Music Perception
Summer 2005, Vol. 22, No. 4, 663–714

A Model of Melodic Expectation

ELIZABETH HELLMUTH MARGULIS
Northwestern University

A model of melodic expectation is proposed. The model assigns ratings to the expectedness of melodic events. The ratings depend on the hierarchic implementation of three primary factors–stability, proximity, and direction–and one secondary factor–mobility. The model explicitly links expectancy ratings to aspects of listeners' experiences of tension in melody. An approach to temporal expectations is discussed but not quantified.

The model is situated within a framework for thinking about a type of schematic melodic expectations. This article assesses the position of these expectations within the broader cognitive processes invoked in listening to music. It suggests methods for investigating the expectations empirically. Additionally, it outlines connections between the theorized expectations and the dynamic, affective contours of musical experience.

Received November 14, 2003, accepted October 13, 2004

EXPECTANCY has long been cited as a generator of musical affect. Leonard Meyer (1956) sparked contemporary theorists' interest in the subject. In *Emotion and Meaning in Music,* he appropriated MacCurdy's (1925) modification of Dewey's (1894) Conflict Theory of Emotions, which suggested that affect, in general, arises from the inhibition of tendency. Meyer proposed that in music, more specifically, "affect . . . is aroused when an expectation—a tendency to respond—activated by the musical stimulus situation, is temporarily inhibited or permanently blocked" (Meyer, 1956, p. 31). This proposal has profound ramifications. According to Meyer, it suggests that "granted listeners who have developed reaction patterns appropriate to the work in question, the structure of the affective response to a piece of music can be studied by examining the music itself. . . . the study of the affective content of a particular work . . . can be made without continual and explicit reference to the respons-

Address correspondence to Elizabeth Hellmuth Margulis, School of Music, Northwestern University, 711 Elgin Rd., Evanston, IL 60208. (e-mail: e-margulis@northwestern.edu)

ISSN: 0730-7829, electronic ISSN: 1533-8312. Please direct all requests for permission to photocopy or reproduce article content to University of California Press's Rights and Permissions website, at www.ucpress.edu/journals/rights.htm.

es of the listener or critic. That is, subjective content can be discussed objectively" (Meyer, 1956, p. 32).

Indeed, the most attractive aspect of the expectational approach to musical affect is the potential it affords for directly connecting the structure of the musical line to the contours of listener experience. Yet this possible connection, which ostensibly motivated the extensive work on expectancy that followed Meyer, has proven difficult to establish.

According to the psychological theories Meyer (1956, p. 20) cites, emotions are

> themselves undifferentiated, affective experience is differentiated because it involves awareness and cognition of a stimulus situation which itself is necessarily differentiated. ... [This] explains and accounts for the existence and nature of the intangible, non-referential affective states experienced in response to music. For in so far as the stimulus situation, the music, is non-referential . . . there is no reason to expect that our emotional experience of it should be referential. The affective experience made in response to music is specific and differentiated but it is so in terms of the musical stimulus situation rather than in terms of extramusical stimuli.

Expectancy violations, in other words, do not translate into affective experiences in any obvious, transparent way. How might one articulate the sensations that violations cause? How might one systematize the relationship between melodic line and listener experience via the conduit of expectancy?

The most expansive treatment of expectation in melody is Narmour's implication-realization theory (1990, 1992). Although comprehensive and penetrating as a theory of melody, the work does not focus on connections between expectancy and affect. Many possible melodic structures are described, but their experiential consequences are mentioned only occasionally. The intermittent references to affect usually entail the identification of moments that might surprise or "shock" a listener, rather than the exposition of a system of dynamic expectancy-based affective fluctuations: a problem given that such moment-to-moment ebbing and flowing seems to characterize musical experience better than isolated injections of surprise.

The present study builds on existing theories of melodic expectation, notably those of Narmour, Larson (1993, 2004), and Lerdahl (2001), reformulating and supplementing the principles in a way that enables the specification of meaningful connections to listener experience. Rather than employ strings of symbols or draw metaphoric relations, the present theory assigns expectedness ratings (predictions about degree of expectedness) to melodic events and associates these ratings with three distinct types of tension (surprise-tension, denial-tension, and expectancy-tension). This approach enables the graphical representation of moment-to-moment expectancy-based fluctuations in affect across a melody's course.

As a candidate for a mechanism of affect, expectation possesses several advantages. The phenomenon is known to be a basic strategy of the human mind; it underlies the ability to bring past experience to bear on the future, the ability to prevent computational overload by constraining the number of outcomes worthy of consideration, and the ability to efficiently direct attention and mental resources (Masson, 1986; Posner & Snyder, 1975). Bharucha (1987) proposes that perceptual facilitation (i.e., expectation) arises naturally out of spreading activation in a network where connectivity encodes musical relationships. The expectational theory of musical affect gains plausibility by relying on a common cognitive process.

The theory also benefits from its focus on the active, real-time experience of listening. Unlike some music theories that seem to assume a listener possesses a mental score of the entire movement before the second note has sounded, a theory built around expectancy absorbs the reality of temporal experience into its foundation, preserving the distinction between past events, which have become fact, and future events, which remain uncertain.[1]

Another significant advantage resides in the theory's appreciation of the complexity of musical experience. Rather than simplistically equating the minor mode with sadness, or a quick tempo with excitement, an expectational model examines a musical landscape as a dynamic succession of events, preserving the constant fluctuation in sensation that characterizes musical experience.

Conceptual Framework

Meyer (1956, p. 25) observes that

> all tendencies, even those which never reach the level of consciousness, are expectations. For since a tendency is a kind of chain reaction in which a present stimulus leads through a series of adjustments to a more or less specified consequent, the consequent is always implied in the tendency, once the tendency has been brought into play.

Narmour's implication-realization theory treats bottom-up, deeply schematic expectations as well as top-down, stylistic-based ones. The present study restricts itself to expectations of the former sort. It uses the term "melodic expectations" narrowly to designate deeply schematic, automat-

---

[1] For a discussion of the gap between final-state theories and real-time listening, as well as an exposition of some potential bridges between the two, see Bigand (2003).

ic primings not available for direct access, primings possibly instantiated within a Fodorian module (Fodor, 1996; Justus & Bharucha, 2001). Although other senses of the term expectation almost certainly apply to listeners' experience of melody, the model treats only those expectations that the listener is not aware of as such. For example, a listener hearing Beethoven's Fifth Symphony for the eightieth time sustains very specific expectations about which melodic events will ensue. Yet the informationally encapsulated module engaged in projective primings cannot access that broader knowledge and continues to produce the same uninformed predictions it would have on the first hearing (see Jackendoff, 1991).

The description "deeply schematic" relates to the distinction that Bharucha (1987) and others make between schematic and veridical expectancies. According to Bharucha, schematic expectations are automatic predictions based on the implicit extrapolation of typical patterns from an extensive corpus of music, whereas veridical expectations represent specific knowledge about the way a particular piece goes. Bharucha uses the deceptive cadence (V-vi) to illustrate the difference; schematic expectations predict continuation to the tonic, even when the listener knows (via a veridical expectation, such as prior experience with the piece) that vi will occur instead.

Although the distinction between schematic and veridical broadly categorizes expectation types, the category of schematic itself encompasses more than one variety of expectation. Specifically, schematic expectations inhabit a continuum from relatively deep to relatively shallow, where depth relates to availability for direct access (from little to much availability), susceptibility to change through exposure (from little to much susceptibility), and scope of application (from more universal to more limited). Examples of increasingly shallow schematic expectations might be: expectations for closure; expectations for cadential closure in tonal music; expectations for common cadence types in music from the classical period; and expectations for common cadence figures in the music of Mozart, where these expectations are increasingly available for access, increasingly susceptible to change through exposure to new pieces within the relevant repertoire, and increasingly limited in scope. The expectations discussed in this article are thought to lie on the extreme deep side of the schematic spectrum, representing little conscious access availability, little susceptibility to change through exposure, and a wide scope of application. To emphasize these characteristics, the expectations are sometimes referred to not simply as "schematic," but as "deeply schematic."

The friction between these deeply schematic expectations and the actually occurring events, the model claims, registers in the listener as experiences of tension. The underlying expectancies are *only* naturally discernible in these effects on tension and affective experience. Accordingly, the sort of melodic expectations discussed here are not equivalent to what composers usually do, to what sounds best, or to the continuations a listener might sing. Composers more likely seek to create an optimal mix of expectedness and unexpectedness, a mix that listeners might try to emulate when singing continuations. The notion of such an "optimal mix" is not new to aesthetics. In a list of rules on the art of writing beautiful melodies, eighteenth-century theorist Johann Mattheson repeatedly advises composers to make judicious use of expected structures—"steps and small intervals are preferred to large leaps"—as well as of unexpected deviations—"one should cleverly vary such small steps" (Mattheson, 1739, in Harriss, 1981, p. 312). In the most comprehensive application of information theory to music, Abraham Moles (1966, p. 169) observes:

> In order for the musical signal to be "intelligible" in the sense of integrally perceived, the average transmitted information, integrated over the maximum extent of presence, must be on the order of the limiting rate of apperception. If it is much lower than this rate, the signal seems uninteresting; if it is much higher, the signal overwhelms the listener and destroys his attention.

Some empirical evidence supports this idea. Simon and Wohlwill (1968), for example, showed that listeners preferred an original, complex musical passage over simplifications of it. Crandall (1967) presented people with sequences of nonsense words and asked them to rate individual words on a good-bad scale. Words that generated moderate uncertainty about the next element in the sequence were rated closer to the good end of the scale than either words that generated absolute certainty or words that generated no certainty. In a study of reactions to printed items, Maddi (1961) showed that people experienced a more positive affect in relation to small deviations from expectedness than they did in relation to large deviations or no deviations.[2] These and other studies suggest that "when novelty is manipulated through degree of deviation from the familiar, . . . an intermediate degree of deviation is preferred" (Berlyne, 1971, p. 194). A melody featuring only the most expected continuations would not possess the rich fluctuations in expectancy-based tension that (in part) make for good, interesting music. It can be broadly assumed that composers try to write interesting music. It can be similarly assumed that listeners try to sing continuations that sound "good" to them. The problem with interpreting these continuations as measures of deeply schematic expectedness is that "good" does not straightforwardly result from such expectedness. If "good" relates to deeply schematic expectancies at all, it relates to the patterns of tension that a mix of expected and unexpected continuations creates.

Several experiments have investigated melodic expectations in the broader (i.e., other than deeply schematic) sense by playing melodies and

---

[2] For a discussion of these and similar studies, see Berlyne (1971).

asking participants to sing continuations (Carlsen, 1981; Lake, 1987). Others have played listeners melodies with different probe tones appended. Listeners were asked to rate how well the different continuations fit (Cuddy & Lunney, 1995; Schellenberg, 1996; Schmuckler, 1989). These experiments probably uncovered a broader type of expectation than that targeted by the present model. First, the experiments required listeners to access their expectations by explicitly evaluating the expectedness of various continuations or by singing expected continuations. The expectations proposed here are theorized to be unavailable for such access. Second, listeners in these experiments were required to rate the goodness of fit of continuations or to sing continuations: both tasks could be understood to reveal (the first explicitly, the second implicitly) what participants thought sounded "good." As already observed, "good" most likely does not arise directly out of expectedness (in fact, constant capitulation to expectedness would most likely sound boring—or "bad"); therefore, these tasks constitute unsatisfactory measures of deeply schematic expectedness. Third, the expectations exposed by the experiments were likely susceptible to the influence of familiarity. For example, had the subjects been played the opening measures of "America the Beautiful" to the "A" of America, they most likely would have sung (or most highly rated) a continuation to "mer" on an ascending major sixth. Yet the present model suggests that even for a memorized tune, schematic expectations will continue to favor a different continuation than the major sixth; it is precisely this expectation, the theory claims, that contributes to the particular affective quality of that leap. If expectations are responsible for the generation of affect, and if repeated exposure could nullify expectations, then familiar music should lose its affective power. Experience suggests otherwise.

Some expectations, and it is these that the present model explores, might operate in an informationally encapsulated way, impervious to the influence of knowledge or familiarity. These are not expectations in an everyday sense, but can be thought of more profitably as schematic primings. To isolate their effects, experiments will have to follow listeners' real-time ratings of tension and correlate them with the expectancy model's predictions (cf. Schubert, 2001–2002; Eerola, Toiviainen, & Krumhansl, 2002), or use event-related-potential (ERP) studies (cf. Besson & Faïta, 1995; Granot & Donchin, 2002; Tervaniemi, Huotilainen, Brattico, Ilmoniemi, Reinikainen, & Alho, 2003) or priming paradigms (cf. Bharucha & Stoeckig, 1986) to reveal expectancy violations. Tillman and Bigand (2004) suggests that the expectations uncovered by priming paradigms are automatic and unaffected by the prior repetition of expectancy-violating passages. For example, a listener will continue to expect the dominant to resolve to the tonic, even after hearing it repeatedly progress to the submediant. This finding supports the distinction proposed between the access availability of schematic and veridical expectations; although, if asked, the listener may attest to expecting the submediant (a veridical expectation), the priming paradigm will reveal a (schematic) expectation for the tonic.

It is important to note that "deeply schematic" does not necessarily suggest "innate." In its stability parameter, the model outlined here incorporates tonal context. Studies have suggested that principles of tonal organization can be learned implicitly through simple exposure at an early age, resulting in their later automatic and modular application (cf. Peretz & Coltheart, 2003). "Deeply schematic" implies the inability of these expectations to be directly accessed. Contrast this with shallowly schematic, more easily accessed and described expectation types (e.g., the expectation for a lyrical second theme in a sonata movement). Studies have shown that priming in chord sequences depends on tonal relationships that are seemingly learned, rather than innate (Justus & Bharucha, 2001), and that these harmonic primings resist repetition priming (Tillman & Bigand, 2004). The robustness of the tonality-based primings uncovered in these experiments suggests that tonal principles absorbed in relatively early stages of development may play a role in informationally encapsulated, schematic expectancies.

The model provides a baseline association between melodic structure and listener experience. This association is an issue best explored empirically. The model outlined here contributes to that project by: (1) clearly specifying the nature of the expectations that might link structure and experience, (2) explicitly specifying the nature of the relationship between structure (expectancy and expectancy violation) and experience (tension types), and (3) providing a foundation on which future inquiries can be built.

The Model

The model assigns expectedness ratings to melodic events. Expectedness ratings are predictions about the amount of deeply schematic expectedness listeners will have for melodic events. Although the symbols in Narmour's implication-realization (I-R) model clearly categorize melodic segments on the basis of interval size and direction, they only secondarily denote the theorized expectedness of each segment. By directly assigning a prediction of expectedness to each event, the present model attempts to make its claims about expectancy more straightforward and accessible. Because musical experience is dynamic and fluctuates in quality from moment to moment, it is advantageous that the model can assign ratings from event to event across the course of a melody. The I-R model's analyses do not address the expectational content of all melodic events, but only of some (as discussed later in the comparative analysis of the opening measures of Mozart K. 282).[3] The present model assigns ratings by hierarchically applying three primary factors—stability, proximity, and direction—and one secondary one, mobility. Temporality is discussed but not quantified. The expectancy ratings are correlated with three tension types, facilitating the graphic depiction of fluctuations in tension across the course of a melody. This article outlines each factor in turn before discussing the factors' combination and joint operation, as well as their theorized impact on the experience of tension.[4]

Of the primary factors, two—proximity and direction—stem from parameters in Narmour's I-R model. One, stability, stems from accounts of tonal pitch space and melodic attraction in Lerdahl (2001). Accordingly, a brief overview of these two extensive theories will permit a better understanding of the present model's approach.

THE I-R MODEL

Narmour's implication-realization theory examines the way Gestalt principles extract implications from style shapes (configurations of a musical parameter such as interval size and directional contour). His basic hypothesis is as follows:

> if in any one parametric simplex the elements of a style-shape relation of similarity, proximity, or common direction occur, then implication of a further style-shape element of similarity, proximity, or common direction takes place, except in cases where implication of intra- or extraopus style-structural complexes are of such empirical conformance so as to interfere. . . . [In other words] . . . If a parametric style shape of elements $a_1 + a_2$ occurs, then $a_1 + a_2$ implies a_3, unless a_2 of a relevant conformant style structure of $a_1 + a_2$ implying b interferes. (Narmour, 1990, p. 70)

---

[3] As an example, consider the three-note stepwise ascent C-D-E. Narmour would bracket these notes together as an instance of process, labeled [P]. The label signifies that the interval C-D created an implication for continued stepwise ascent that the E fulfilled. Thus, the label says something about E's expectedness—namely, that it was highly expected—but it says nothing about D's expectedness in light of the preceding C, or about C's expectedness in light of whatever preceded it. The present model assesses the expectedness of every event, and avoids the gapped quality of the analyses in the I-R model, where expectedness is sometimes assessed only for every third note.

[4] It should be noted that the model directly considers only pitch and time in the production of expectancy ratings. Other parameters most likely relate—particularly to the resultant melodic tensions. For example, a performer might choose to compound a tense event with an additionally unexpected dynamic accentuation, or, contrarily, to soften the effect by withholding dynamic emphasis. Although expectancies regarding other parameters (timbre, dynamics, etc.) probably apply, the model is limited in scope and does not address them formally. For an account that may provide a starting point for an investigation of timbre and expectation, see Tsang (2002).

According to this hypothesis, similarity implies more similarity, and differentiation implies further differentiation. Configurations of implication fulfillment and denial in various parameters are categorized into "basic structures" notated by strings of letters. [P], for example, symbolizes process, a chain of small intervals all ascending or descending. It embodies fulfillment of both intervallic and registral direction implications. [R], on the other hand, symbolizes reversal, a large interval followed by a small one in the opposite direction, again constituting realization of both intervallic and registral tendencies. Structures may be characterized by a mix of fulfillment and denial. [IP], for example, designates a case where small intervals follow one another in opposite directions, embodying fulfillment of intervallic implications, but denial of registral ones. Narmour's melodic analyses are made up of chains of these symbols, often accompanied by annotations, in smaller type, of metric, harmonic, dynamic, and other extramelodic factors contributing to the segmentation and analysis.

The theory's conceptual background is problematic first in its foundational reliance on Gestalt principles (questioned for their dependence on a priori notions of "good" and "best"—vague and definition-resistant terms), and second in its seemingly unjustified assertion that small intervals embody similarity and large intervals embody differentiation. Building on this categorization, Narmour postulates that small intervals imply continuation to small intervals (a case of similarity implying similarity), but large intervals imply reversal to small intervals (a case of differentiation implying differentiation). Several questions arise: although pitches separated by a small interval are nearer to one another in frequency, and could in this respect be viewed as more similar than distant ones, do pitches a semitone apart really seem more similar than pitches a fifth or an octave apart (pitches more closely related in a tonal sense)? If large intervals, as an instantiation of differentiation, imply further differentiation, should they not imply more large intervals, since large intervals by definition represent dissimilarity? Why do the principles apply only to interval and registral direction? According to the theory's rules, should dynamic and rhythmic differentiation not suggest further differentiation as well?

Although the theoretical foundation seems open to question, Schellenberg (1996, 1997) showed that parts of the I-R theory could be excised without damage to the model's predictive power. For example, Schellenberg observes that the notion of small intervals implying small intervals and large intervals implying small intervals can be reduced to the idea that pitches imply continuation to other nearby pitches—proximity—a principle that has garnered empirical support (Schellenberg, 1996, 1997).

Narmour's basic hypothesis, quoted above, embodies another of the theory's core tensions. On the one hand, the I-R theory seeks to describe robust, innate, universal perceptual processes. On the other hand, it seeks to acknowledge the effect that learned schemata, stylistic idiosyncrasies, and other parameters (rhythm, harmony, etc.) can have on implication formation. The result is a highly rigid and precise theory that—in analytic practice—relies almost entirely on except clauses.

Presumably, expectancy interests theorists because of its potential role in the generation of affective experience in listening. Yet the I-R theory addresses affect only occasionally and in passing. When the subject is broached, musical experience sometimes seems to be depicted as an implausible ricochet between shock (thought to arise when expectations are denied) and satisfaction (thought to arise when expectations are fulfilled).

The notational system of the I-R theory remains relatively opaque with regard to expectational content. The symbology consists primarily of chains of letters that designate "basic structures." The structures are categorized, however, not according to their degree of realization or denial, but according to their component parts (interval of this size in this direction followed by one of that size in that direction). The method often causes raw taxonomy to supersede and obscure expectational content in analyses and discussion. Although a reader steeped in Narmourian classification might understand, at the sight of an [IP], that partial realization has occurred, a casual reader would probably have to flip to the glossary of symbols in the back and reread the entry for intervallic process to find, couched among other points regarding the sequence of intervals and their direction, the summary of [IP]'s realizational status.

Because the symbols do not speak directly to the implications and realizations they purport to describe, they necessarily diffuse some of the theory's focus away from its chief aim: an examination of expectancies. Narmour's inclusion, as Appendix 5 of his first book, of a catalog of the 200-odd possible combinations of [IR]s, [D]s, and other structures that may occur in melodies reflects a turn away from expectancy and toward categorization—a turn promoted by the notation. Indeed, at times, Narmour (1990, p. xiv) restates his aim in terms of categorization: "Analytically, I characterize the whole journey as an explication of the 'genetic code' of melody, with the aim of discovering a consistent taxonomy of structural types." Narmour ultimately succeeds more in this endeavor than in the explication of the dynamic expectations that characterize a listener's experience.

Key connections between the implication-realization theory and the present model are the commitment to an account of purely schematic expectations, and the centrality of notions of distance (related to intervallic size in Narmour's model and proximity in the present one) and direction. Key distinctions include the incorporation of tonal relatedness (couched as stability within the present model, but excluded from the core of Narmour's account), the treatment of hierarchy (explicitly formalized in the present model but not in Narmour's) and of extramelodic factors such as meter and harmony (allowed to enter Narmour's model in informal ways, but only permitted to enter the present model in formalized ones), the symbolic representation (category letters in the I-R model, and numeric predictions of expectedness in the present one), the theorized connections to affect (relatively sparse and informal in the I-R model, but relatively extensive and systematized in the present one), the centrality of the concept of closure (essential in the I-R model, but absent in the present one), the theoretical background (relying on Gestalt notions vs. not), and the analytic focus (categorization vs. the tracking of expectancies).[5]

MELODIC ATTRACTIONS

In contrast to Narmour's approach of excluding tonal relationships to focus on interval size and direction, Lerdahl (2001) presents a model of tension and expectation based primarily on scale degree function. Given a context chord and key (determined by event governance rules), pitch classes are assigned an anchoring strength, which represents the degree to which that pitch can attract surrounding ones. For example, in a I of C major context, the root C receives anchoring strength 4, the other members of the tonic triad (E and G) anchoring strength 3, other diatonic pitches (D, F, A, and B) anchoring strength 2, and all chromatic pitches anchoring strength 1. According to the theory, pitches tend to be attracted to the nearest pitch with a higher anchoring strength (i.e., A, with anchoring strength 2, is attracted to neighbor G, with anchoring strength 3). "Tension," in this account, corresponds to instability.

Lerdahl quantifies the tendency of one pitch to resolve to another by dividing the anchoring strength of the potential attractor by the anchoring strength of the pitch under consideration, and multiplying the result by the inverse square of the semitone distance between them. For example, A would be attracted to neighbor G by $(3/2) \times (1/4)$ or .375.

Unlike the I-R theory, Lerdahl's model explicitly addresses the degree to which each melodic event is theorized to be expected (in his terminology, the degree to which the present event attracted the preceding one). Additionally, the predictions are presented in transparent notation (graphs of attractional fluctuations positioned across the top of scores) and with explicit references to a proposed effect on experienced tension. However, the model does not incorporate direction, an omission that results in somewhat inflexible analyses. (In a C-major context, no directional con-

---

[5] For a more comprehensive comparison between the implication-realization theory and the present model, see Chapter 3 of Margulis (2003).

text can make F more attracted to G than E). Additionally, its quantification—particularly the specification of proximity as the inverse square of semitone distance—causes problems. Because the semitone distance between a pitch and its repetition is 0, the model misleadingly predicts the expectancy for pitch repetition to be infinite. Moreover, the formula privileges movement by semitone to an inordinate degree, assigning huge attractional ratings to continuations by semitone and comparatively minuscule ratings to other possibilities. But experiments have suggested (Lake, 1987) that in many circumstances listeners expect a relatively broad range of continuations—not exclusively movement by semitone.

The present model takes the core of its event governance rules for the parameter of stability from Lerdahl's theory of melodic attractions. It also adopts the general approach of assigning a numeric prediction to each melodic event—in the case of Lerdahl, the prediction is of attraction, in the case of the present model, the prediction is of expectedness. It uses Lerdahl's approach of linking predicted values to tension and graphing the tension fluctuations across the top of the score. However, tensions in the present model pertain exclusively to melody and are separated into three different types; where Lerdahl's model associates tension with instability, the present model proposes three tension types that arise out of all aspects of expectancy (not just stability). Also in contrast with Lerdahl's theory of melodic attractions (a small part of a large study of pitch space), the present model attempts a thorough exposition of melodic expectations, including an extended account of direction and proximity, and a system for addressing hierarchic expectancies. As already discussed with respect to proximity, the quantified predictions of the present model differ substantially from Lerdahl's.

MELODIC EXPECTATION

In comparison to these studies, a primary contribution of the current paper is the integration of the two frameworks into a single approach. One interpretation of Peretz and Coltheart (2003) suggests that tonal encoding might operate modularly, relying on principles of tonal organization learned implicitly and automatically in reaction to early exposure. Stability, therefore, might be considered to act in conjunction with other, more Narmourian aspects of melodic projection. In comparison to the I-R theory, the present model is more specific about the type of expectations under consideration—deeply schematic, automatic primings. It is also more consistent about pursuing its predictions on their own terms, without allusion to the contradicting effect of other expectancy types. The present model was constructed around a theory of the way in which schematic expectancy connects to musical affect, so expectancy predictions always connect explicitly to notions of melodic tension. Via the parameter of mobility, the present model successfully addresses the pitch repetition issues that are problematic in Lerdahl (2001). Finally, in the form of the hierarchy parameter, the present model features a formalized approach to expectancies that are deeply schematic and automatic but incorporate more context then just the two immediately preceding notes.

The next sections describe the model parameters—stability, proximity, direction, mobility, and their hierarchic implementation—one at a time.

Stability

Stability captures the intuition that, in general, listeners expect relatively stable melodic events. Tonal stability is a central concept in both music theory (cf. Lerdahl & Jackendoff, 1983) and music cognition (cf. Krumhansl, 1990 and Bharucha, 1996). Its potential role in expectancy has been articulated by Larson (1993) and Lerdahl (2001), but has not been treated within the I-R model. One aim of the present model is to integrate within a single framework the expectational approaches suggested by existing theories.

Events are experienced differently in different tonal contexts.[6] C, for example, is maximally stable in a context based on I of C major, but minimally stable in one based on I in F♯ major. With slight adaptations, the model adopts Lerdahl's (2001) event governance rules to select an operative chord and key in which stability may be evaluated. The bottom half of Table 1 outlines the model's event governance rules. Figure 1 illustrates their operation in sample contexts. The boxes in the left column feature nonchord tones, which shift the chord context, triggering an expectancy for the dissonance to resolve to a chord member. The boxes in the bottom row feature secondary chords, which shift the key context, triggering an expectancy for movement to members of the tonicized key. Only the upper right box, featuring a chord tone over a diatonic harmony, generates an expectancy in the default context of I in C major.

The top half of Table 1 presents the model's stability ratings for events within a given chord and key context. For major key contexts, it is clear what pitches count as diatonic. For minor key contexts, with their alternate versions of $\hat{6}$ and $\hat{7}$, some clarification is needed. Both versions of $\hat{6}$ and $\hat{7}$ are considered diatonic until a particular version (e.g. raised $\hat{6}$) occurs in the melody. An occurrence of the raised version of either $\hat{6}$ or $\hat{7}$

---

[6] The stability parameter requires a governing chord and key context. Atonal or non-Western contexts may not allow for the extraction of such a context. For a listener not versed in these styles, the other expectational parameters (i.e., proximity, direction) might supersede. For a listener who had been immersed in these styles, the stability parameter might have developed with different specifications appropriately attuned to the relevant musical vocabularies.

TABLE 1
Event Governance Rules Select the Tonal Context (a Chord and Key) in Which an Event Is Evaluated

Stability Rating	Pitches in Context of a Chord and Key
6	Chord root (and, after a seventh in the melody, the pitch one diatonic step down from it)
5	Chord third and fifth
4	Other diatonic pitches
2	Chromatic pitches

Event Governance Rules

Context is I in current key, EXCEPT…

When a secondary chord occurs, the context shifts to I in the tonicized key.
When a melody note constitutes a nonchord tone with respect to the current harmony, the context shifts to the current chord in the current key.
When a melody note constitutes the seventh of the current chord, its lower diatonic neighbor is promoted to the highest stability rating.
When a strong predominant chord (such as an augmented sixth or Neapolitan) occurs, the context shifts to V in the current key.

NOTE—The default context is I in the current key, except in the cases listed. Within the chord and key context established by the event governance rules, stability ratings are distributed as listed among the root of the context chord, its remaining members, pitches diatonic to the context key, and pitches foreign to it. Lerdahl (2001) also distributes pitches according to the categories of root, third/fifth, diatonic, and chromatic, but uses a different quantification. The event governance rules are based on Lerdahl (2001), with added treatments for chord sevenths and strong predominants.

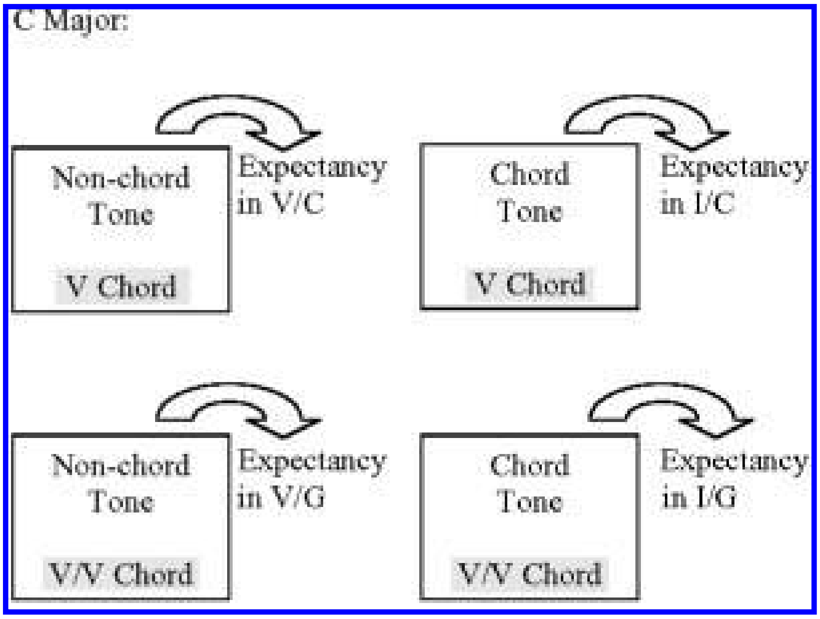


Fig. 1. The event governance rules establish the tonal context for expectancies about the next event. If a nonchord tone occurred over a dominant harmony in C major, the context key would remain C, but the context chord would shift to V. From the standpoint of stability, the next event would be expected to resolve the preceding one to a chord tone of V. Three other scenarios are depicted. Stability ratings, depicted in Table 1, depend on the event governance rules' assignment of a tonal context.

selects the raised version of both scale degrees as diatonic for the expectancy targeting the immediately ensuing event (or, at higher levels, immediately ensuing span, as discussed later). The lowered versions of each are then consigned to chromatic status. Exactly the reverse occurs following the sounding of the lowered version of either: both lowered versions are then considered diatonic and both raised versions, chromatic.

Each model factor assigns a rating to each melodic event; stability ratings are ultimately combined with ratings from the other factors to produce overall expectancy ratings for each melodic event. Generally, expectancy ratings represent the predicted expectedness of melodic events, with high ratings signifying high expectedness. The component parts of the overall rating—the stability rating, the proximity rating, and so on—represent the portion of expectedness attributable to the individual factor. A stability rating, in other words, captures the degree to which a melodic event is predicted to be primed solely on the basis of its position within the governing tonal context.

Regarding the quantification of the model, it is important to note that numeric values are invoked simply to manage the complexity of the predictions. It is not the precise value assigned to each event that is important, but rather the position in which that value situates the event with respect to surrounding events and other possible events. Quantification allows for the relative ordering of expectancy predictions for the otherwise unmanageably many possible melodic contexts. Additionally, it permits a specificity in the model's claims that makes empirical investigation easier.

The quantification of the model was established secondarily by interpreting the results of existing experiments (continuation-sing-or-rate paradigms) as reflections of an optimal mixture of expectedness and unexpectedness, and primarily by consulting intuition to try to trace sensations of tension to the originating expectancies. The details of the precise quantification of the expectancies are not themselves as important as the framework for future inquiry that the structure of the model provides. Experimental study can establish the best specification of each parameter. But such study could not occur in the absence of a conceptual frame that linked structure to experience with explicit predictions. This model aims to provide that conceptual frame.

Consider the stability ratings. In the default I of C-major context, C receives a stability rating of 6, signifying strong expectedness. E and G receive stability ratings of 5, D, F, A, and B receive stability ratings of 4, and the remaining chromatic pitches receive stability ratings of 2. The particularly low ratings for chromatic pitches reflect the experimental finding (Lake, 1987) that diatonic events are much more strongly expected that nondiatonic ones. However, the primary motivation behind this and other assignments of specific numerical ratings is a consideration of the experiences of tension predicted to arise from them.

As described later in this article, unexpectedness is theorized to create a type of tension (termed surprise-tension) that registers in the listener as an experience of intensity or dynamism. Intuitive reflection reveals that pitches foreign to the governing key (i.e., chromatic pitches) generate a special intensity, more pronounced in comparison to that generated by diatonic pitches than the intensity produced by the same diatonic pitches seems in comparison to that produced by members of the governing chord. Accordingly, chromatic pitches receive a rating that is much less than that given to diatonic pitches, but diatonic pitches receive a rating that is only moderately smaller than that given to members of the governing triad. On account of the inverse relationship between expectancy and surprise-tension discussed later, chromatic pitches' particularly low expectancy ratings translate into particularly high amounts of surprise-tension.

These sorts of concerns determined the general relationship among the different stability ratings. The specific numeric values chosen reflect a consideration of the effect the numbers would have when combined with the values assigned by other parameters (proximity, etc.) to produce overall ratings. Working backward from experiences of melodic tension, the numbers were set at a level that in combination with other factors produced the desired tension predictions. As discussed later, the combined formula for expectancy ratings makes it easy to list the core ratings for events after any pitch in any key. Because these ratings are correlated with several tension types, scrutiny of intuitive experience in multiple situations that shared melodic successions allowed the predictions to be weighed, and suggested, on occasion, reorderings or reassignments of parameter values. In this way, the model was built downward from experience to prediction. It is reasonable to assume that empirical study across wider populations will further modify the values.

Proximity

Proximity captures the intuition that listeners expect subsequent events to be relatively proximate to preceding ones. The I-R model theorizes that listeners expect small intervals to continue to additional small intervals in the established direction and large intervals to reverse to small intervals in the opposite direction. This amounts in part to theorizing that listeners expect subsequent events to be proximate to preceding ones.

Schellenberg (1997) suggests that the principle of proximity underlies a large part of listener expectations. Table 2 presents the model's proximity ratings, which are distributed differently than the proximity ratings in Schellenberg (1997) and in other models. Several features make this quan-

TABLE 2

Proximity Ratings Increase as Semitone Distance Increases, Reflecting the Expectation That Pitches Will Proceed To Relatively Nearby Continuations

Pitch Distance in Semitones	Proximity Rating
1	36
2	32
3	25
4	20
5	16
6	12
7	9
8	6
9	4
10	2
11	1
12	0.25
13	0.02
>14	0.01

tification an improvement over those produced by more elegant formulas, such as the simple inverse and inverse square of semitone distance. First, the ratings for pitches less than an octave distant employ no decimals and are easily combined with other factors. Second, the ratings decrease less (4) from 1 to 2 semitones than they do from 2 to 3 semitones (7). The placement of the steepest drop between 2 and 3 semitones rather than between 1 and 2 semitones reflects a preference for stepwise motion affirmed in counterpoint textbooks (Kennan, 1998), statistical surveys of Western and world music (Vos & Troost, 1989), and psychological studies (Bregman, 1990; Schellenberg, 1996). In Western tonal music, as well as in many other styles, steps are composed of 1 or 2 semitones. Composers use a preponderance of steps, and listeners almost always, on an expectancy scale, rate stepwise continuations higher than leaping ones (Schellenberg, 1996). Steps as a class are clearly more different from leaps as a class than semitone steps are from whole-tone steps. The proximity ratings reflect that perceptual distinction. They avoid the undesirably high ratings given to semitones by Lerdahl's inverse-square rule (2001) and better follow empirical studies (Lake, 1987; Schmuckler, 1989), suggesting that listeners expect a variety of possible continuations with comparable strength and do not radically favor semitone neighbors.

After 3 semitones, proximity ratings drop increasingly gradually until, for very large intervals, expectancy ratings are equally low. The model predicts that a pitch 3 octaves away and a pitch 2 octaves away are expected by about the same amount, but a pitch 1 octave away is much less expected than a pitch 1 semitone away. At a smaller scale, the same principle applies. The model predicts that a pitch 10 semitones away is expected slightly less than a pitch 7 semitones away, but a pitch 6 semitones away is expected considerably less than a pitch 3 semitones away.

The specific increments between the proximity ratings were selected, like the increments between the stability ratings, with end tension predictions in mind. In the case of proximity, the potential impact of direction was especially attended to. As the formula (discussed later) for combining individual factor ratings reveals, stability and proximity ratings are multiplied and direction ratings are added. Thus after a given pitch, say D in C major (see Table 4), the stability- and proximity-based rating for virtual (possible future) continuations can be listed (first C, than E, etc.). D's directional context then shifts and reorders the ratings. For example, if D were leapt to from a much higher pitch, the directional impulse for reversal might add enough impetus for E to surpass C as the most highly rated continuation. By considering the tension created by multiple musical examples in which specific scale degrees were leapt—or stepped—to, the desired interaction between proximity and direction for each possible configuration was determined. Numeric values were chosen so that direction reordered ratings (placing E before C after D) when intuitive experience warranted it, but preserved the default ranking (e.g., C before E after D) when appropriate (when, for example, a progression from D to E instead of C created more surprise-tension).

Direction

Support for Narmour's (1990) idea that small intervals imply directional continuation but large intervals imply directional reversal has been mixed. Schellenberg (1996, 1997) suggests that the expectation for reversal after large intervals may be considerably stronger than the expectation for continuation after small ones. Figure 2 depicts the present model's distribution of direction ratings.

Within the category of small intervals, the more exaggeratedly small the interval is, the stronger the expectation for continuation is proposed to be. Likewise, within the category of large intervals, the more exaggeratedly large the interval, the stronger the expectation for reversal is proposed to be, until a cutoff point beyond which expectations are theorized to be equivalently strong. (Extremely large intervals might suggest a melody that is processed as a polyphonic combination of two component melodies). Between the categories of small and large, however, it can be seen that reversal expectations after even moderately large intervals are theorized to grow much stronger than continuation expectations after the smallest intervals.

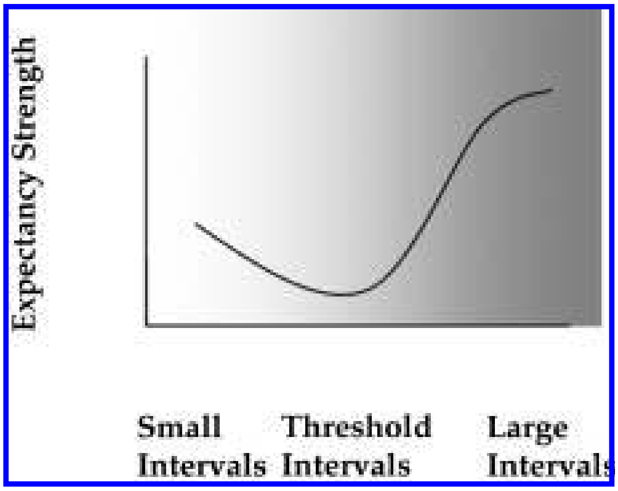


Fig. 2. The *x* axis represents interval size, and the *y* axis represents strength of directional expectation. Dark shading represents an expectation for reversal, and light shading indicates an expectation for continuation. The influence of the direction parameter varies according to the size of the interval between the two preceding notes. Very large intervals are theorized to generate a strong expectation for reversal, but very small intervals are theorized to generate only a moderate expectation for continuation.

Table 3 shows the specific direction ratings after differently sized intervals, ratings calculated in consideration of the proximity ratings to produce tension predictions that matched experience. After a pitch repetition (interval size 0, for example the succession G-G), direction ratings weakly favor continued lateral motion. Ratings after other intervals favor

TABLE 3
Ratings Given To Events That Lie in the Direction Implied by the Interval Between the Preceding Two Notes

Interval Size	Direction Rating
0	6 for continuation
1	20 for continuation
2	12 for continuation
3	6 for continuation
4	0
5	6 for reversal
6	12 for reversal
7	25 for reversal
8	36 for reversal
9	52 for reversal
≥10	75 for reversal

NOTE—Events lying in the opposite direction receive a rating of 0. For example, after an ascending leap of a minor sixth (8 semitones), 36 is given to any event that constitutes a descent, 0 to anything that constitutes an ascent. (These values are different than those of later quantifications of the registral direction parameter in Narmour's implication-realization theory, cf. Schellenberg, 1996, 1997.)

either continuation or reversal. After an ascending interval, an event qualifies as a continuation if it lies above the last pitch and a reversal if it lies below it. After a descending interval, an event qualifies as a continuation if it lies below the last pitch, and a reversal if it lies above it. Narmour (1990) suggests that after a large interval pitch repetition should also qualify as reversal because of the differentiation between the initial ascending or descending motion and the ensuing lateral motion. According to this notion, a repetition of A should count as a partial reversal after the ascending leap C-A. In such cases, the absence of directional continuation seems enough to qualify the repetition as a partial reversal. In the present model, pitch repetitions fulfill directional implications at only 1/3 of their full value (rounded to the nearest whole number). For example, after the ascent C-A, a descent to neighboring G would receive a directional rating of 52, but repetition of the A would receive a directional rating of 17, one third of 52.

Mobility

The secondary factor mobility captures the general intuition that a melody will move. Expectancies for self-repetition (the expectancy, for example, that D will continue to D) have caused problems in previous models. Lerdahl (2001) excludes them from consideration, since his proximity rule (the inverse square of semitone distance) implies an infinite expectation for self-repetition (where semitone distance is 0). The difficulties stem from repetition's in-between status; it counts neither as a continuation of ascending or descending motion nor as a reversal of it, and it counts neither as movement to a proximate or nonproximate pitch, but rather as something different. Before the inclusion of mobility, the present model predicted dubiously strong and robust expectations for repetition. Mobility tempers these predictions by capturing the general expectation that a melody will move. The principle operates within the model by lowering expectation ratings for self-repetitions. Specifically, mobility works by multiplying the stability and proximity rating for repetition by the constant 2/3, to penalize possible continuations that do not move to a new pitch level. In all other cases, mobility is 1 and does not alter the overall expectancy rating.

Factor Combination

Before hierarchic levels are considered, an event's expectancy rating may be generated by using the following formula.

EXPECTANCY FORMULA (BEFORE HIERARCHIC LEVELS ARE INCLUDED)

A pitch x is expected to follow pitch y by amount z:

$$z = (s \times p \times m) + d,$$

where s = the stability rating of x (see Table 1), p = the proximity rating of x (see Table 2), d = the direction rating of x (see Table 3), and m = the mobility rating of x (2/3 if x repeats y and 1 in all other cases).

All factors that can apply, do, all that cannot, do not. For example, expectations concerning the second note of a piece use only s and p because no direction has been established. Consider the melody in Figure 3. Expectancy ratings could be generated for each event. Take the G in measure 1 as an example. In the default I of C major context, it receives a stability rating of 5. Two semitones away from the preceding A, it receives a proximity rating of 32. G does not constitute a repetition of A, and therefore receives a value of 1 (rather than 2/3) for mobility. Directionally the G satisfies the expectation generated by the preceding 9-semitone leap from C to A. Before hierarchic levels are considered, G's total expectedness (212) is calculated by multiplying 5, 32, and 1, its stability, proximity, and mobility ratings, and adding 52, the reversal rating generated by the leap. As a contrasting example, take the F in measure 2. Its stability rating is 4, and its proximity rating (lying 5 semitones away from the preceding C) is 16. Again, its mobility rating is 1. Its directional rating is 0 because it does not lie in the direction predicted by the preceding descending step. Its total rating before hierarchic levels are considered, therefore, is (4 × 16 × 1) + 0, or 64.

A model overview can be provided by listing the core stability and proximity ratings after different events in a given tonal context. Such listings are valuable because they make general ratings trends apparent. By studying the distance between top-rated events, it can easily be seen how different directional contexts would affect the ratings. Table 4 depicts some sample core stability/proximity ratings in an I of C major context. Although a pitch's theorized expectedness is calculated by combining all factors, examining the stability/proximity ratings in isolation gives an

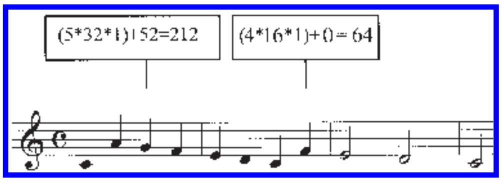


Fig. 3. A melody, for every note of which an expectancy rating could be calculated. The expectancy ratings signify the extent to which the continuation is predicted to be primed. Sample rating calculations are included for the G in measure 1 and the F in measure 2.

TABLE 4
Sample Core Expectancy Ratings in an I of C Major Context

	216	**192**	**180**	**160**	**150**	**144**	**128**	**125**	**120**	**100**	**96**	**80**
D		C		E						B F	D	G
B	C						A			D G	B	E
F		E	G							D	C F	A
G							F A	E	G		C	B

NOTE—The leftmost column represents a pitch in C major. Subsequent columns show the combined (multiplied) stability and proximity ratings of pitches found within the span of one tritone below to one tritone above. For example, in the row for D, C refers to the C one step below, because it is the closest instance of C, and the only one within the tritone-to-tritone span. F refers to the F a minor third above D for the same reasons. As described in the text, the direction factor can reorder the listings presented here.

excellent sense for the inner workings of the model; provided a tonal context, the ratings for possible continuations after any pitch can be concisely listed. Because direction depends not only on the identity of the previously heard pitch, but also on the one before it (i.e., the interval created by the two), if it were included on the table, the range of possible implicative situations—even in a single tonal context—would become impracticably large; in C major, for example, ratings would have to be listed not just after C, but after C when it followed B, after C when it followed D, and so on. Instead, a much more intuitive sense of the model's predictions can be gained by viewing a restricted list of stability/proximity ratings and imagining the shifts that different directional circumstances create.

Contrast expectancy ratings after B, shown in row 2, with expectancy ratings after G, shown in row 4. Leading-tone B is theorized to trigger a maximal expectation for the tonic, semitone-neighbor C. The next most expected continuation, A, receives a dramatically lower rating. Because the highest possible direction rating is 75, and A's rating is 88 less than C's, no directional context can cause A to surpass C as the most expected continuation after B. Consider row 4, however. When only stability and proximity are considered, diatonic neighbors F and A tie as the most expected continuation after G. Following $\hat{5}$, directional context is always the determining factor in selecting $\hat{4}$ or $\hat{6}$ as the most expected continuation. Aside from the distinction in degree of susceptibility to directional influence, expectations after B and G differ in another important way. The most expected continuations after G receive a rating of 128, but the most expected continuation after B receives a rating of 216. The discrepancy indicates that expectations after B are strong, specific, and robust, but that expectations after G are weak and diffuse. Situations in between these two extremes can be seen in rows 1 and 3, which list expectancy ratings after D and F in a default C-major context.

Hierarchy

Even deeply schematic expectancies depend not only on the last heard note, or the last heard pair of notes, but also on a larger context. Hierarchic rules drawn from principles in Lerdahl and Jackendoff (1983) segment the music into time spans, choose a "head," the most structurally significant event, for each span, and specify the head's influence on melodic expectations. More important melodic events are predicted to play a correspondingly larger role in expectancy formation.

Time spans can "be thought of as apprehended rhythmic units in terms of which pitch structure is heard" (Lerdahl & Jackendoff, 1983, p. 124). Since many of the segmentation rules are couched in terms of preference rules, some room for disagreement over the selection of a head exists. Such issues could be resolved by formalizing the preference rules more rigidly, but the task lies outside the scope of this essay. The present theory does introduce one additional preference rule, to supplement Lerdahl and Jackendoff's nine rules for time-span reduction. The additional rule is meant to resolve ties in the case of more than one plausible head choice. Called "Expectancy Fulfillment," it reads: *Of the possible choices for head of a time-span* T, *weakly prefer a choice that fulfills expectations generated by* e, *the head of the time-span immediately preceding* T *at the same level.* Since expectancies usually privilege stepwise motion, with the highest rated continuations lying a whole or half step from the previous pitch, the expectancy fulfillment rule captures a listener's attempt to form stepwise progressions beneath the musical surface. Other time-span reduction rules specify, for example, to prefer the selection of a head that is metrically strong, intrinsically consonant, relatively closely related to the tonic, and, in the case of melody, relatively higher in terms of register (Lerdahl & Jackendoff, 1983).[7]

According to the theory proposed here, listeners treat the heads of the time spans at a particular level as a sort of background melody about which expectations pertain. The principles of proximity, stability, mobility, and direction apply to the background melody just as they apply to the notated one. Expectations about the level of the notated melody operate simultaneously alongside expectations about background levels, for example quarter-note, half-note, and whole-note levels. In other words, a listener is theorized to have a sense of not only what might happen next

---

[7] For a full listing of the time-span reduction rules, see Lerdahl and Jackendoff (1983).

locally, but also what might happen next slightly more globally. The notated melody is sometimes referred to in this article as the note-to-note level, to emphasize that it is composed of notes rather than heads. However, "notes" in this reading are entities within a tonal network, and event governance rules influence how stability is assessed within that network.

Figure 4 shows the time-span reduction for measures 1–4 of Mozart's Piano Sonata in E♭ major, with the score on the bottom line and reduction levels on the lines above it—from quarter-note to half-note, measure, and two measure. Although only the melodies are analyzed, the full scores of musical excerpts are included in this article. The harmonies influence the selection of chord and key contexts, which contribute to the stability ratings. Elements outside the melody can also influence the formation of time-spans and the selection of heads, important in the hierarchic implementation of the parameters of stability, proximity, and direction. At a given level, notes depicted with the appropriate rhythmic value (i.e., quarter note at the quarter-note level, half note at the half-note level) represent the time-span heads: this notation is consistent with that used in Lerdahl and Jackendoff (1983). The unstemmed noteheads (a new notational element) result from the implementation of a final-state theory within a temporally situated context. They represent considered but ultimately discarded head candidates. At the start of a span, a listener cannot know the identity of the span's best head. Only the best candidate in the part of the span already heard can be known. The idea is that the listener assumes the best candidate from among the pitches already heard to be head until further evidence accrues. Consider the start of measure 3. For lack of better options, C is initially assumed to be head (observe the unstemmed notehead representing C at the quarter-note level). B♭ soon supplants it, forming the actual head of that quarter-note span (observe the stemmed B♭ on the same level). Actual heads are notated differently than merely considered ones because they ultimately play a role in the formation of expectancies about future spans, as the next example explains.

Observe that the C at the start of measure 3, dissonant with the underlying harmony, is a highly unsatisfactory head candidate. Nevertheless, at the time point before any other notes have yet occurred within the span, it is the only candidate. The theory holds that listeners represent as head the best candidate from among those already heard. Accordingly, it is predicted that listeners will often temporarily entertain the possibility of a poor head, especially at the start of a span, where often the only notes that have yet sounded are nonchord tones or similarly inferior options. As soon as something more satisfactory occurs, the listener supplants his or her representation of the span's head with the better choice.

Note that in Lerdahl and Jackendoff's use of this notation, the final-state assessment of a pitch's structural importance is represented by the

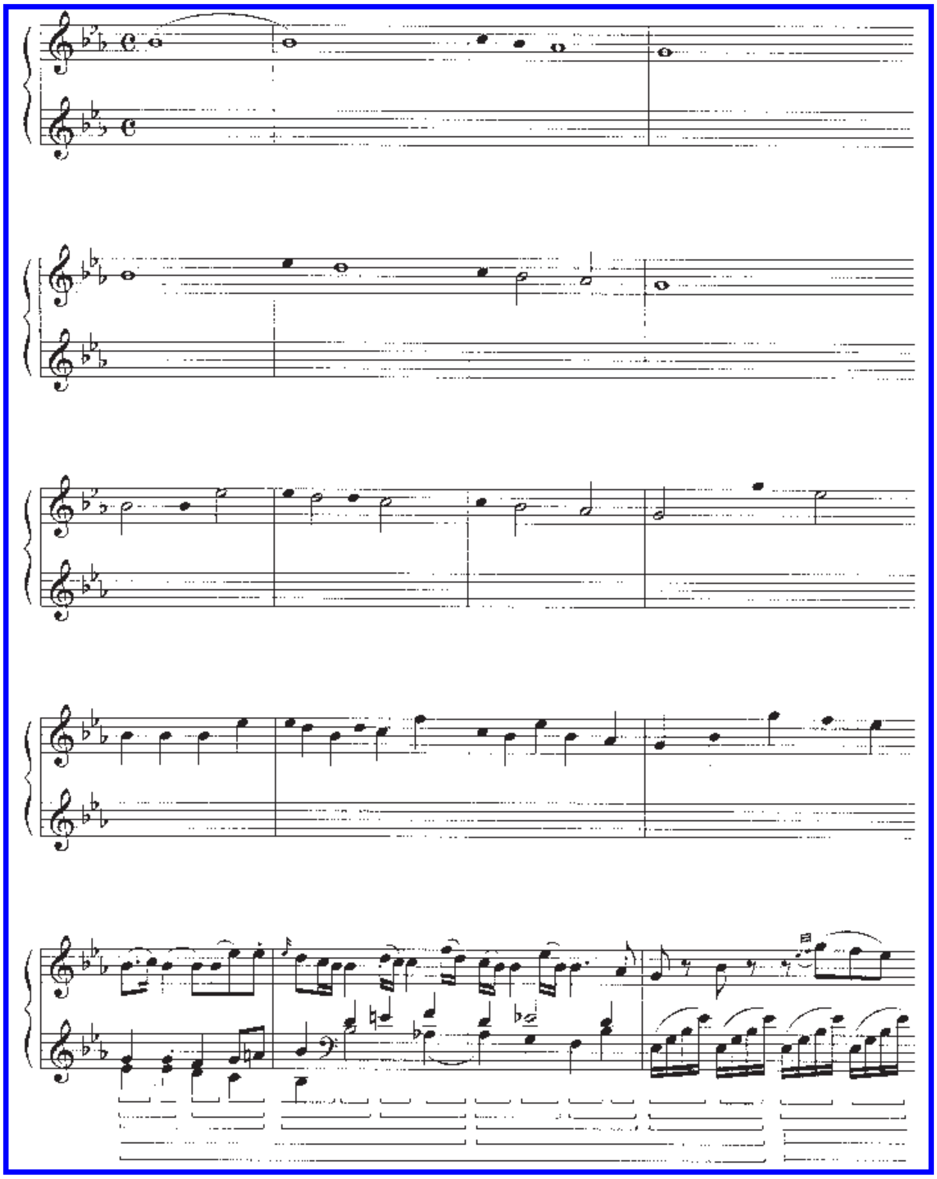



Fig. 4. Time-span reduction for measures 1–4 of Mozart, K. 282. Quarter-note, half-note, measure, and two-measure levels are depicted on separate staves. Unstemmed noteheads represent considered but ultimately discarded head candidates. Expectancies project at each level, targeting adjacent heads the way expectancies at the note-to-note level target adjacent pitches.

highest level at which that pitch is retained; the higher the level of retention, the more structurally important the pitch. However, in the present use of the notation, no final-state assessment is implied. The inclusion as a stemless notehead of the C in measure 3 at not only the quarter-note, but also the half-note, measure, and two-measure levels does not indicate that the C has been retrospectively judged to function structurally at all of these levels. Rather, it indicates that measure 3 initiates spans at not only the quarter-note level, but also all higher ones through the two-measure level. Listeners are theorized to simultaneously project continuations at all these levels of the notated hierarchy, and for the first sixteenth note of all of the spans starting at the beginning of measure 3, they are theorized to temporarily sustain C as the only available head candidate. In other words, C's notational presence on multiple levels at the start of measure 3 says something not about a final-state assessment of the pitch, but rather about a theorized projectional and evaluative experience during one sixteenth note's duration at the start of measure 3.

At the note-to-note level, expectations about the downbeat of measure 3 (the C) are formed largely by the preceding two notes, F and D. At the quarter-note level, expectations about that same downbeat are formed largely by the preceding two heads, C and F. At the half-note level, expectations about the same downbeat are most afected by the identity of the previous two heads at that level, C and D. The C at the start of measure 3 is more expected at some levels than at others. The way in which its expectancy ratings at different levels are combined to form an overall expectancy rating is discussed shortly.

Lerdahl and Jackendoff use time-span reduction as a derivational step toward prolongational reduction. For the purposes of the current model, time-span reduction suffices. It avoids the complications of prolongational reduction, which depends more substantially on final-state awareness and resists a left-to-right temporal implementation. On its own, time-span reduction effectively selects the most expectation-influencing pitch from within individual spans. In Lerdahl and Jackendoff, prolongational reduction provides the connection to experience (position in the resultant tree is theorized to link to tension), but in the present model, the connection to experience is made via expectancy ratings that evolve as the piece progresses, obviating a final-state tree structure.

Expectations cannot apply to levels with arbitrarily long time spans. As Justin London (2002) remarks, "the constraint on the scope of larger temporal patterns is correlated with our sense of the psychological present" (p. 536), a period whose upper limit researchers have variously placed between 5 and 10 seconds. London locates another important threshold between 1.5 and 2 s, beyond which rhythmic synchronization becomes difficult or impossible. These values suggest a scaling of expectancy ratings, according to which ratings for levels with spans longer than 2 s diminish in importance.

Although at the shallowest level expectancy ratings measure the note-to-note fluctuation of tension, hierarchic ratings measure a sort of background tension. Consider Figure 5, which presents a melody followed by

Fig. 5. A melody consisting of a segment that is sequenced an octave higher.

a sequence of it an octave higher. If only note-to-note expectations were represented, the profiles for each statement, excepting the first two notes, would be identical. Yet the second statement (starting in m. 4) possesses an elevated degree of tension when compared with the first one, because of its registral distance and accordingly low background expectancy. Hierarchic expectancy ratings register the background proximity violation and assign a low value to the second statement, which, when combined with other levels' ratings, produces slightly lower overall expectancy ratings for events within the second statement.

The method for computing and combining hierarchic expectancy ratings is presented here.

COMPLETE EXPECTANCY FORMULA[8]

1. At a given level, the sequence of heads of each of that level's time spans forms a background melody.
2. The time-span segmentation, well-formedness, and preference rules determine the head of each time-span at a given level.
    a. At any point, the best candidate for head is selected from the time span's already heard pitches. A listener may consider several different head possibilities over the course of a span.
3. Expectancy ratings at the note-to-note level are calculated by applying the previously presented formula, $(s \times p \times m) + d$. Expectancy ratings at other levels are calculated by applying the same formula to events (heads and head candidates) in the background melody.
4. To determine a pitch's overall expectancy rating, produce a weighted average of the pitch's expectancy rating at each level.
    a. A pitch's rating at the shallowest level is calculated by using the expectancy formula, $(s \times p \times m) + d$.
    b. A pitch's rating at a given hierarchic level is the rating for the best yet-heard head candidate of the time span within which the pitch falls at that level.

---

[8] Peter Desain implemented the model in LISP. The program accepts representations of pitches and associated tonal contexts, produces expectancy ratings, and graphically depicts the fluctuations in tension theorized to occur across the course of the melody. A demo is planned for user access at www.nici.kun.nl/mmm under "demos."

c. Each level receives a weight reflecting its theorized salience.
   i. Note-to-note ratings receive a weight of 15.
   ii. Ratings at levels beyond the note-to-note level, up to and including levels with spans of 2 s duration, receive a weight of 5.
   iii. Ratings at levels with time-span durations from 2 s up to and including 6 s receive a weight of 2.
   iv. No levels with time-span durations longer than 6 s form.
d. Do not include hierarchic levels whose time spans occupy a duration equivalent to or less than the duration between the current and the last heard note. This rule prevents a span from receiving a weight both at the note-to-note level and a hierarchic level.

5. The formula for overall expectedness can be expressed as

$$\frac{\Sigma\, w_i[(s_i \times p_i \times m_i) + d_i]}{\Sigma\, w_i}$$

where i = the level under consideration, w_i = the weight of the level under consideration (15 for the note-to-note level, 5 for levels with spans up to 2 s, 2 for levels with spans up to 6 s), s_i = the stability rating for the pitch or head candidate at that level, p_i = the proximity rating for the pitch or head candidate at that level, m_i = the mobility rating for the pitch or head candidate at that level, and d_i = the direction rating for the pitch or head candidate at that level.

Using the weighted average of different levels' ratings ensures that expectations derived from adjacent events play a greater role than more distant hierarchic ones in the determination of overall expectancy ratings. To produce a weighted average, an event's expectancy rating at each level is multiplied by that level's weight (specified in the preceding formula). The resultant products, composed of a rating multiplied by a weight, are summed. Finally, the resultant sum is divided by the sum of the weights to produce the overall expectancy rating. The values of the weights reflect the diminished salience spans are theorized to possess over longer periods, as suggested by London (2002). They were calculated to produce ratings that matched intuitive tension profiles, but are quite susceptible to future modification. Certain factors, for example, likely cause listeners to shift their attending to deeper levels, whereas other factors most likely steer their attending toward shallower ones. These factors are not treated within this model and would most likely change the quantification of the level weights.

The formulation of the parameters' interrelationships affords the model certain conceptual and procedural advantages. Consider the fact that many melodies proceed primarily by step. The impact of direction is theorized to be minimal after small intervals. Accordingly, there are many readily available musical situations in which expectancy-based tension results primarily from factors other than direction (e.g., stability and proximity). However, there are notably fewer musical situations in which expectancy-based tension results primarily from factors other than stability, or factors other than proximity; in tonal music, there is usually a hierarchy of stability relations at play, and in most music, there is usually some distance from the preceding pitch to the current one. In other words, proximity and stability do not drop out in numerous situations the way that direction does. The configuration of the primary factors within the formula captures this dissimilarity: stability and proximity apply multiplicatively but direction is added, sometimes dramatically reordering the ratings (e.g., after a leap in a certain direction), other times leaving them fundamentally intact (e.g., after a small interval in a certain direction).

This conceptualization of factor combination made it easier to trace the tensions to their sources in the different parameters of expectancy. By looking at the many situations where direction applied weakly or not at all, predictions about the other, relatively isolated factors could be made more easily. For example, after introspecting about sensations of tension in various circumstances, it seemed clear that after the pitch F in C major, continuation to G is normally more surprising than continuation to E. It was then possible to assess the impact of direction by examining expectancies after F in situations where F was itself preceded by different notes, to determine the point at which E became a more surprising continuation than G (this might occur, for example, when F was leapt to from further above). The additivity of the parameter of direction contributed strongly to the logic and organization of this procedure. Similarly, the formulation of the model is intended to contribute to the ease of understanding its claims. For example, Table 4 lists stability- and proximity-based expectancy ratings after pitches in C major. This kind of table makes the model's predictions compact and easier to understand; if ratings based on stability, proximity, and direction were listed, the table would explode in complexity. Ratings would have to be included not just for D when it followed C, but also for D when it followed the two-note sequence B-C, the two-note sequence A-C, and the two-note sequence G-C, and so on, making it much harder to grasp the gist of the model. Given the additivity of direction, readers can examine a table based solely on stability and proximity and imagine the effect of directional context by adding the appropriate increment to events lying in a specified direction.

Effects of Deeply Schematic Expectations on Listener Experience

In broad terms, the model proposes three forms of association between expectancy and affect. One of the persistent problems afflicting expectational theories is the difficulty of correlating expectancy with anything but a unidimensional account of the clearly multidimensional phenomenon of musical affect. Although expectancy is thought to be interesting because of its impact on the generation of musical affect (Meyer, 1956), the nature of this connection remains obscure. Existing accounts tend to be impoverished, postulating a simple alternation between shock (in the case of unexpected events) and contentment (in the case of expected ones). Yet, as literary theorist Charles Altieri (2003, p. 10) observes, the arts should be capable of providing a fertile ground for exploring the subtleties of affective experience in all domains:

> Many affects have power in our lives because they emerge as immediate aspects of the kind of attention we pay to the world and to ourselves. And how we feel is often shaped less by belief per se than by how we experience the fit of various elements. Here works of art are instructive because so much depends on their internal dynamics, that is, on matters of structure and pacing and angle of perspective.

Aesthetics, in other words, is a domain that encourages a deeper engagement with affect, rather than a more superficial one. But a deeper engagement requires willingness to confront the elusive: “rather than dismiss what seems inchoate or indefinite, we may have to treat these qualities as fundamental features our affective lives are constantly negotiating” (Altieri, 2003, p. 12). For this to occur, “the first step . . . has to be developing a vocabulary” to articulate and distinguish among music’s many near-ineffable affective aspects (Altieri, 2003, p. 46).

The model presented here attempts to take a step in this direction by offering a vocabulary for different aspects of affective engagement with melody. This vocabulary respects the notion that, in the arts, “different versions of intentionality come into play, especially modes of intentionality connected to values like intensity and connectedness rather than to discursive propositions that evaluate possible actions” (Altieri, 2003, p. 3). It includes terms related to the experience of intensity (surprise-tension), to the highlighting of a melody’s apparent intentionality (denial-tension), and to the impression of desire or forward-directedness in melody (expectancy-tension). This taxonomy was inspired by Lerdahl’s (2001) separation of the phenomena of melodic attraction and implicative denial. If it is true that “affects are ways of being moved that supplement sensation with at least a minimal degree of imaginative projection” (Altieri, 2003, p. 47), then it makes sense that musical affects would be fleeting and projective in the way described by these tension types. However, as people are typically not accustomed to introspecting to this degree about their ephemeral and dynamic melodic impressions, it makes sense that it might take some time, and some practice, to come to match the tension types with the experiences they describe.

The first expectancy-related tension type the model proposes is termed *surprise-tension*, and it correlates inversely with expectancy ratings. Highly predictable events (those with high expectancy ratings) generate little surprise-tension, but extremely unpredictable events (those with low expectancy ratings) generate considerable amounts. Although named for surprise, the tension deriving from unexpectedness registers not as a conscious experience of shock, but rather as a subtle experience of intensity and dynamism. It motivates closer attention from the listener.

The sforzando leap to D in the second full measure of the Mendelssohn excerpt (Figure 6) marks a local peak in surprise-tension. Because D denies a strong expectation for G generated by the preceding F♯, it also creates high *denial-tension*. This second tension type correlates directly with implicative denial. High denial-tension creates a sense of will, intention, or determinedness.

Implicative denial is calculated with the following formula, taken from Lerdahl (2001).

IMPLICATIVE DENIAL FORMULA

$$E_m - E_r,$$

where E_m = the amount by which the maximally expected pitch was expected and E_r = the amount by which the actual realization was expected.

Whereas the expectancy rating measures the expectedness of a given continuation in relation to all possible continuations for all possible events, the implicative denial rating measures the continuation's expectedness in relation only to those continuations possible following the actually occurring event. Figure 7 illustrates the difference between expectancy

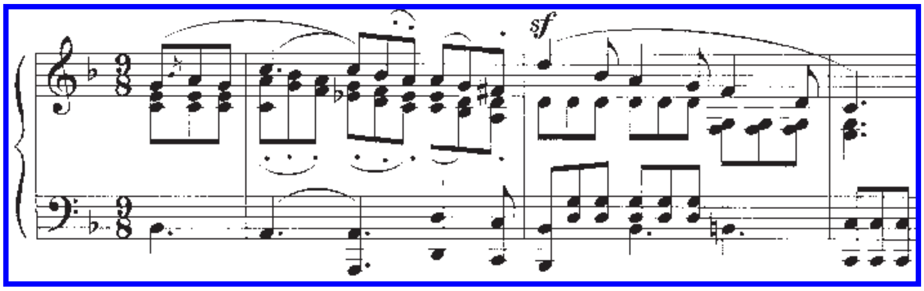
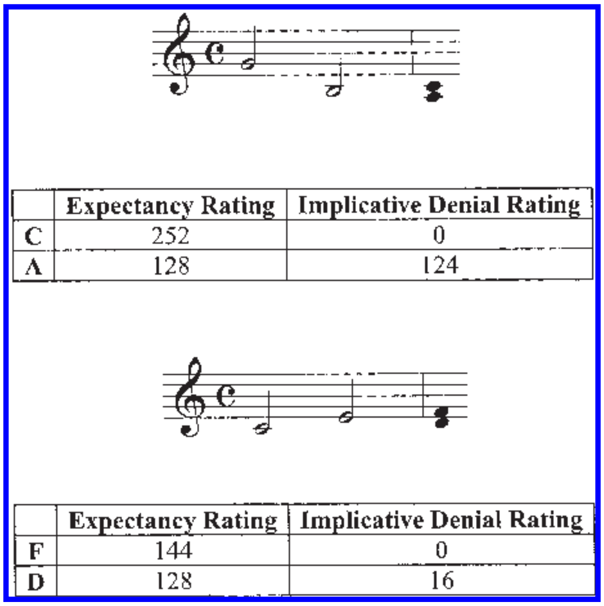


Fig. 6. Mendelssohn, Song Without Words, Op. 53, No. 4, measures 5–8.



	Expectancy Rating	Implicative Denial Rating
C	252	0
A	128	124

	Expectancy Rating	Implicative Denial Rating
F	144	0
D	128	16

Fig. 7. The unstemmed noteheads represent two possible continuations of the melody in half notes. In both examples, one continuation (A in the first excerpt, D in the second) receives an expectancy rating of 128. Because surprise-tension is inversely proportional to expectancy rating, both continuations are predicted to generate the same amount of surprise-tension. In the first example, however, the A thwarts a powerful expectation for continuation to C. Since the difference between the most highly rated possible continuation (252 for C in the first example, and 144 for F in the second example) and the actual continuation (128 in both cases, for A and D, respectively) constitutes the predicted denial-tension, the A is theorized to create much more denial-tension than the D, despite their equivalent expectancy ratings.

and implicative denial. Although one event, a descending minor sixth leap to B in a C-major context, for example, might permit a continuation, in this case to C, with an expectancy rating of 252, another, for example a leap from C to E in C major, might permit a maximally rated continuation of only 144, in this case to F. Had the B from the first scenario instead proceeded to A, with an expectancy rating of 128, it would have generated enormous implicative denial (124), but had the E from the second scenario proceeded to D, also with an expectancy rating of 128, it would have generated little implicative denial (16). The distinction stems from the fact that certain circumstances (a large leap to the leading tone) can produce stronger, more dramatic expectations than others (a skip to the mediant).

A third tension type, *expectancy-tension,* pertains not to the degree to which an event satisfies or denies expectations created by preceding events, but to the strength of expectation generated by an event about future ones. It correlates with the maximum expectancy an event generates (E_m for the next event). Events that trigger strong expectations generate high expectancy-tension, but events that generate mild expectations generate low expectancy-tension. Expectancy-tension creates an impression of strain and desire in a melody. Unstable events generally produce more expectancy-tension than stable ones because strong expectations (and high values for expectancy-tension) occur when a proximate pitch is also stable—a situation that does not typically apply after a stable event. In the Mendelssohn excerpt, the F♯ preceding the leap to D generates high expectancy-tension. Table 5 outlines the three tension types.

Consider the core ratings in C major shown in Table 4. If A occurred after B, it would receive an expectancy rating of 128. Likewise, if A occurred after G, it would receive an expectancy rating of 128. Since surprise-tension depends only on the expectancy rating, the surprise-tension generated by the two continuations would be the same. However, after B, A denies a strong expectation (of 216, for C), generating an implicative denial of 88. After G, A's implicative denial is 0: no more highly rated continuation was possible. Accordingly, in the former case, A would spark significant denial-tension, but in the second case it would create none.

Examine Figure 8, which situates the three tension types within a series of events. Surprise-tension and denial-tension relate to the expectedness of an event given previous ones—the expectedness, for instance, of event Z given events X and Y. Expectancy-tension, on the other hand, relates to the expectedness generated about a future event—the expectedness generated by event Y, for instance, about future event Z. The three tension types are produced differently and signify different experiences. For clarity's sake, the following discussion will explain the function of each tension type without referencing the hierarchic implementation of expectancies—only projections regarding adjacent events will be examined. The mechanisms for the calculation of tension are the same when hierarchic expectancies are included. Note that Figure 8 has been simplified in order

TABLE 5

An Outline of the Three Tension Types, Their Qualitative Aspects, and Their Source in Melodic Expectancy

Tension Type	Associated Experience	Expectancy Source	Formula
Surprise	Intensity, dynamism	Inversely proportional to expectancy rating	$\dfrac{1}{\dfrac{w_i[(s_i \times p_i \times m_i) + d_i]}{w_i}}$
Denial	Will, intention, determinedness	Directly proportional to implicative denial	$E_m - E_r$
Expectancy	Yearning, strain	Directly proportional to expectancy rating of most-expected possible continuation	E_m (of next event)

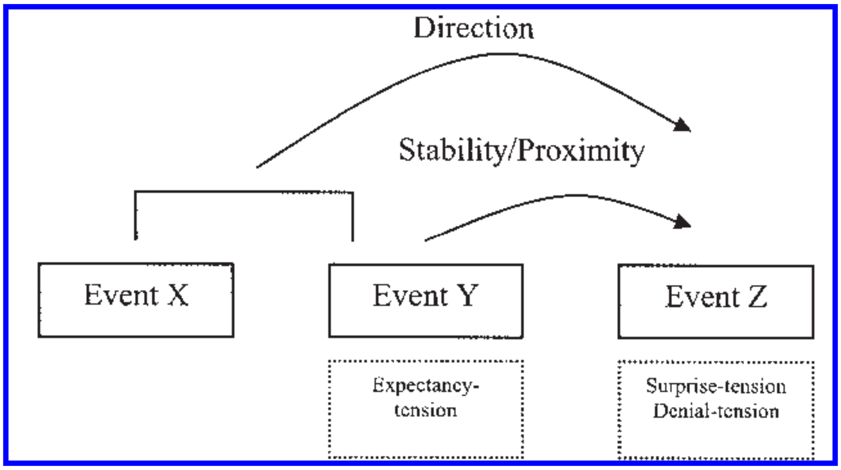


Fig. 8. Expectancy and tension types in context. Based on the interval between events X and Y, the status of event Y within the governing tonality, and the pitch height of Y, expectancy ratings for possible future event Zs can be calculated. The highest of these ratings constitutes Y's expectancy-tension. The rating for the actually occurring event Z constitutes Z's surprise-tension. The difference between this rating and the highest one possible for an event that could have occurred in its place constitutes Z's denial-tension.

to facilitate presentation. In reality, every note is simultaneously functioning as event X in the generation of an expectancy about the note 2 events away, as event Y in the generation of an expectancy about the next note, and as event Z in the realization or denial of a previously extant expectancy regarding it. It may be helpful to think of Figure 8 as a sliding window, where melodic pitches can move from event box to event box as the diagram is pulled forward. Each note can bear surprise-tension, denial-tension, and/or expectancy-tension simultaneously. A note with more expectancy-tension may seem saliently future directed, and a note with more surprise-tension more retrospective in character.

To determine an event's surprise-tension, the event's expectancy rating must be calculated using the formula $(s \times p \times m) + d$. In Figure 8, the pitch height of event Y determines the proximity rating for event Z: the closer event Z is to event Y, the higher its rating. Event Y, considered within the existing tonal context, also determines the stability rating for event Z: the more stable event Z is within the context established by event Y and the underlying chord and key, the higher its rating. Event X must be brought into the picture to determine event Z's direction rating. Event X and event Y together form an interval that determines the direction rating for event Z: if event Z lies in the predicted direction, it receives a higher rating. The multiplication of the stability and proximity ratings for event Z, and the addition of its direction rating, creates event Z's expectancy rating. Event

Z’s surprise-tension is inversely proportional to its expectancy rating. The less expected the event, the more surprise-tension it produces.

Whereas surprise-tension is backward-looking, registering the effect of an event in light of previous ones, expectancy-tension is forward-looking, registering an expectation about the future. To determine an event’s expectancy-tension, calculate the highest expectancy rating it could assign to a hypothetical subsequent event. The expectancy-tension of event Y in Figure 8, for example, depends on the strength of expectancy it generates about event Z. To find the strongest expectancy generated by Y, compare the expectancy ratings of possible event Zs. Event Y determines the possible proximity and stability ratings, and the interval produced by the distance between event X and event Y determines the possible direction ratings. The highest $(s \times p \times m) + d$ combination of these constitutes the strongest expectancy generated at event Y, and, accordingly, that event’s expectancy-tension. Expectancy-tension varies directly with the value of the strongest expectancy. It is, to use Bharucha’s terminology, the tension of “yearning,” or powerful expectancy (Bharucha, 1984).

Denial-tension, like surprise-tension, interprets an event in light of past ones. The denial-tension of an event relates to the degree to which it fulfills or denies the specific expectations that existed about it before it occurred. Denial-tension depends jointly on the two factors that surprise-tension and expectancy-tension depend on individually. To calculate event Z’s denial-tension, subtract its expectancy rating (which correlates inversely with Z’s surprise-tension) from the expectancy rating the most expected event would have received (this rating correlates directly with Y’s expectancy-tension). The result assesses the amount Z fulfilled the expectations created by Y.

By separating the experiences stemming from expectations into three categories, the model makes a preliminary step toward a richer taxonomy of the multiple dimensions of musical experience. Music does not seem merely a linear succession of more and less tense junctures; rather, it seems qualitatively rich and multidimensional. Expectancy-tension is an inherently forward-looking, prospective phenomenon, and events with high values for it should seem saliently implicative. Denial-tension, on the other hand, is inherently backward-looking, registering an event in light of specific preceding ones. Events with high values for denial-tension should seem more saliently connected to the past than to the future. Reducing these different sensations to a single parameter would depict musical experience as poorer than it seems in reality. If one goal of music cognition research is to uncover and articulate aspects of the listening experience that seem “purely musical” and difficult to describe, theories should seek to expand and enrich the vocabulary for musical impressions, not artificially limit it. Because the tensions described here are associated with the temporal directedness of attention, experimental paradigms created to gauge attention might be successfully adapted to track the predicted tension types.

Expectancy-Based Tension in the Opening of Mozart K. 282

An analysis of the opening of Mozart's Piano Sonata K. 282 will illustrate the proposed tensional consequences of expectancy and expectancy violation. Figure 9 shows measures 1–8, also analyzed extensively in a special issue of *Music Perception* (Vol. 13, No. 3, Spring 1996). For any tonal excerpt, the model can generate a graph of the fluctuations in each tension type across the course of the melody. Figure 10 shows, as an example, the changes in surprise-tension predicted across measures 1–8 of the Mozart excerpt. Because surprise-tension varies inversely with expectancy ratings, the *y* axis's reversal (height signifies a *decrease* in ratings) enables graph peaks to represent tension peaks. The three peak events in measures 1–8 are the high G in measure 4, the high A♭ in meas-

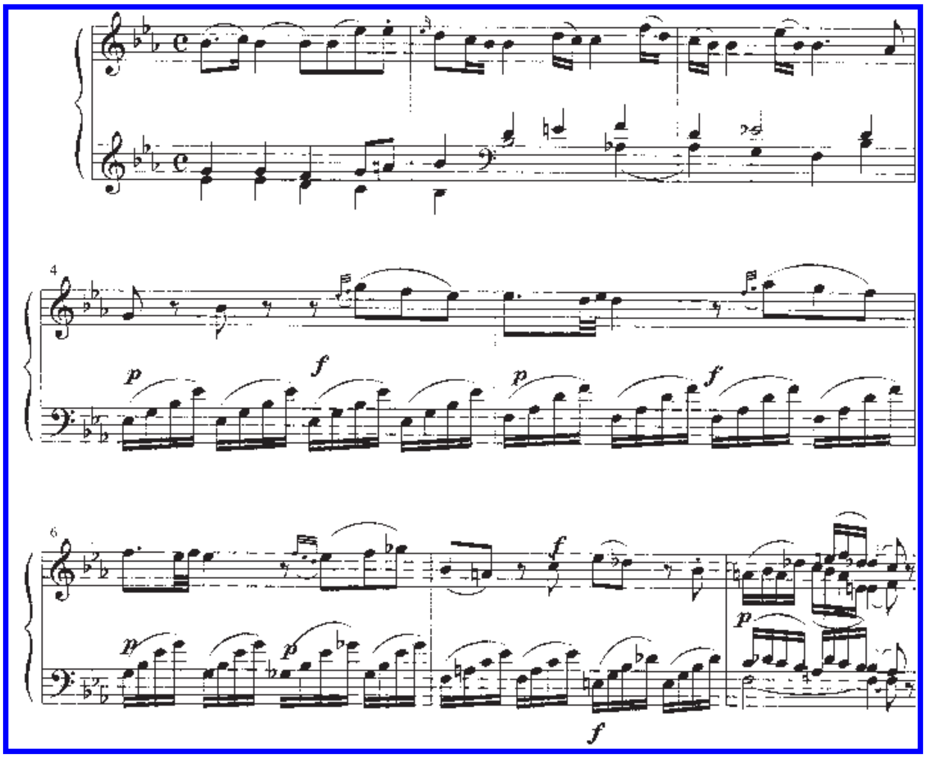



Fig. 9. Mozart, Piano Sonata K. 282, measures 1–8.

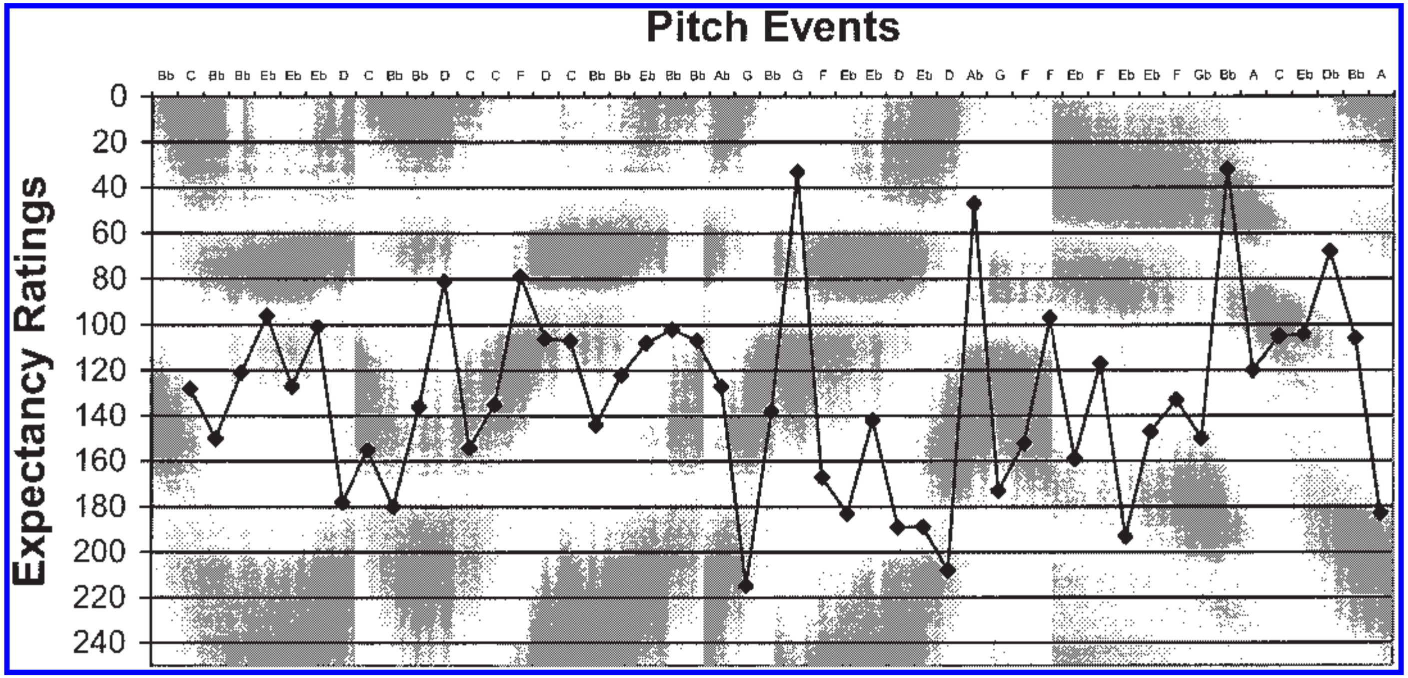


Fig. 10. Predicted overall surprise-tension across the course of the melody of measures 1–8 of Mozart K. 282, where the tempo is quarter = 48. The *x* axis represents pitch events, and the *y* axis represents expectancy ratings, listed from high to low so that the inversely proportional surprise-tension can be seen to increase with graph height.

ure 5, and the first B♭ in measure 7. However, only the A♭ in measure 5 and the B♭ in measure 7 represent equally noticeable denial-tension peaks (see Figure 12). Although the A♭ in measure 5 denied a strong expectation for the preceding leading-tone D to progress to E♭, and the B♭ in measure 7 denied a strong expectation for the preceding chromatic G♭ to descend to F, the G in measure 5 follows B♭, $\hat{5}$, the least implicative scale degree, after which expectations are typically weak and vague.

Noticeably, the chromatic G♭ does not count among the excerpt's lowest-rated events; in fact, it receives a relatively high expectancy rating and produces 0 implicative denial. The preceding nonchord tone F shifts the context to the underlying e♭-minor chord, elevating G♭ to the most expected continuation. G♭ remains a salient event, by virtue of its registral prominence, its chromaticism, and its high expectancy-tension, but not an unexpected one. Unexpectedness, although it can produce a type of tension (namely, surprise-tension), is not synonymous with tension. Events, like the G♭ in measure 6, can be expected yet quite arresting—either on account of expectancy-tension, or of some harmonic, dynamic, or timbral feature not treated by the model. Imagine that the Mozart excerpt arrived at the second beat E♭ in measure 6, and continued to a melodic G♭ at the start of the third beat over the change to the borrowed chord, in place of the eighth rest, as shown in Figure 11. Because unprepared by the harmony, that G♭ would generate surprise-tension in addition to the expectancy-tension present in the unaltered version, increasing the severity of G♭'s effect.

The expectancy ratings depicted in Figure 10 reflect interactions among the four primary parameters—stability, proximity, direction, and mobility—applied hierarchically as described earlier. To gain a better understanding of how the computations work, consider the C at the start of measure 3. Because the preceding D can be understood as a chord tone within the underlying diatonic harmony, it does not shift the governing chord and key context from the default tonic (I) in E♭ major. Within this context, C, as a member of the governing key but not the governing chord, receives a stability rating of 4. Lying a whole step

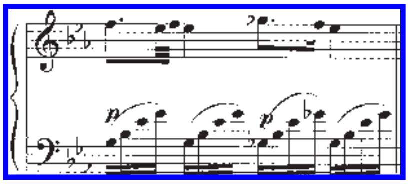


Fig. 11. An alternate version of measure 6 from K. 282. Since the melodic G♭ coincides with the harmonic shift (rather than being prepared by it), it garners a low expectancy rating and, accordingly, is predicted to generate high surprise-tension.

away from D, it receives a proximity rating of 32. The preceding interval, a descending minor third from F to D, confers a direction rating of 6 on C for continuing the downward motion. The mobility constant is 1 because C does not repeat the preceding pitch. Before hierarchic effects are considered, C's expectancy rating is (4 × 32 × 1) + 6, or 134. At the quarter-note level, however, (see Figure 4), C's rating is (4 × 16 × 1) + 6, or 70, because of its decreased proximity in comparison to that level's preceding span head, F. C receives the same rating at the eighth-note level.[9] At the half-note level, C's rating is (4 × 36 × 2/3) + (12 × 1/3), or 100. At this level, C follows a span whose head was also C. Therefore, the proximity rating is maximally high (36), but the mobility constant (2/3) applies. Additionally, the preceding small interval from D-C generated a directional impulse of 12 for continued descent, which lateral motion partially fulfills. As the section on direction above explains, lateral motion is theorized to fulfill directional tendencies by a factor of 1/3. At the measure level, C receives a rating of (4 × 32 × 1) + 0, or 128, because it follows head D but receives no directional boost from the preceding ascending interval from B♭ to D.

At a typical tempo for this movement (48 to the quarter note), spans at the eighth-note level last 0.625 s, spans at the quarter-note level last 1.25 s, spans at the half-note level last 2.5 s, spans at the measure level last 5 s, and spans at the two-measure level last 10 s. Because a span of 10 s exceeds most estimates of the duration of the perceptual present (London, 2002),[10] the two-measure level in this example is theorized not to play a role in the expectancies under consideration. The two-measure level, in other words, receives a weight of zero. Because there seems to be another perceptual cutoff around the duration of 2 s, beyond which rhythmic synchronization becomes difficult or impossible (London, 2002), the model assigns a higher weight to spans with durations shorter than 2 s, reflecting the theorized greater role they play in priming. In this example, the eighth-note and quarter-note levels receive the higher weight of 5, and the half-note and measure levels the lower weight of 2, reflecting in this example the notion that the proximity violation from F to C is more salient than the continuity between C and the preceding measure's D. When the ratings for C at each level are combined via the weights, C receives an overall rating of 109, reflecting the joint influence of levels at which it is relatively expected (the note-to-note level) and levels at which it is relatively unexpected (the quarter-note level). Because another continuation,

---

[9] Because the eighth-note level applies so rarely in this excerpt (see rule 4.d in the complete expectancy formula—levels whose spans match or go below notated durations drop out), it is not included in the reduction in Figure 4.

[10] For example, Fraisse (1963) places the limit around 5 s. Clarke (1987), at the long end of the range, places it around 10.

E♭, had it occurred in place of the C that starts measure 3, would have fulfilled expectations at all levels, including the note-to-note and the quarter-note, C receives a large implicative denial rating of 96. The model proposes, accordingly, that C's most intense effect is one of willedness, as if the melody had pushed in a direction it would not have followed without the intervention of agency.

The influence of hierarchic expectancies is particularly noticeable at several points in the Mozart example. Consider the A♮ in measure 7—at the note-to-note level, a strongly expected resolution of the nonchord tone B♭. However, at the quarter-note level and beyond, the A♮ strongly violates expectancies in terms of stability (as a chromatic pitch), proximity, and direction. The background violation might draw attention away from the note-to-note level, and produce a sensation of melodic intensity at the A♮, despite its expectedness at shallower levels. Contrastingly, consider the E♭ in measure 3. Before hierarchic levels are considered, it receives a relatively low rating. However, at levels beyond the quarter note, it is understood as an elaboration of the highly expected preceding B♭. In this case, hierarchic factors raise the overall rating for E♭. Consider, finally, the end of measure 3, which features a notable confluence of expectation: each level targets G as the maximally expected continuation. The unanimity and force of that projection makes the G at the start of measure 4 uniquely inevitable-seeming.

Figure 12 graphs the predicted overall denial-tension across the course of the same measures from Mozart K. 282. Consider the C that starts measure 3: it occupies a middle low region on Figure 10, indicating moderate surprise-tension. However, on Figure 12, the same C represents a peak in denial-tension. The contrast indicates that although C is diatonic (not particularly unstable) and proximate to the previous D, it does deny a strong, specific expectancy for the preceding D to resolve to E♭. Accordingly, the C is not predicted to be experienced as a maximally intense or poignant moment, but rather as one where the melody seems to assert it will, to become energized and endowed with intent. Consider next the high G in measure 4 and the initial B♭ in measure 7—the maximal theorized generators of surprise-tension in the excerpt. Figure 12 shows that although the B♭ is also a projected creator of large denial-tension, the G is predicted to create a much smaller amount of this tension type. The distinction is due to G following the relatively non-implicative fifth scale degree, in contrast to B♭ following the highly implicative chromatic G♭. The theorized experiential consequence of this distinction is that the B♭ sounds not only intense, but also like a willful push in a new direction, but the G only sounds intense—there is no sense of a determined thrust—making the G an overall less marked occurrence.

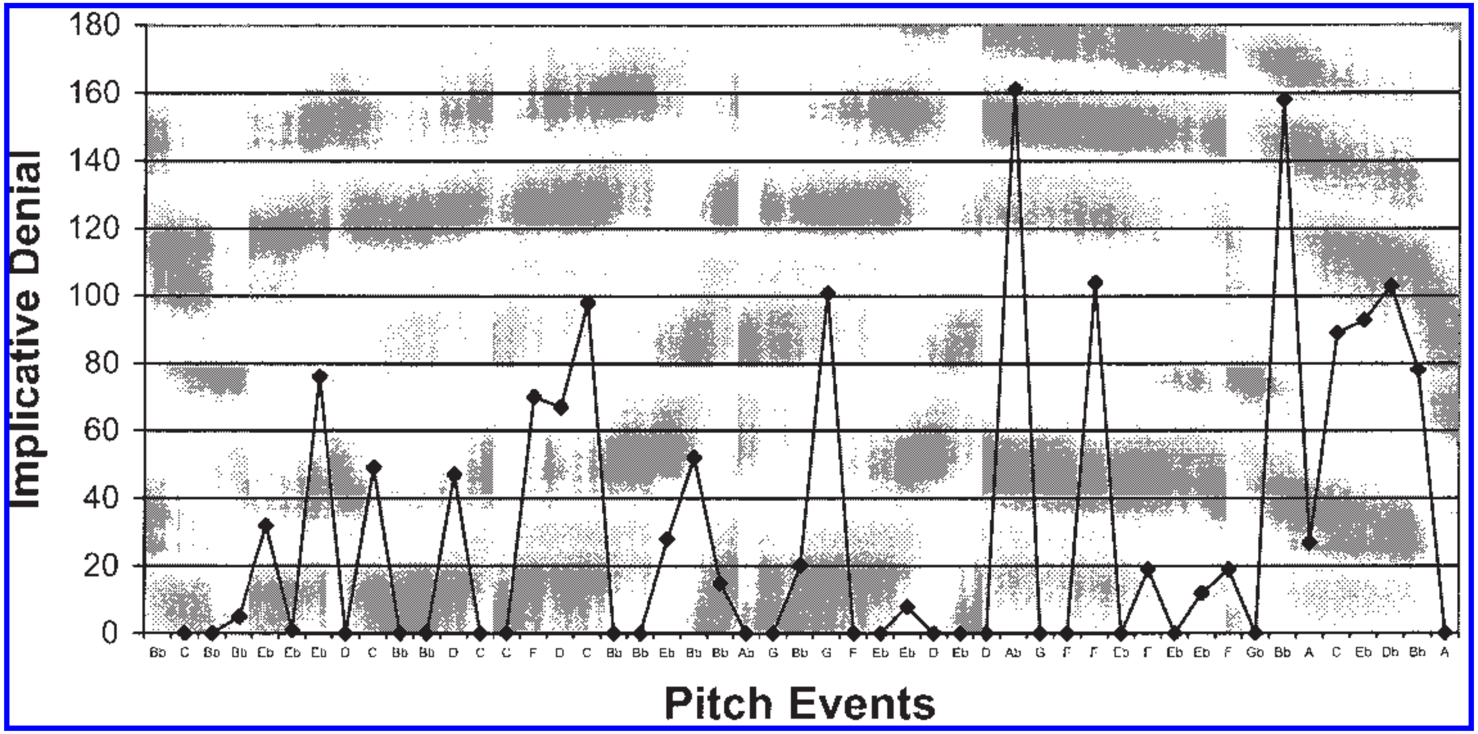


Fig. 12. Predicted overall denial-tension across the course of the melody of measures 1–8 of Mozart K. 282, where the tempo is quarter = 48.

Figure 13 shows the predicted overall expectancy-tension across the course of the same passage. The A♭ at the end of measure 3 represents the excerpt's expectancy-tension peak. As the 7th of a V7 chord, it marks a maximally implicative moment, and is theorized to generate a sense of yearning for resolution to G (indeed, part of the relaxedness theorized to arise when G does occur on the downbeat of measure 4 stems from its dispelling of the forward looking expectancy-tension of the preceding A♭). Note that expectancy-tension relates to the expectancies an event generates, not the ones it fulfills: the A♭ itself receives a rather high expectancy rating (i.e., produces relatively little surprise-tension). It is also noteworthy that the piece begins with some of the lowest expectancy-tension of the excerpt. In default contexts (governed by the tonic chord) in major keys, the fifth scale degree (B♭ in this excerpt) produces the lowest expectancy-tension of any, because it lies a whole step (rather than a half-step, which would increase the proximity expectancy) from both its diatonic neighbors—the fourth and sixth scale degrees—neither of which is particularly stable (i.e., a member of the governing tonic triad). This leads to an opening without significant expectancy-tension, an opening that seems rather composed, and lacks a strong forward impulse until it moves to the E♭s at the end of measure 1.

By way of contrast, consider Lerdahl's analysis (1996) of the same measures. Lerdahl's attractional analysis evaluates the sum tension created by relations in tonal pitch space (including harmonies), and the current one evaluates only the tension created by melodic expectancy. Lerdahl's analysis requires a final-state awareness—future events influence the rating of current ones—whereas the present model, because it captures expectancy, by definition disallows the influence of future events. In keeping with their different approaches, the models produce markedly different results. A particularly noticeable divergence relates to the rating for the leap to high G in measure 4. As the largest leap in the excerpt, the current model assigns G a high implicative denial rating and a markedly low expectancy rating. Lerdahl's model, on the other hand, rates it as one of the excerpt's most relaxed events. The current model captures melodic G's unexpectedness, and Lerdahl's model captures the stability of the underlying harmony. For similar reasons, Lerdahl's model assigns a high tension rating to the G♭ in measure 6, an event the current model assesses as expected. Lerdahl's model captures the foreignness of the chromatic pitch, and the current model captures its resolution of the preceding nonchord tone.

Another point of contrast is offered by Narmour's analysis (1996) of the same measures. Because of its notational complexity, heavily segmented nature, and exclusion of stability factors, Narmour's analysis, though perhaps closer in aim, proves more difficult to compare.

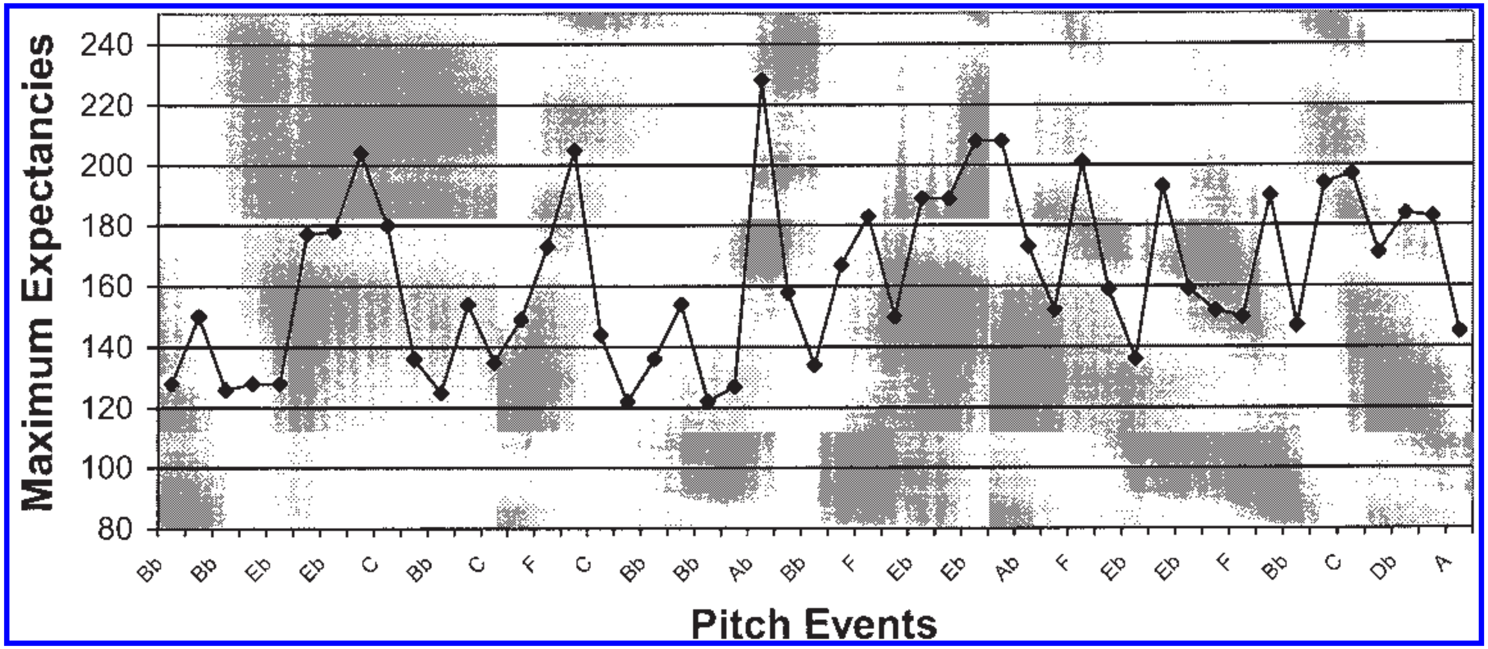


Fig. 13. Predicted overall expectancy-tension across the course of the melody of measures 1–8 of Mozart K. 282, where the tempo is quarter = 48.

The expectational content behind the symbols must be narrated for readers who are not already well acquainted with his theory. Additionally, his analysis does not describe the expectational content of every event. Within a bracketed group, at least two notes must transpire before an expectation may be articulated, and then the expectation refers only to the third event, not to either of the first two. Consider the opening neighbor-tone elaboration of B♭, bracketed and labeled ID in Figure 14, Narmour's analysis of the excerpt. The abbreviation stands for "intervallic duplication" and indicates that the second B♭ satisfies the expectation for a small interval, but denies the expectation for continued ascent. No account of expectations about the neighbor-tone C is provided. The next bracketed structure, the leap from B♭ to E♭, emerges as a dyad (4), signifying an interval whose implications are suppressed. Accordingly, two more notes must transpire before any expectations apply. The next note that the notation describes is the first E♭ in measure 2, which, as a result of duplication (D), satisfies both intervallic and registral implications. Among the first 7 melodic pitches of the excerpt (from the start through the downbeat of m. 2), Narmour's analysis contains information only about the expectancy fulfillment or denial of two events: the second B♭ in measure 1 and the first E♭ in measure 2. Although his implication-realization theory successfully articulates many of the factors at play in melodic expectancy, it cannot be a complete theory, if only because of its omission of expectational assessments of many, and in some cases most, events in each melody.

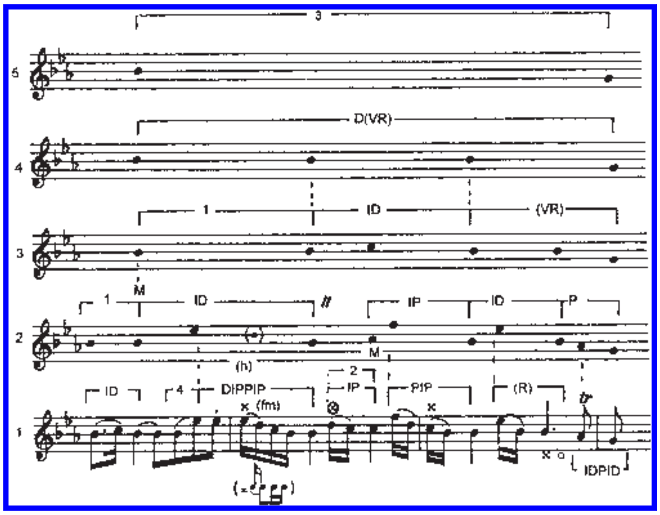



Fig. 14. Narmour's analysis of measures 1–4 from Mozart K. 282. (Narmour, 1996).

Conclusion

The model outlined here generates expectancy ratings for melodic events. It associates these ratings with listeners' experiences of tension across melodies. The specificity of its predictions makes the model particularly amenable to empirical study. Much recent research on music and emotion has focused on more static aspects of listening experience, such as perceived moods (Juslin & Sloboda, 2001). Expectational analysis may be one way of examining those aspects of listening experience that are dynamic, fleeting, and difficult to articulate.

Experimental study is the next logical step in the development of the model. Three approaches seem particularly natural: priming studies, ERP research, and tension measures. All three possess the potential to reveal the dynamic course of expectancies without requiring the subject to consciously access them. Bharucha and Stoeckig (1986) adapted a priming paradigm to the study of harmonic expectancy, using reaction times and accuracy on a tunedness judgment task to assess the degree to which different chords are primed. Margulis and Levine (2004a, 2004b) employ reaction times and accuracy on a timbre judgment task to gauge the degree to which various melodic continuations are primed. Faster reaction times and improved accuracy, reflecting the facilitation of processing created by priming, were predicted to occur for expected continuations. Figure 15 shows an example of the results obtained in this experiment: in comparison to accuracy on the timbre judgment task for the pitches presented in isolation, accuracy for the pitches presented after a melodic context was substantially increased for C, a stable and proximate continuation, somewhat increased for A, a bit less stable and less proximate continuation, and decreased for A♯/B♭ and G, unstable and nonproximate continuations, respectively. Granot and Donchin (2002) used the P300 component of the ERP to measure the degree of expectancy violation in tonal sequences. A similar experiment using a wider variety of possible continuations could assess the validity of the model's ranking of expectancy violations.

Another empirical approach to implicit expectancies might be an investigation of the tensional profile the expectancies are theorized to create. Krumhansl (1996), for example, asked subjects to move a slider in response to the degree of tension experienced across the first movement of Mozart's K. 282. Listeners responded to the full context—not just the melody—and in broad outline the tension judgments seemed to correlate with segmentation, tempo variation, contour, pitch height, density change, and dynamics (Krumhansl, 1996). Tension can be viewed as a composite phenomenon, with the above factors as well as harmonic distance (Lerdahl, 2001), dissonance (Smith & Cuddy, 2003), and expectancy con-

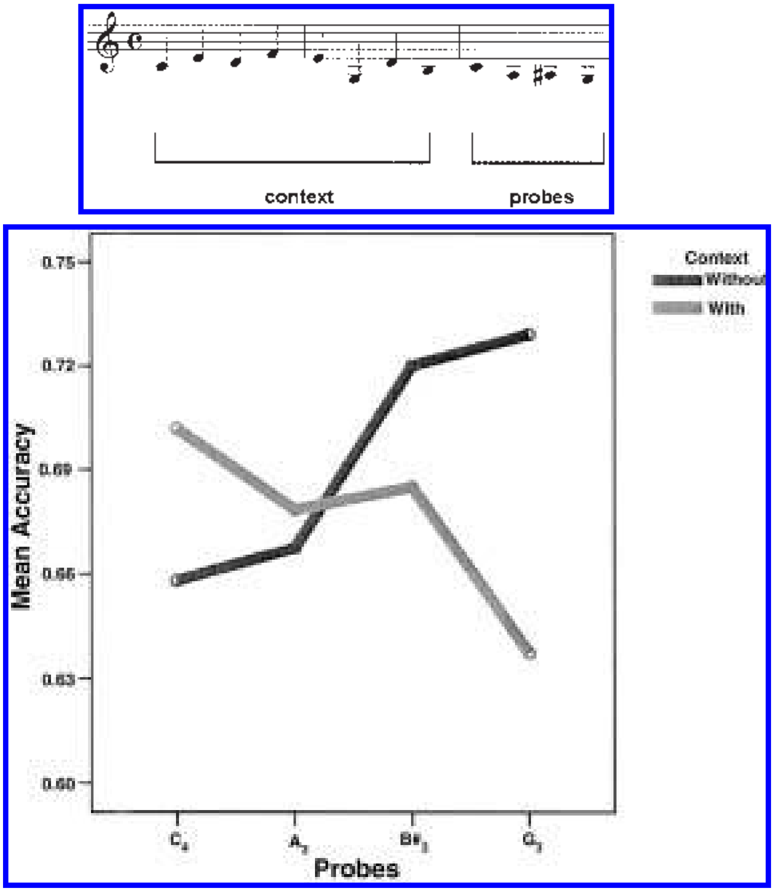


Fig. 15. Accuracy on a timbre judgment task (Margulis & Levine, 2004a, 2004b). Four probe tones played in one of two electric piano–like timbres followed the context melody played in a pianolike timbre. Listeners were asked to identify the timbre of the probe tone as quickly and accurately as possible. A comparison of accuracy on the timbre judgment task for the probe tones when they followed the displayed melodic context with accuracy on the task for the probe tones when they were presented in isolation illustrates that performance was facilitated for continuations that the model theorizes to be expected.

tributing in important ways. Because of the numerous factors at play in the generation of tension in this full musical excerpt, it is surprising and compelling that the model's predictions for expectancy-tension correlate positively with tension ratings in Krumhansl's experiment ($R = .257$, $N = 44$, $p < .05$), as do the model's predictions for denial-tension ($R = .323$, $N = 44$, $p < .05$). The correlation for the model's predictions of surprise-tension do not rise to significance, but correlate positively as expected ($R = .095$, $N = 44$, $p > .05$). The present article does not claim that musical tension arises solely from melodic expectancy; rather, it claims that expectancy is one important factor in the generation of the complex phenomenon of experienced musical tension. The results of the comparison between the model's tension predictions and Krumhansl's tension data suggests that expectancy-based tension forms an important part of the generation of overall experiences of tension. An even better way to empirically investigate the ties to affect proposed here would be first to establish a baseline by asking listeners to adjust a slider representation tension heard across the course of a musical excerpt, and second to adjust only the expectancy profile of the melody, controlling as much as possible for other theorized expectancy sources (e.g., segmentation, dissonance, tempo, and dynamics). A comparison of the tension ratings produced in the first and second cases should reveal the tension fluctuations due exclusively to expectancy. Given phenomenological descriptions of the theorized tension types (e.g., "yearning" for expectancy-tension, "willful" for denial-tension), listeners may be able to rate fluctuations in individual tension types across the course of melodies, again allowing comparison with the model's predictions.

Perhaps the prospective, or forward-looking quality of events (as embodied in the concept of expectancy-tension) could be investigated with a memory task; because high expectancy-tension increases the amount of attention directed at subsequent events, memory of events following notes with high expectancy-tension should be improved relative to memory of the same events following notes with low-expectancy tension. Likewise, the backward-looking quality of events (as embodied particularly in the concept of denial-tension) may be susceptible to investigation via a different memory task. Because the influence of preceding events lingers more when ensuing events possess high denial-tension, performance on a memory task for events preceding a pitch with high denial-tension should be improved relative to performance on a memory task for the same events when placed before a pitch with low-denial tension. It may also be possible for a better taxonomy of musical tension to tease out separate elements of the qualitative experience typically lumped together under the generic rubric of "tension." The categorization in the present model attempts to make a step in this direction.

Additionally, it is hoped that experimental work can suggest the best ways for expanding the model in two directions: first, incorporating grouping structure and, second (relatedly), integrating temporal expectancies with the other types. Grouping boundaries most likely suppress expectancies. They may also cause structurally higher hierarchic expectations to become relatively more salient than structurally lower ones. Regarding temporal expectancies, the model suggests that the earlier or later than expected an event comes, the more it violates temporal expectancies. Yet these temporal expectancies likely intersect with pitch-based expectancies. For example, as an event is increasingly delayed, the passing time might create a more intense violation of temporal expectancies, but the delay most likely starts to create a grouping boundary and weaken expectancies about the next event's pitch. The temporal placement of an event, then, affects the way in which groups are formed and the way in which structural pitches are integrated into background lines—both factors that feed back into the time-span reduction and thus influence the production of pitch hierarchies. To a certain extent, these factors are integrated into the model via the inclusion of time-span reduction rules, but a fuller account of temporality remains a topic worthy of expanded treatment in the future.

Mari Riess Jones (1981, 1982, 1989, 1990) has advanced a theory of dynamic attending—a theory of the ways in which listeners temporally guide and focus their attention when listening to music. "The gist of this idea is that listeners 'use' pitch/time relationships to anticipate the 'where' of pitch and 'when' of time in an ongoing event" (Jones & Holleran, 1992). The expectations are stratified, and can apply at lower, note-to-note levels, termed *analytic attending,* and high, phrase and form levels, termed *future-oriented attending.* Different time-span levels can cause listeners to concurrently hold different expectations about the same melodic point: analytic attending may cause listeners to expect one thing, and future-oriented, another.

Considering the work of Jones in light of the present model, it seems there are two temporal expectancy types an expanded theory might address: time-point expectancies and metric-level expectancies. If a pitch is expected at a particular time point, its actual arrival might come early or late in comparison to that point. However, the descriptives of early and late do not adequately capture the complexity of temporal expectancy. A pitch that arrived late with respect to a particular time point might arrive on a later hypermetric downbeat, or on a later thirty-second-note division of a weak beat. There is, in other words, another component to temporal expectancy—a metric component—that predicts the pitch to occur at a certain level within the metric hierarchy. In comparison to the predicted level, the actual event may occur at a point that is metrically subordinate or metrically superordinate. Additionally, the occurrence of an event predicted by larger scale future-oriented attending most likely constitutes a materially different experience than the occurrence of an event predicted by smaller scale analytic attending. Temporal expectancies, in sum, involve not only earliness or lateness, but also metric subordination or superordination, and can occur over broader spans (future-oriented attending) or short spans (analytic attending).

As an example, consider the D at the start of measure 2 in the Mozart K. 282 excerpt. This D not only constitutes a continuation implied by the immediately preceding (grace note) E♭, but also by the E♭ that constituted the head of the second half-note span in measure 1. Because the half-note level expectancy was stronger, its resolution to D might seem more salient than the resolution of the surface E♭, a circumstance that might steer attention to the half-note level. Over the next three measures, the half-note level is characterized by an expectancy-fulfilling stepwise descent to G in measure 4. Assuming that the half-note heads that form the components of this stepwise descent (E♭-D-C-B♭-A♭-G) are expected to occur at the starts of their respective spans, it could be claimed that each one seems to come a bit late. The D is delayed by the grace note, the C by the preceding sixteenths, and the Ab is delayed most of all—not occurring until the final eighth note of its span. Observe that the sixteenth-note anticipation of C that occurs within the third beat of measure 2, over the E♮, seems too metrically subordinate to constitute a fulfillment of the expectation for descent to C generated by the D head of the preceding half-note span. It could be argued that this C is experienced more as an adjacent consequence of the immediately preceding D, and it is only the next quarter-note C that is experienced as a consequence of the half-note-level D.

These issues require a detailed system for the assignment of grouping structure, one that prospectively identifies the ending and starting points of segments, phrases, and larger sections. They also require an understanding of the ways in which musical events can steer temporal attending. Such a system currently lies outside the scope of the model, but it remains a possible area for expansion.[11]

References

Altieri, C. (2003). *The particulars of rapture: An aesthetics of the affects.* Ithaca, NY: Cornell University Press.

Berlyne, D. E. (1971). *Aesthetics and psychobiology.* New York: Meredith Corporation.

Besson, M., & Faïta, F. (1995). An event-related potential (ERP) study of musical expectancy: Comparison of musicians with nonmusicians. *Journal of Experimental Psychology: Human Perception and Performance, 21,* 1278–1296.

Bharucha, J. J. (1984). Anchoring effects in music. *Cognitive Psychology, 16,* 485–518.

Bharucha, J. J. (1987). Music cognition and perceptual facilitation: A connectionist framework. *Music Perception, 5,* 1–30.

Bharucha, J. J. (1996). Melodic anchoring. *Music Perception, 13,* 383–400.

Bharucha, J. J., & Stoeckig, K. (1986). Reaction time and musical expectancy: Priming of chords. *Journal of Experimental Psychology: Human Perception and Performance, 12,* 403–410.

Bigand, E. (2003). Contributions of music to research on human auditory cognition (D. Dusinberre, Trans.). In S. McAdams & E. Bigand (Eds.), *Thinking in sound: The cognitive psychology of human audition* (pp. 231–277). New York, NY: Oxford University Press.

Bregman, A. S. (1990). *Auditory scene analysis.* Cambridge, MA: MIT Press.

---

[11] This article benefited greatly from comments and suggestions from Fred Lerdahl, as well as from Eric Clarke and three anonymous reviewers. Carol Krumhansl helpfully provided data from her 1996 study.

Carlsen, J. C. (1981). Some factors which influence melodic expectancy. *Psychomusicology, 1,* 12–29.

Clarke, E. (1987). Levels of structure in the organization of musical time. *Contemporary Music Review, 2,* 211–239.

Crandall, J. E. (1967). Familiarity, preference, and expectancy arousal. *Journal of Experimental Psychology, 73,* 374–381.

Cuddy, L. L., & Lunney, C. A. (1995). Expectancies generated by melodic intervals: Perceptual judgments of melodic continuity. *Perception and Psychophysics, 57,* 451–462.

Dewey, J. (1894). The theory of emotion. *Psychological Review, 1,* 553–569.

Dowling, W. J. (1990). Expectancy and attention in melodic perception. *Pyschomusicology, 9,* 148–160.

Eeorla, T, Toiviainen, P., & Krumhansl, C. (2002). Real-time prediction of melodies: Continuous predictability judgments and dynamic models. Proceedings of the 7th International Conference on Music Perception and Cognition, Sydney, 2002.

Fodor, J. A. (1996). *The modularity of mind.* Cambridge, MA: MIT Press.

Fraisse, P. (1963). *The psychology of time* (J. Leith, Trans.). New York: Harper and Row.

Granot, R., & Donchin, E. (2002). Do re me fa sol la ti—constraints, congruity, and musical training: An event-related potentials study of musical expectancies. *Music Perception, 19,* 487–528.

Harriss, E. (1981). *Johann Mattheson's "Der vollkommene Capellmesiter": A revised translation with critical commentary.* Ann Arbor, MI: UMI Press.

Jackendoff, R. (1991). Musical parsing and musical affect. *Music Perception, 9,* 199–230.

Jones, M. R. (1981). Music as a stimulus for psychological motion. Part I: Some determinants of expectancies. *Psychomusicology, 1,* 14–31.

Jones, M. R. (1982). Music as a stimulus for psychological motion. Part II: An expectancy model. *Psychomusicology, 2,* 1–13.

Jones, M. R. (1989). Dynamic attending and responses to time. *Psychological Review, 96,* 459–491.

Jones, M. R. (1990). Learning and the development of expectancies: An interactionist approach. *Psychomusicology, 9,* 193–228.

Jones, M. R., & Holleran, S. (Eds.). (1992). *Cognitive bases of musical communication.* Washington, DC: American Psychological Association.

Juslin, P., & Sloboda, J. (2001). *Music and emotion: Theory and research.* Oxford: Oxford University Press.

Justus,T. C., & Bharucha, J. J. (2001). Modularity in musical processing: The automaticity of harmonic priming. *Journal of Experimental Psychology: Human Perception and Performance, 27,* 1000–1011.

Kennan, K. (1998). *Counterpoint* (4th ed.). New York: Pearson Education.

Krumhansl, C. (1990). *Cognitive foundations of musical pitch.* New York: Oxford University Press.

Krumhansl, C. (1996). A perceptual analysis of Mozart's Piano Sonata K. 282: Segmentation, tension, and musical ideas. *Music Perception, 13,* 401–432.

Lake, W. (1987). *Melodic perception and cognition: The influence of tonality.* Unpublished doctoral dissertation, University of Michigan, Ann Arbor.

Larson, S. (1993). Modeling melodic expectation: Using three "musical forces" to predict melodic continuations. In *Proceedings of the Fifteenth Annual Conference of the Cognitive Science Society: June 18 to 21, 1993, Institute of Cognitive Science, University of Colorado, Boulder* (pp. 629–634). Hillsdale, NJ: Lawrence Erlbaum Associates.

Larson, S. (2002). Musical forces, melodic expectation, and jazz melody. *Music Perception, 19,* 351–385.

Larson, S. (2004). Musical forces and melodic expectations: Comparing computer models with experimental results. *Music Perception, 21,* 457–498.

Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music.* Cambridge, MA: MIT Press.

Lerdahl, F. (1996) Calculating tonal tension. *Music Perception, 13,* 319–364.

Lerdahl, F. (2001). *Tonal pitch space.* New York: Oxford University Press.

London, J. (2002). Cognitive constraints on metric systems: some observations and hypotheses. *Music Perception, 19,* 529–550.

MacCurdy, J. (1925). *The psychology of emotion.* New York: Harcourt, Brace and Co.

Maddi, S. R. (1961). Affective tone during environmental regularity and change. *Journal of Abnormal and Social Psychology, 62,* 338–345.

Margulis, E. H. (2003). *Melodic expectation: A discussion and model.* Unpublished doctoral dissertation, Columbia University, New York.

Margulis, E. H., & Levine, W. H. (2004a). Melodic expectation: a priming study. In S. D. Lipscomb, R. Ashley, R. O. Gjerdingen, & P. Webster (Eds.), *Proceedings of the Eighth International Conference on Music Perception and Cognition* (pp. 364–366). Adelaide, Australia: Causal Productions.

Margulis, E. H., & Levine, W. H. (2004b). *Priming effects and expectancy in melody.* Manuscript submitted for publication.

Masson, M. E. (1986). Comprehension of rapidly presented sentences: The mind is quicker than the eye. *Journal of Memory and Language, 25,* 588–604.

Meyer, L. B. (1956). *Emotion and meaning in music.* Chicago: University of Chicago Press.

Moles, A. (1966). *Information theory and esthetic perception* (J. Cohen, Trans.). Urbana: University of Illinois Press.

Narmour, E. (1990). *The analysis and cognition of basic melodic structures.* Chicago: University of Chicago Press.

Narmour, E. (1992). *The analysis and cognition of melodic complexity.* Chicago: University of Chicago Press.

Narmour, E. (1996). Analyzing form and measuring perceptual content in Mozart's sonata K.282: A new theory of parametric analogues. *Music Perception, 13,* 265–318.

Peretz, I., & Coltheart, M. (2003). Modularity of music processing. *Nature Neuroscience,* 6, 688–691.

Posner, M. I., & Snyder, C. R. R. (1975). Attention and cognitive control. In R. L. Solso (Ed.), *Information processing and cognition: The Loyola Symposium* (pp. 55–85). Hillsdale, NJ: Lawrence Erlbaum Publishers.

Schellenberg, E. (1996). Expectancy in melody: Tests of the implication-realization model. *Cognition, 58,* 75–125.

Schellenberg, E. G. (1997). Simplifying the implication-realization model of melodic expectancy. *Music Perception, 14,* 295–318.

Schmuckler, M. (1989). Expectation in music: Investigation of melodic and harmonic processes. *Music Perception 7,* 109–150.

Schubert, E. (2001–2002). Correlation analysis of continuous emotional reponse to music: Correcting for the effects of serial correlation. *Musica Scientiae, Special Issue 2001–2002,* 213–236.

Simon, H. A., & Wohlwill, J. F. (1968). An experimental study of the role of expectation and variation in music. *Journal of Research in Music Education, 16,* 227–238.

Smith, N. A., & Cuddy, L. L. (2003). Perceptions of musical dimensions in Beethoven's Waldstein sonata: An application of tonal pitch space theory. *Musicae Scientiae, 7,* 7–34.

Tervaniemi, M., Huotilainen, M., Brattico, E., Ilmoniemi, R.J., Reinikainen, K., & Alho, K. (2003). Event-related potentials to expectancy violation in musical context. *Musica Scientiae, 7,* 241–262.

Tillman, B., & Bigand, E. (2004). Musical priming: Schematic expectations resist repetition priming. In S. D. Lipscomb, R. Ashley, R. O. Gjerdingen, & P. Webster (Eds.), *Proceedings of the Eighth International Conference on Music Perception and Cognition* (pp. 674–676). Adelaide, Australia: Causal Productions.

Tsang, L. (2002). Towards a theory of timbre for music analysis. *Musica Scientiae, 6,* 23–52.

Vos, P. G., & Troost, J. M. (1989). Ascending and descending melodic intervals: Statistical findings and their perceptual relevance. *Music Perception, 6,* 383–396.