[3]
Registral Process, VP	Same[2]	Larger
Retrospective Reversal, (R)	Different	Smaller
Retrospective Intervallic Reversal, (IR)	Same[2]	Smaller
Retrospective Registral Reversal, (VR)	Different	Larger
For Large Implicative Intervals:		
Reversal, R	Different[2]	Smaller[3]
Intervallic Reversal, IR	Same	Smaller[3]
Registral Reversal, VR	Different[2]	Larger
Retrospective Process, (P)	Same	Similar
Retrospective Intervallic Process, (IP)	Different[2]	Similar
Retrospective Registral Process, (VP)	Same	Larger

[1]If registral direction of implicative and realized intervals is the same, smaller means by more than a minor third, similar-sized means within a minor third, and larger means by more than a minor third. If registral direction is different, smaller means by more than a major second, similar-sized means within a major second, and larger means by more than a major second.
[2]Satisfies principle of registral direction.
[3]Satisfies principle of intervallic difference.

process, (IP), and *retrospective registral process*, (VP). As before, each of these retrospective melodic structures for large implicative intervals corresponds to a prospective melodic structure for small implicative intervals. The correspondence between retrospective and prospective structures (retrospective process, (P), to process, P; retrospective intervallic process, (IP), to intervallic process, IP; and retrospective registral process, (VP), to registral process, VP) can be see by comparing the regions in Figure 4 and the definitions in Table 2.

The melodic structures of the implication-realization model, just described, derive from the two principles of registral direction and intervallic difference. Three additional principles of melodic expectancy are implicit in the model. The first of these is *registral return*, and it refers to proximity between the first tone of the implicative interval and the second tone of the realized interval. *Exact registral return*, denoted aba, is the case in which these tones are the same, for example, C_4–A_3–C_4 (3 semitones, 3 semitones different direction). *Near registral return*, denoted aba[1], is the case in which these tones are within a major second of one another (but are not identical), for example, C_4–G_3–B_3 (5 semitones, 4 semitones different direction). The shaded cells in Figure 5 show the combinations of implicative and realized intervals that have either exact or near registral return. This principle is treated here as all-or-none but could, in an elaboration of the model, be treated as graded in strength.

The next principle is *proximity*, and it refers to the size of the realized interval. This principle is independent of the size and direction of the implicative interval. On the parametric scale of musical interval size, intervals that are a perfect fourth or smaller are *proximate*, and all larger intervals are *non-proximate*. The principle of proximity can be represented as in Figure 6, with the degree of proximity grading off from maximum for the unison to minimum for the largest proximate interval, the perfect fourth. All larger intervals are non-proximate and the corresponding region is not shaded. The sequence C_4–E_4–F_4 (4 semitones, 1 semitone same direction) would be strongly proximate, C_4–E_4–A_4 (4 semitones, 5 semitones same direction) would be weakly proximate, and the sequence C_4–E_4–B_4 (4 semitones, 7 semitones same direction) would be non-proximate. Note that, unlike the other principles described so far, proximity is treated here as graded in strength.

The final principle implicit in the implication-realization model, *closure*, describes closural effects of different pairs of implicative and realized intervals. As noted earlier, the model holds that a number of factors, including pauses, changes of tone duration, meter, and harmony, produce closure, and these factors are important for determining *where* points of implication occur. The principle of closure is a restricted sense of closure and specifies, *given* that an implicative interval has occurred, *which* realized intervals will produce closure (independent of these other factors). Closure occurs either when the registral direction changes (for example, C_4–F_4–E_4, 5 semitones, 1 semitone different direction) or when

Figure 5: The principle of registral return. Registral return refers to proximity to the first tone of the implicative interval.

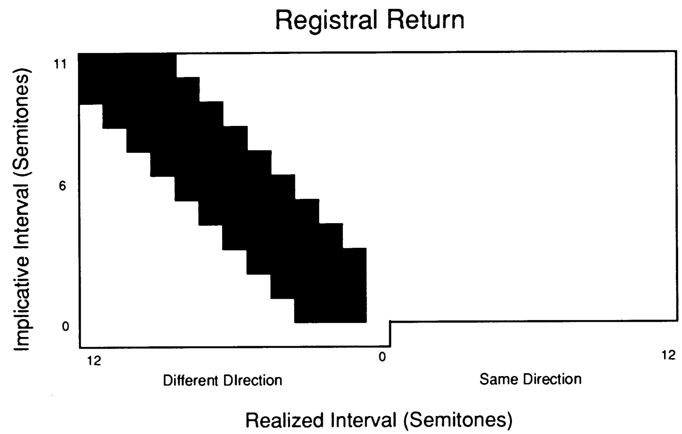
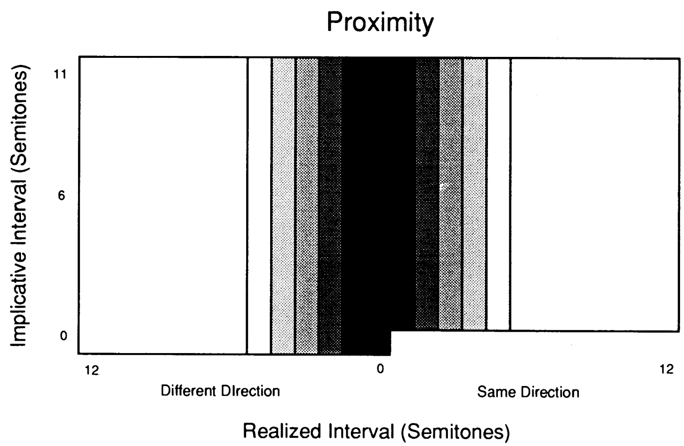
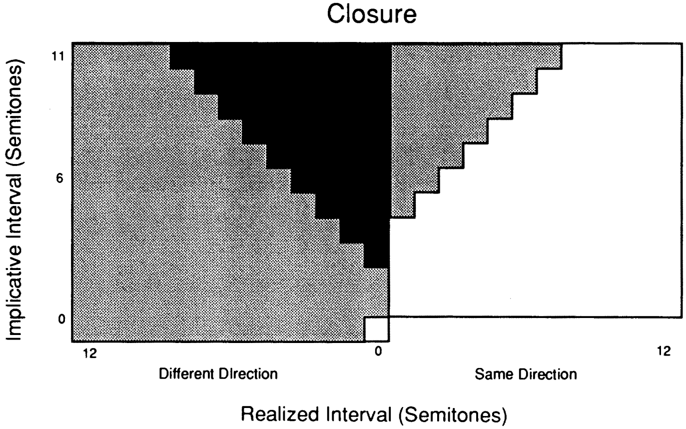


Figure 6: The principle of proximity. Proximity occurs when the size of the realized interval is a perfect fourth or smaller.



a large implicative interval is followed by a smaller realized interval (for example, C_4–G_4–A_4, 7 semitones, 2 semitones same direction). In cases in which both occur, closure would be strongest (for example, C_4–A_4–G_4, 9 semitones, 2 semitones different direction). The different strengths of closure are indicated in Figure 7: both conditions (dark shading), only one condition (light shading), and neither condition (no shading). Thus, this principle is also treated as graded in strength.

Figure 7: The principle of closure. Closure occurs when the registral direction changes, or when a large interval is followed by a smaller interval.



EXPERIMENTAL DESIGN

The five principles formalized as described above provide a basis for developing a quantitative model that can be tested against the results of perceptual experiments. In the experiments reported here, melodic fragments ending on unclosed, implicative intervals were extracted from musical pieces in three different styles. A single tone, the "continuation tone," was added at the end of the fragments in the rhythmic position of the next tone in the original music. The continuation tone varied from trial to trial until each tone in the set of continuation tones for that experiment had been presented with each melodic fragment. The sequences were played in a synthesized piano timbre under computer control, which was also used to record and analyze the data.[37]

The listeners were told that they would hear fragments of melodies that began at the beginning of a phrase but ended in the middle of a phrase. They were instructed to rate how well the additional single tone (the continuation tone) *continued* the melodic fragment. For this, they used a numerical rating scale ranging from 1 ("extremely bad continuation") to 7 ("extremely good continuation"). Listeners were told that each melodic fragment may still sound incomplete even with the added tone, and they should *not* rate how well the test tones *completed* the fragments, but rather how well the test tones *continued* the fragments. Listeners heard a number of practice trials so they could ask any questions they might have about the instructions. Then the actual experiment began with a block of trials for each melodic fragment. In each experiment, the order of the melodic fragments was random, and the order of the continuation tones was random. At the beginning of each block of trials, the fragment was presented three times so that listeners could become familiar with it.

The final, implicative intervals of the fragments varied in both size and direction, so that both small and large implicative intervals were equally represented in both ascending and descending forms. The first experiment employed eight

[37]Adrian Robert wrote the computer program that controlled the experiment and collected the data. E. Glenn Schellenberg selected the musical excerpts, conducted the experimental sessions, carried out preliminary statistical analyses, and drafted some of the figures.

melodic fragments from British folk songs (shown as Examples 1.1–1.8).[38] The continuation tones were all diatonic scale tones within the two-octave range centered on the final tone of the fragment. Ten of the listeners who participated in the experiment were musically trained, and the remaining ten listeners were untrained. The second experiment employed eight fragments from Webern's *Lieder*, Op. 3, 4, and 15 (shown as Examples 2.1–2.8). The continuation tones were all chromatic scale tones within the two-octave range. The listeners in this experiment were generally unfamiliar with the atonal style; thirteen were musically trained and thirteen were untrained. The third experiment employed twelve fragments from Chinese folk songs (shown as Examples 3.1–3.12).[39] Continuation tones were all pentatonic scale tones within the two-octave range. Eight listeners were native Chinese; eight listeners were native Americans.[40]

EXPERIMENTAL RESULTS AND DISCUSSION

The results will be presented and discussed in three parts. The first part summarizes tests of the five bottom-up principles underlying the implication-realization model. The statistical analysis compares the listeners' numerical ratings of the melodic continuations with a quantitative formulation of the principles. The data used in this first part were averaged across listeners because preliminary analyses found the listeners gave similar responses. The second part compares the

Examples 1.1–1.8: The melodic excerpts from British folk songs used in the first experiment.

responses for the different melodic structures with the model's predictions concerning their relative degree of surprise. Again, these analyses were done on the data averaged across listeners. The third part presents tests showing consistent

[38]Cecil J. Sharp, *English Folk Songs*, Vols. 1–2, selected ed. (London: Novello & Company, 1920); R. Palmer, ed., *Folk Songs Collected by Ralph Vaughan Williams* (London: J. M. Dent & Sons, 1983).

[39]*Chung-kuo Min Kuo Hsuan* (Chinese Folk Songs) (People's Republic of China: People's Music Publishing Company, 1980).

[40]The Chinese listeners had resided in the United States for an average of 2 years and 7 months and reported having been primarily exposed to Chinese music.

Examples 2.1–2.8: The melodic excerpts from atonal songs used in the second experiment.

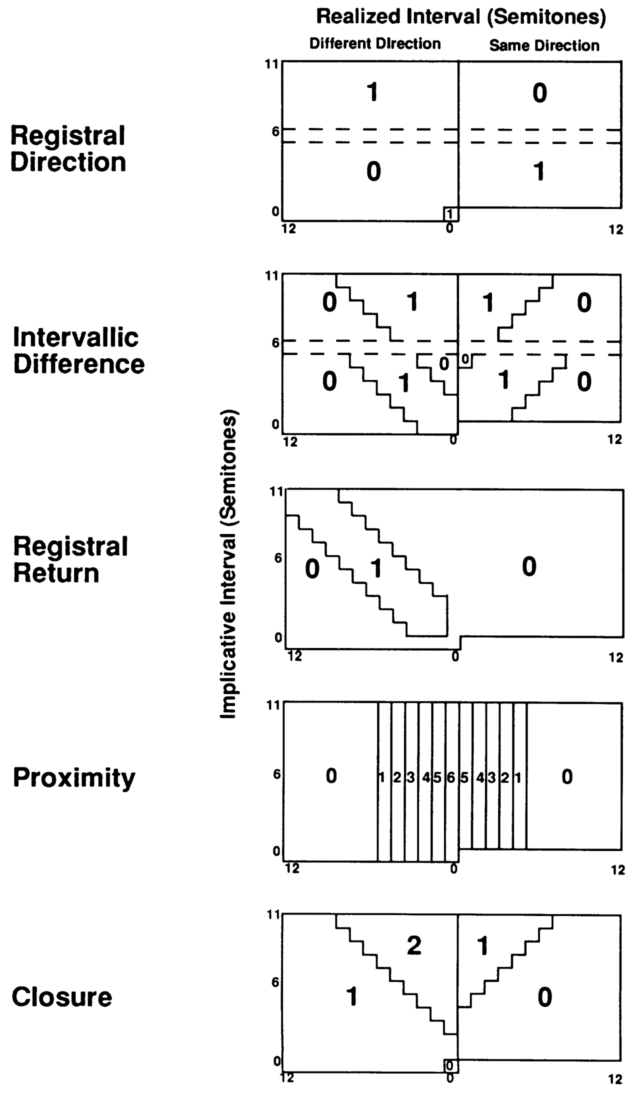
Examples 3.1–3.12: The melodic excerpts from Chinese folk songs used in the third experiment.

response patterns across listeners despite differences in musical training and familiarity with the styles of music.

The first step in testing the five bottom-up principles described above (registral direction, intervallic difference, registral return, proximity, and closure) was to translate them into numerical values. Figure 8 shows the numerical values assigned to each combination of implicative and realized intervals. When a principle is all-or-none, 1 was assigned to combinations that satisfy the principle and 0 was assigned to the combinations that do not satisfy the principle. When a principle is graded in strength, the numbers were chosen to represent the gradation. The five principles coded in this way are called predictor variables and are denoted RD, ID, RR, PR, and CL, respectively.

The second step was to assign the appropriate values of the predictor variables to each trial in each of the three experiments. To take a concrete example, consider Example 1.1 in which the implicative interval (A_4–G_4) is a major second. On one trial, the continuation tone presented following the fragment was E_4. For this particular combination of implicative and realized intervals (2 semitones, 3 semitones same direction), the value of RD = 1 because the continuation tone is in the expected direction (small implicative intervals imply registral direction will continue). The value of ID = 1 because the implicative interval (A_4–G_4) and realized interval (G_4–E_4) are similar in size (small implicative intervals imply similar-sized intervals). The value of RR = 0, because the continuation tone (E_4) is farther than a major second from the first tone (A_4) of the implicative interval. The value of PR = 3 because the continuation tone (E_4) is moderately proximate to the second tone of the implicative interval (G_4). And, the value of CL = 0, because neither condition for closure obtains (reversal of direction, or large interval followed by a smaller interval). So, this gives five values for this particular combination of implicative and realized intervals. To take another example, the continuation tone on another

Figure 8: Quantified predictor variables for the perceptual tests of the implication-realization model. The numbers represent the degree to which the pairs of implicative and realized intervals satisfy each of five principles.



trial was B_4 (2 semitones, 4 semitones different direction), with RD = 0, ID = 1, RR = 1, PR = 2, and CL = 1. Five numerical values were found in this way for each of the 120 trials of the first experiment, the 200 trials of the second experiment, and the 132 trials of the third experiment, each representing different combinations of implicative and realized intervals.

The analysis of the data used the statistical technique of multiple regression.[41] This technique assesses whether a dependent variable, Y, can be explained well as a weighted sum of a number of predictor variables. In the present case, the dependent variable was the listeners' ratings of the continuation tones, and the predictor variables were the numerical codings of the five principles for each trial. The analysis tested whether listeners tended to give higher ratings to continuation tones that fit with the model's predictions for melodic expectancy as formalized in the five principles. Specifically, it tested whether weights, w, could be found such that the ratings of the continuation tones, Y, were approximately equal to $w_{RD}RD + w_{ID}ID + w_{RR}RR + w_{PR}PR + W_{CL}CL$ (plus a constant). The weights allow for the possibility that some of the principles may be stronger than others. The statistical method finds the weights that provide the best possible fit to the data, evaluates the overall fit of the model, and evaluates the contribution of the individual predictor variables.

The results of the statistical analysis strongly supported the implication-realization model. Table 3 summarizes the results for the data averaged across listeners in each of the three experiments. The multiple correlation, R, shown at the top

[41] Richard B. Darlington, *Regression and Linear Models* (New York: McGraw, 1990); Allen L. Edwards, *Multiple Regression and the Analysis of Variance and Covariance* (San Francisco: Freeman, 1979). Computer algorithms for performing multiple regressions are included in most statistical packages, such as *Systat* and *Statview*.

Table 3. Results of Multiple Regression Testing Model

	British Folk Songs	Atonal Songs	Chinese Folk Songs
Multiple Regression Model	$R = .846$ $p < .0001$ $N = 120$	$R = .714$ $p < .0001$ $N = 200$	$R = .849$ $p < .0001$ $N = 132$
Registral Direction	$w_{RD} = .17$ $p < .005$	$w_{RD} = .32$ $p < .0001$	$w_{RD} = .12$ $p < .05$
Intervallic Difference	$w_{ID} = .17$ $p < .1$	$w_{ID} = .09$ n.s.	$w_{ID} = .16$ $p < .05$
Registral Return	$w_{RR} = .31$ $p < .0001$	$w_{RR} = .21$ $p < .0005$	$w_{RR} = .16$ $p < .01$
Proximity	$w_{PR} = .50$ $p < .0001$	$w_{PR} = .51$ $p < .0001$	$w_{PR} = .67$ $p < .0001$
Closure	$W_{CL} = .21$ $p < .005$	$W_{CL} = .16$ $p < .05$	$W_{CL} = .30$ $p < .05$
Unison	$W_{UN} = -.06$ n.s.	$W_{UN} = -.19$ $p < .05$	$W_{UN} = -.31$ $p < .01$
Tonality	$W_{TN} = .17$ $p < .005$	$W_{TN} = .13$ $p < .05$	$W_{TN} = .11$ $p < .05$

Note: n.s. = not statistically significant

of each column, is a measure of how well the dependent variable was explained by the model. An $R = 1$ means that the dependent variable can be predicted perfectly by the predictor variables; an $R = 0$ means that the dependent variable cannot be predicted at all by the predictor variables. The obtained values of R were quite high, which means that the ratings of the continuation tones were quite well predicted by the principles as quantitatively coded (with the addition of two other variables described below). The statistical significance of each multiple correlation, R, is measured by the corresponding value of p, shown below. The p value indicates the probability that this value of R would occur by chance (that is, if the dependent variable were randomly chosen from a distribution of values that are in fact unrelated to the predictor variables). For all three experiments the probability values, p, were less than .0001, which means that R values this large are highly unlikely to occur by chance. (By convention, any p value less than .05 is considered statistically significant.) The N values shown are simply the number of trials in each experiment; they are included in the table because the statistical significance of a multiple correlation depends on this value.

The analyses generally validated the five principles as they were numerically coded. The relative strength of each of the predictor variables is indicated in two ways in the results of the regression analysis, by the weight, w, and by the probability, p, of each of the variables. These are also given in the table. Most of the variables were individually significant as measured by the p values. Overall, the weakest contribution was made by intervallic difference, and the strongest contribution was made by proximity. The somewhat different patterns of weights across the three experiments may suggest possible differences in the relative strengths of the principles across styles, but they may also simply reflect differences between the particular excerpts chosen for the three experiments.

Preliminary analyses suggested that two additional predictor variables should be added to the model. Listeners tended to rate realized intervals of unisons (0 semitones) somewhat lower than predicted, so a compensatory variable called "unison" (which received a negative weight) was added to the model. This suggests that the perceptual effects of unisons may be qualitatively different from other realized intervals and not governed by the same principles. A variable called "tonality" was also added to the model for each experiment. These variables were developed from previous psychological results. For the first experiment, the tonality variable was the experimentally quantified tonal hierarchy for the major or minor key of the melodic fragment.[42] For the second experiment, the tonality variable was a measure of the degree to which the realized interval fits tonally with the implicative interval.[43] For the third experiment, the tonality variable coded whether or not the pentatonic scale tones belonged to the basis set of the mode.[44]

In considering these results, it is important to bear in mind the model's precise parameterization along the scale of musical pitch. This gives the sharply defined boundaries for the principles in the grid representation. The good fit of the data by the quantified principles suggests that these boundaries are at least approximately correct. One contribution that can be made by experimental tests is to evaluate whether perceptual effects might be modeled better by a somewhat different formulation. An analysis that considers alternative formulations with a more comprehensive set of data will be presented elsewhere[45]; the analysis presented here used the principles as they are currently parameterized in the implication-realization model.

A second issue concerning the model testing deals with the interrelationships among the five principles. The five principles are not all mutually compatible. Realized intervals that satisfy one principle do not necessarily satisfy the others, but all the principles are presumed to influence melodic expectancies to some degree. This is supported by the good fit of the weighted sum of the quantified principles. A problem of a somewhat technical nature arises, however, in interpreting the weights. This is because the principles are not all independent. For example, realized intervals that satisfy the principle of intervallic difference also tend to be proximate and produce closure. Thus, trade-offs between the weights for the related principles are possible. This problem of non-independence is also treated elsewhere.[46]

[42]Carol L. Krumhansl and Edward J. Kessler, "Tracing the Dynamic Changes in Perceived Tonal Organization in a Spatial Representation of Musical Keys," *Psychological Review* 89 (1982): 334–68. In the experiment, a strong key-defining context such as a IV–V–I cadence in a major or minor key was followed by each of the twelve tones of the chromatic scale. Listeneɪs rated these tones in terms of how well they "fit with" the key-defining context. The rating of the tonic was highest, followed by the third and fifth scale degrees, then the other scale degrees, and finally the nondiatonic tones.

[43]This measure was the maximum correlation between the three tones of the implicative and realized interval and the tonal hierarchy of the twenty-four major and minor keys as quantified by Krumhansl and Kessler, "Tracing the Dynamic Changes." The measure is based on the key-finding algorithm of Carol L. Krumhansl and Mark A. Schmuckler described in Krumhansl, *Cognitive Foundations of Musical Pitch*. It was used in Krumhansl, Sandell, and Sergeant, "The Perception of Tone Hierarchies and Mirror Forms in Twelve-Tone Serial Music," which also found an influence of tonal implications on the perception of excerpts from Schoenberg's *Wind Quintet*, op. 26, and *String Quartet No. 4*, op. 37.

[44]The basis set consists of the three structurally significant tones within the mode. See Sin-Yan Shen, *Chinese Music and Orchestration: A Primer on Principles and Practice* (Chicago: Chinese Music Society of North America, 1991). This variable coded as 1 those tones in the basis set of the mode, and coded as 0 the other scale tones. Recall that the continuation tones were all members of the pentatonic scale. A similar coding of tones based on partitionings of the octatonic collection was used by Krumhansl and Schmuckler, "The *Petroushka* Chord: A Perceptual Investigation."

[45]An extensive analysis of alternative formulations of the bottom-up component of the implication-realization model will be presented in Carol L. Krumhansl, "Tonal and Melodic Implications of Musical Intervals," manuscript in preparation. Briefly, the experiment to be reported used implicative intervals ranging from a descending major seventh to an ascending major seventh. The continuation tones were all tones in the chromatic scale in two octaves. These data supported the principle of registral direction, and suggested that both proximity and registral return should be modified to include broader ranges of pitch differences. When these modifications were made, neither intervallic difference nor closure had residual effects. The unison effect was again found, and also an effect of octaves. Moreover, consonance and dissonance were found to influence the results over and above the effects of tonality.

[46]Ibid. The data from that experiment was used to develop a model containing independent predictor variables. These include the principle of registral direction as originally formulated in the implication-realization model, and modified versions of the principles of proximity and registral return.

A third issue concerns the effects of tonality and other influences of style knowledge. Although this article focuses on the bottom-up component of the implication-realization model, effects of style knowledge would be expected, given the numerous psychological studies cited earlier documenting influences of such factors as scale, harmony, and key on perceptual judgments. In the present test of the implication-realization model, additional variables were entered into the analyses representing effects of scale and tonality. These are extrinsic to the implication-realization model itself, but are not incompatible with it. The model posits both intra-opus and extra-opus knowledge as additional sources of melodic expectancy, but without the precise specification of the bottom-up component. Consequently, the model tested here constructed variables representing style knowledge from prior experimental results. These variables were independent (in the technical sense introduced above) of the five principles, and the results of the analyses suggest that they neither dominate over nor interact strongly with the five principles.

The second part of the results concerned the model's predictions for the degree of surprise of the melodic structures. The theory predicts a graded degree of surprise along a continuum.[47] The order of the melodic structures along the continuum from least to most is shown in Table 4, first for structures with small implicative intervals and then for structures with large implicative intervals. When the implicative interval is small, process, P, has the effect of no surprise; intervallic process, IP, has a small effect of surprise; followed by registral process, VP; and finally retrospective registral reversal, (VR), with the greatest degree of surprise. When the implicative interval is large, reversal, R, has the effect of no surprise; intervallic reversal, IR, has a small effect of surprise; followed by retrospective intervallic process, (IP); and finally retrospective process, (P), with the greatest degree of surprise. Analyses examined whether the ratings of the continuation tones confirmed these predictions. High average ratings would be expected to correspond with low degrees of surprise, and vice versa. The average ratings for all pairs of implicative and realized intervals for each of the melodic structures were computed and are also shown in the table.

Table 4. Predicted Degree of Surprise of the Basic Melodic Structures and Corresponding Perceptual Judgments

	British Folk Songs	Atonal Songs	Chinese Folk Songs
For Small Implicative Intervals:			
Process, P	4.73	4.67	4.81
Intervallic Process, IP	5.22	4.77	5.23
Registral Process, VP	2.33	3.47	2.30
Retrospective Registral Reversal, (VR)	2.89	2.98	3.10
For Large Implicative Intervals:			
Reversal, R	4.92	5.05	5.33
Intervallic Reversal, IR	3.89	4.03	3.98
Retrospective Intervallic Process, (IP)	4.48	4.52	3.90
Retrospective Process, (P)	2.25	3.03	2.43

The ratings conformed quite well with the predictions, with three exceptions. First, the ratings for intervallic processes, IP, tended to be somewhat higher than the ratings for processes, P (lines 1 and 2 of the table). Second, the ratings for retrospective intervallic processes, (IP), tended to be somewhat higher than the ratings for intervallic reversals, IR (lines 6 and 7 of the table). Both intervallic processes, IP, and retrospective intervallic processes, (IP), may have received relatively high ratings because they satisfy the principle of registral return. This can be seen by noting that the regions

[47]Narmour, *Analysis and Cognition of Basic Melodic Structures*, 343. Only those melodic structures with a sufficient number of experimental observations to make reliable estimates are considered.

for these two melodic structures in Figure 4 are included in the shaded region for registral return in Figure 5. Third, the ratings for retrospective registral reversals, (VR), tended to be somewhat higher than registral processes, VP (lines 3 and 4). This finding may be explained by the different degrees of closure produced by the two melodic structures. Retrospective registral reversals, (VR), are more closed than registral processes, VP. This can be seen by comparing the regions for these two structures in Figure 4 with the corresponding regions in Figure 7. In sum, analyzing the data in this way provides further support for the idea that melodic expectancies are influenced not only by registral direction and intervallic difference (which define the basic melodic structures) but also by the additional principles of registral return and closure.

Finally, the degree of consistency found across listeners in these experiments might be the most unexpected result. Listeners produced similar melodic continuation judgments despite substantial variation in their musical training and familiarity with the musical styles. In the first and third experiments, each listener's judgments correlated significantly with those of every other listener. In the second experiment, agreement between listeners was also strong, with the exception of four moderately trained listeners who responded idiosyncratically to the atonal melodies.[48] Moreover, the fit of the quantified model was statistically significant for every individual listener in the first and third experiments, and for all but two listeners in the second experiment.

This consistency bears emphasis for two reasons. First, the bottom-up component of the implication-realization model is directed at identifying general principles of music cognition that operate independently of the listener's musical experience. The degree of consistency found across listeners suggests that musical expectancy is a phenomenon that is compatible with this goal. Second, the consistency suggests that musical expectancy is, in part, a basic psychological response that does not depend on specialized training, extensive experience with the particular musical style, or knowledge of technical concepts or vocabulary.

In this connection, a few words might be said about the task that was used in the experiments. A natural question is what effect the particular instructions used in the experiment might have had on the pattern of responses. One approach to this question is to vary the instructions and determine if this produces a changed pattern. The first experiment was, in fact, repeated with two different sets of instructions with a separate group of listeners. The two alternative instructions were to rate how "pleasant" the continuation tone sounded in the context of the melodic fragment, and to rate how "interesting" the continuation tone sounded in the context of the melodic fragment. Again, half of the listeners were musically trained, and half were untrained. The "pleasant" instructions and the instructions to judge the goodness of the continuations produced virtually identical results. For most listeners, the "interesting" ratings were essentially the opposite of the "pleasant" ratings, that is, pleasant continuations were judged uninteresting. For a few listeners, "interesting" was responded to as though it were a synonym of "pleasant." The essential point is that all three sets of verbal instructions resulted in essentially the same underlying continuum of responses.

This strategy of using multiple methods to study a psychological phenomenon is called "converging operations." Any particular method may impose its own distinctive characteristics on the data, so it is important to use different methods. If the results converge, one can gain confidence that

[48]Two of these tended to give higher ratings to continuation tones that had appeared more recently in the fragment; two showed the opposite pattern. Similar response patterns were found in the study by Krumhansl, Sandell, and Sergeant, "The Perception of Tone Hierarchies and Mirror Forms in Twelve-Tone Serial Music."

the data reflect the underlying psychological phenomenon and not the measurement technique itself. This is especially true if the two methods are quite different. In the present case, the research of James Carlsen and collaborators provides a source of converging evidence for the predictions of the implication-realization model.[49] Recall that their method was to present listeners with two successive tones, ranging from a descending octave to an ascending octave. The listeners, who were music students, were instructed to respond by singing the tones that they expected would follow the stimulus interval in a melody.

The first tone produced by the listeners in their melodic continuations can be considered analogous to the continuation tones of the present experiments. Statistical analyses showed that the production data are well predicted by the implication-realization model. The fit of the data by the model coded as described above was highly significant statistically, and each of the quantified predictor variables was significant individually. The only notable difference from the results reported here was that the contribution made by the principle of closure was considerably weaker than in the experiments reported here. This difference, however, would be expected given their instructions to continue the melody. The degree to which the perception and production measures agree is remarkable given the very different tasks and stimulus materials. The convergence between the results suggests that they are both tapping into the same underlying system of musical expectancy.

CONCLUSIONS

These experiments have sampled a restricted range of musical styles and listeners, and the rules of statistical inference prevent generalizing beyond the particular musical segments employed in the experiments and beyond listeners similar to those who participated. Nevertheless, the uniformity with which the present results supported the implication-realization model encourages the view that the model has successfully codified psychological principles governing melodic expectancy. Assuming that further psychological tests continue to support the model or related music-theoretic proposals, what can be taken as the significance of the results in the broader sense? Psychologically, two points seem particularly noteworthy. The first concerns the role that principles of perceptual organization may play in melodic expectancy. The second concerns the absence of effects of musical training and enculturation.

Narmour considers the bottom-up component of the implication-realization model to embody general principles of perceptual organization operating in music cognition. These principles, often attributed to the Gestalt psychologists, specify the properties on the basis of which perceptual wholes are built up from the parts. These laws are named, and most often illustrated, with reference to visual patterns, so that it is difficult to be precise in applying them to another domain. But the principles of the implication-realization model, especially those for small implicative intervals, do lend themselves to this kind of comparison. In essence, the model says that given a small implicative interval, listeners expect that the melodic direction will continue (in Gestalt terms, good continuation); the next interval will be small (proximity) or similar in size (similarity); or the melody will reverse direction and return to the earlier pitch range (symmetry). In Gestalt theory, these principles determine how parts are combined into wholes. Thus the psychological function of these principles in music cognition may be to join successive tones into coherent melodic patterns.

The principles describing melodic expectancies for large implicative intervals are harder to derive from general prin-

[49] Carlsen, "Some Factors which Influence Melodic Expectancy"; Unyk and Carlsen, "The Influence of Expectancy on Melodic Perception."

ciples of perceptual organization. Gestalt principles of symmetry and proximity play a role for large implicative intervals just as they do for small implicative intervals. But the core concept for large implicative intervals is reversal, that is, the expectancy that the melodic direction will reverse and the interval size will decrease. These conditions, according to the theory, are precisely those that engender the strongest sense of closure. To speculate on a possible psychological basis, large intervals, which are relatively infrequent in tonal melodies, may create a sense of instability, which in turn implies that resolution or closure will follow. At a more general level, closure is important for establishing the boundaries of coherent, complete, and well-formed units, which is important for efficient cognitive representation. Thus from a psychological point of view, the principles of the implication-realization model may serve to facilitate grouping and segmentation of the on-going perceptual information.

Although the bottom-up component of the model can be related to some degree to general principles of perceptual organization, it should be emphasized that they are precisely stated in terms of musical interval size and direction. That is, the principles take on values particular to the musical domain. According to the theory, these principles are innately specified and universal. Although this claim would be difficult to assess, the results of these preliminary studies are remarkable for the degree of consistency found across musical styles and across listeners with different levels of musical training and enculturation, as the model would predict. If future research supports the generality of these precisely-specified principles, then it would strengthen the possibility that the mind has modes of processing and representation special to music, and that these modes do not require extensive learning.

This possibility generally runs counter to current conceptions in music psychology that emphasize the importance of knowledge of musical style presumably acquired through extensive experience. Research with infants, however, has increasingly found early sensitivity to music and music-like patterns. Sandra Trehub recently summarized experiments showing that infants may be selectively predisposed to the intervals of octaves and fifths,[50] and I have shown in collaborative research with Peter Jusczyk that infants as young as $4\frac{1}{2}$ months of age are sensitive to musical phrasing.[51] Mechthild Papoušek and Hanus Papoušek have documented the presence of many musical elements in infants' early vocalizations.[52] As the data accumulate, psychologists will continue to clarify the relative contributions of uniquely musical abilities, general psychological principles, and acquired knowledge to various musical behaviors. Whatever the balance, it should be clear that the implication-realization model poses a challenge to psychologists by offering a precisely specified proposal for an aspect of music perception that does not require extensive learning and thus might be exhibited quite generally by listeners.

I close with a few general remarks concerning the relationship between music psychology and music theory. Music-theoretic proposals that are oriented toward psychological

[50]Sandra E. Trehub, "The Perception of Musical Patterns by Human Infants: The Provision of Similar Patterns by their Parents," in *Comparative Perception*, Vol. I, *Basic Mechanisms*, ed. M. A. Berkley and W. C. Stebbins (New York: Wiley, 1990), 429–59; Sandra E. Trehub and Laurel J. Trainor, "Listening Strategies in Infancy: The Roots of Music and Language Development," in *Cognitive Aspects of Human Audition*, ed. S. McAdams and E. Bigand (Oxford: Oxford University Press, 1993), 278–327.

[51]Carol L. Krumhansl and Peter W. Jusczyk, "Infants' Perception of Phrase Structure in Music," *Psychological Science* 1 (1990): 70–73; Peter W. Jusczyk and Carol L. Krumhansl, "Pitch and Rhythmic Patterns Affecting Infants' Sensitivity to Musical Phrase Structure," *Journal of Experimental Psychology: Human Perception and Performance* 12 (1993): 627–40.

[52]Mechthild Papoušek and Hanus Papoušek, "Musical Elements in the Infant's Vocalization: Their Significance for Communication, Cognition, and Creativity," in *Advances in Infancy Research,* Vol. 1, ed. L. P. Lipsitt (Norwood, N.J.: Ablex, 1981), 163–224.

issues and informed by the experimental literature are valuable resources for the psychological study of music. In exchange, the experimental results can serve to refine the theoretical proposals, offer complementary techniques, and explicate the psychological foundations underlying musical structure. Nonetheless, the divergent aims, terminology, and methods need to be respected.[53] Music theory seeks to explicate the analytic relations that exist in music, whereas music psychology is concerned with describing the processes underlying musical behaviors. Borrowed terminology may also be problematic unless the disciplinary context that embeds the term is acknowledged. Any music-theoretic proposal is understood in relation to other similar proposals, just as experimental results are interpreted and evaluated with respect to related empirical findings and theoretical claims that have emerged from previous research. Finally, each discipline has its established methodologies, the technical nature of which may impose barriers.

These points of tension between the various disciplines concerned with music were noted also by Helmholtz, who concludes of his own research: "But I can scarcely disguise from myself, that although my researches are confined to the lowest grade of musical grammar, they may probably appear too mechanical and unworthy of the dignity of art, to those theoreticians who are accustomed to summon the enthusiastic feelings called forth by the highest works of art to the scientific investigation of its basis. To these I would simply remark in conclusion, that the . . . investigation really deals only with the analysis of actually existing sensations—that the physical methods of observation employed are almost solely meant to facilitate and assure the work of this analysis and check its completeness—and that this analysis of the sensations would suffice to furnish all the results required for music theory."[54]

ABSTRACT

Music theorists and music psychologists may benefit from increasing awareness of each others' discipline. However, it is necessary to delimit the common ground shared by the disciplines and, at the same time, to clarify basic differences between the two approaches. Toward this end, this article begins with a schematic history of the psychology of music from the point of view of how it has been influenced by music theory. Following this is a brief characterization of the goals, methods, and theoretical commitments of experimental psychology. The article then reports a recent program of research that reflects the direct influence of music theory on psychological experimentation. The experiments test predictions of the bottom-up component of Eugene Narmour's 1990 implication-realization model for melodic expectancy. It is shown that five principles underlie this component leading to the development of an experimentally testable, quantitative formulation. The principles are validated across three different musical styles and listeners varying in music training and experience. Implications for the psychological studies of music are discussed.

[53] See also Eric F. Clarke, "Mind the Gap: Formal Structures and Psychological Processes in Music," *Contemporary Music Review* 3 (1989): 1–13.

[54] Helmholtz, *On the Sensations of Tone*, 6.