

The generation of temporal and melodic expectancies during musical listening

MARILYN G. BOLTZ

Haverford College, Haverford, Pennsylvania

When listening to a melody, we are often able to anticipate not only *what* tonal intervals will occur next but also *when* in time these will appear. The experiments reported here were carried out to investigate what types of structural relations support the generation of temporal expectancies in the context of a melody recognition task. The strategy was to present subjects with a set of folk tunes in which temporal accents (i.e., notes with a prolonged duration) always occurred in the first half of a melody, so that expectancies, if generated, could carry over to an isochronous sequence of notes in the latter half of the melody. The ability to detect deviant pitch changes in the final variation as a function of rhythmic context was then evaluated. Accuracy and reaction time data from Experiment 1 indicated that expectancy formation jointly depends on an invariant periodicity of temporal accentuation and the attentional highlighting of certain melodic relations (i.e., phrase ending points). In Experiment 2, once these joint expectancies were generated, the temporal dimension had a greater facilitating effect upon melody recognition than did the melodic one. These results are discussed in terms of their implications for the perceptual processing of musical events.

When listening to a concert performance, we can often generate expectancies about the future course of a melody and anticipate the recurrence of certain themes and variations, the appearance of particular tonal intervals, and the ending of the composition itself. However, in addition to generating expectancies about the upcoming *what* of a melody, we can often anticipate *when in time* these melodic relations will occur. The generation of temporal expectancies during musical listening has not been investigated well in the previous literature and is the topic of interest here. In particular, two major questions are addressed: First, what types of structural relations in a melody's unfolding context support the generation of temporal expectancies? And second, to what extent are expectancies concerning the upcoming *what* and *when* of a melody—namely, melodic and temporal expectancies—interrelated with one another?

Expectancies in Cognition and Music

The concept of expectancy has long been of interest in the history of psychology and has been investigated in many different behavioral contexts. These include the areas of speech perception and discourse processing (Bartlett, 1932; Duffy, 1986; Ehrlich & Rayner, 1981; Mills, 1980), auditory and visual perception (Pachella & Miller, 1976; Posner & Snyder, 1975), and social cognition (Bel-

more & Hubbard, 1987; Hastie, 1980; Srull & Wyer, 1989). It is also a concept of particular interest to psychomusicologists. To a large extent, this interest stems from the work of Leonard B. Meyer (1956), which suggests that the confirmation and violation of expectancies is critical to aesthetic experience and the expression of emotional meaning in music. Meyer argues that movement toward points of tonal stability within a tune often creates a sense of satisfaction and musical resolution within listeners, but that artful deviations of expectancies are just as important for one's overall interest level and musical enjoyment. In his later work, Meyer has attempted to identify the melodic structures or "processes" that give rise to musical meaning (Meyer, 1967, 1973) and that have psychological validity in an empirical setting (Rosner & Meyer, 1982, 1986). Although this topic has continued to be of interest (e.g., Clarke, 1985), more recent research has been concentrated on the expectancy concept within a wider range of musical behavior.

As noted in a review by Schmuckler (1990), this more recent literature has been guided by two underlying goals. The first is to assess the impact of expectancies on cognitive processing abilities. This issue has been examined in several different contexts, including those of time estimation behavior (Boltz, 1989c, 1991b; Jones & Boltz, 1989), perceptual ratings (Boltz, 1989a, 1989b; Povel, 1985), selective attention (Dowling, Lung, & Herrbold, 1987; Jones, Kidd, & Wetzel, 1981), chord classification (Bharucha & Stoeckig, 1986, 1987), and musical restoration (DeWitt & Samuel, 1990). The second goal is much more fundamental and, in fact, predicates the first in that it is directed toward determining the factors that support the generation of listeners' expectancies. In general, most

This research was presented at the Fourth Rhythm Workshop on Rhythm Perception and Production, in Bourges, France, June 1992. The author thanks Andy McConnell for the collection of data and assistance with the data analysis process. Requests for reprints should be addressed to: M.G. Boltz, Department of Psychology, Haverford College, Haverford, PA 19041 (e-mail: m_boltz@acc.haverford.edu).

researchers agree that musical expectancies can stem from two primary sources that are not always independent of one another. One derives from long-term knowledge acquired from repeated exposure to a given musical event. For example, most people have listened to the tune "Happy Birthday" on so many occasions that they can anticipate the sequence of notes as it is heard once again. The second source of expectancies is more immediately influenced by the unfolding context of a melody. That is, as one listens to a tune, the unfolding arrangement of tonal intervals and chordal progressions often creates expectancies about what types of melodic relations should occur next. Unlike the first, these types of expectancies require listeners to extract invariant relations within a melody that may imply a continuation or reversal of this ongoing motion (Jones, 1990; Narmour, 1989). The aim of this research, then, is to identify the structural relations that support the generation of expectancies, in such a way that their underlying strength and specificity can be quantitatively assessed.

A variety of techniques have been implemented with these goals in mind. One common strategy is to rely on production tasks in which subjects are presented with an initial melodic context and asked to predict a future interval through singing (Carlson, 1981; Unyk & Carlson, 1987) or keyboard responses (Abe & Hoshino, 1990; Schmuckler, 1990). Carlson (1981), for example, presented listeners with a set of two-tone pitch intervals and asked them to sing their continuations. The results revealed that the subjects were more apt to produce small interval continuations (three semitones or less) as opposed to large ones, and that some melodic contexts produced stronger expectancies than did others—namely, the ascending major and minor second, and the ascending major and minor seventh. Unyk and Carlson (1987) later extended this research with a musical dictation task. For each individual subject, a set of expectancy responses was first obtained from the sung continuations of two-tone melodic contexts. These data were then used to create a set of melodies for each subject that displayed variations in the expectancy strength of certain pitch intervals as well as different types of expectancy confirmations/violations. The results showed that relative to the confirmation of strong expectancies, the frequency of dictation errors markedly increased for intervals of weak expectancy strength, or when a strongly expected interval size was violated. Expectancies concerning the underlying contour of a melody, on the other hand, exerted a negligible effect upon performance.

An elegant series of studies by Schmuckler (1989) later converged with these findings by revealing several additional factors that support the generation of expectancies. Instead of two-tone context intervals, subjects were presented with a Schumann composition, and at various stopping points they were asked to rate how well a given continuation tone conformed with their expectancies. Although contour once again was not an important factor, ratings did vary with the tonal stability of intervals (cf.

Krumhansl & Kessler, 1982) and displayed evidence for certain melodic processes such as continuation and reversal (Meyer, 1973; Narmour, 1989). In addition, when subjects were asked to continue the composition with keyboard responses (Schmuckler, 1989, 1990), the results showed that the preceding metrical structure of the piece was also a strong predictor of expectancy responses.

As a set, this previous research has been very useful in revealing what types of melodic contexts are able to generate the strongest and most specific expectancies, as well as the relative importance of different melodic dimensions. It is important to note, however, that this literature has primarily been directed toward melodic structure and expectancies concerning the upcoming *what* of tonal intervals. Although this is, in fact, a fundamental aspect of musical listening, it is also the case that listeners are often able to anticipate *when* in time a given tonal interval will occur. The concept of temporal expectancies has not been investigated well, and little is known about the factors that lead to the formation of temporal expectancies and how these are directed over the course of an unfolding melody. The purpose of the present research was to investigate these issues in the context of a melody recognition task and thereby consider how the inherent temporal structure of a melody and its joint relationship with melodic structure may give rise to temporally based expectancies.

Temporal/Melodic Structure of Music

Within the psychomusicology literature, the structure of music has been described from many different perspectives. The framework offered by Jones and Boltz (1989), however, is particularly relevant here, because it permits one to consider musical structure relative to the cognitive activities of perceiving and attending, which in some cases can be anticipatory in nature and based on the generation of temporal expectancies. The basic assumption is that melodies (and other natural events) vary along a continuum of structural coherence that influences a perceiver's mode of attending. At one end of this continuum are highly coherent events that display a high degree of predictability in which the event's temporal and nontemporal structure are interrelated through a hierarchical scheme. At the other end are less coherent events whose sequence of nontemporal information (i.e., notes) is temporally unpredictable. Although these two levels of structure are necessarily dependent on one another in order for an event to be characterized, each can be described in turn to illustrate its respective pattern of invariant relationships. Of the two, temporal structure is considered to be the more important dimension, because it is the vehicle that drives attending over the course of an unfolding event and, depending on its structural layout, can thereby guide attending toward or away from important nontemporal invariants.

In music, the temporal structure of a melody arises from the interplay of meter and rhythm. Meter supplies the underlying beat of a tune within a mensural framework.

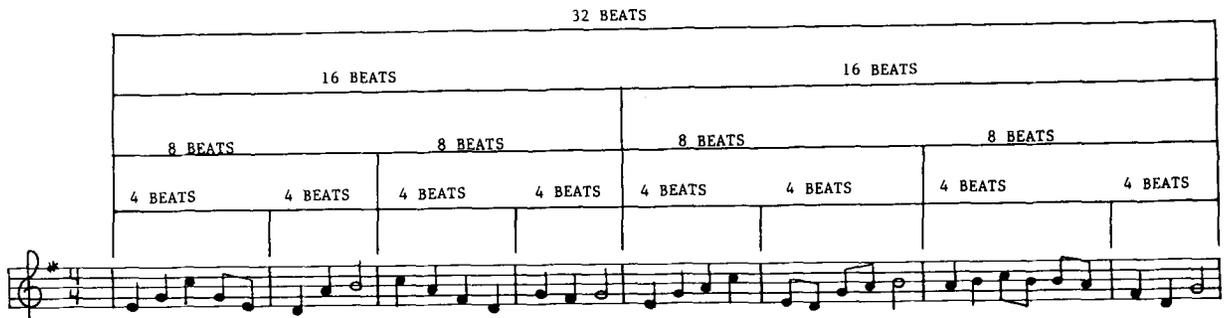


Figure 1. Example of a musical composition (excerpt from "There's Music in the Air" by G. Root) displaying a hierarchical arrangement of temporal/melodic structure. (Adapted from *The International Library of Music: Album of the World's Best Home Songs*, p. 124, by the Editorial Board of The University Society, The University Society Inc., New York, 1964.) This melody displays tonal resolution in that it nonarbitrarily begins and ends on the tonic note, G. Within this total time span, the underlying meter of the melody provides lower level periodicities that are nested within melodic phrases marked by tonic triad members (here, the tonic, G; the mediant, B; and the dominant, D). These phrase ending points are attentionally highlighted by temporal accents (J), and both regularly recur in time to outline higher order phrase relationships.

The melody shown in Figure 1, for example, is based on a 4/4 meter, where a quarter note (J) receives one beat of time and there are 4 beats per measure. Meter, by providing the referent beat period, not only functions as a lower level invariant, but because the first beat of a measure is perceived to be accented (Palmer & Krumhansl, 1990), also yields a set of nested periodicities within a melody's total time span. Rhythm, on the other hand, refers to durational changes within a melody, the pattern of *longs* and *shorts* that arise because some notes (e.g., J) receive fewer beats than the referent, whereas others receive more beats (e.g., J). Rhythm is psychologically important because notes that either follow a pause or are prolonged in duration create temporal accents that stand out and capture one's attention. For example, temporal accents in Figure 1 arise from the relatively longer half notes (J). In addition, given that temporal accents often recur after an invariant number of beats and terminate a measure, they provide a higher order level of nesting within the underlying metrical framework (Longuet-Higgins, 1976, 1978; Yeston, 1976). Thus, if an individual's attention is captured by temporal accents, this in turn will reveal the hierarchical arrangement of temporal structure.

Although a given melody may display the types of temporal relations just described, it does not emerge as a coherent entity unless its nontemporal arrangement of melodic structure is lawfully related to the hierarchy of temporal information. In many forms of Western music, this relationship does occur. Melodic structure is based on the underlying key of a melody, which often can be derived from the strategic placement of tonic triad members (i.e., the tonic, mediant, and dominant). The tonic interval not only is used to end a melody but, along with the mediant and dominant, often serves to mark phrase ending points within a melody's total time span (Kramer, 1982; Piston, 1978; N. Todd, 1985). Melodic phrases represent themes and variations before shifting to new ones and by marking the endings of these with tonic triad mem-

bers, provide momentary points of resolution within a tune and a reaffirmation of the underlying tonality scheme. As can be seen in Figure 1, melodic phrasing also contributes to a hierarchical arrangement of structure because pairs of phrases can be nested within higher level phrases of a longer periodicity (here, 16 beats) and thereby provide a lawful organization of melodic relations.

For a listener, the perceptual pickup of this organization is often facilitated through the temporal accentuation of salient melodic relations. In some cases, this may involve the accentuation of contour peaks or reversals as well as large melodic pitch skips (Boltz & Jones, 1986; Jones, 1981). One of the more common conventions, however, is the temporal accentuation of phrase ending points, which serves to attentionally highlight the melody's underlying tonality scheme and hierarchical organization of melodic relations (Boltz & Jones, 1986; Lerdahl & Jackendoff, 1983; Monahan & Carterette, 1985). This yields a joint accent structure in which the lawful and invariant recurrence of temporal accents can be used to guide attending over the melody's total time span and facilitate attentional tracking. In addition, this type of structural arrangement is assumed to afford a future-oriented mode of attending (Jones, 1990; Jones & Boltz, 1989) in which the regularly recurrent array of joint accents allows listeners to extrapolate the melody's preceding context and generate expectancies about the what and when of upcoming structural relationships. In Figure 1, for example, the invariant periodicity of temporal accents allows one to anticipate the beginnings and ends of upcoming melodic phrases. This includes expectancies about the melody's end, because relative to the preceding periodicities, it too recurs with an invariant beat period.

In sum, the temporal structure of a melody is assumed to play a primary role in the cognitive processing of melodic relations. When the pattern of temporal accentuation reinforces the hierarchical organization of melodic structure, it can be used to guide the course of perceptual pickup and enable listeners to generate expectancies

for anticipatory attending. At the same time, however, the temporal structure of a melody may also give rise to relatively incoherent melodies that hinder cognitive processing activities. This may occur when the arrangement of temporal or melodic structure alone is relatively unpredictable in nature or, alternatively, when the array of temporal accents conflicts with the organization of melodic relations. In either case, these types of melodies are very difficult to attentionally track because the listener is confronted with a seeming array of disjointed items—the melody lacks a unified organizational scheme and appears to contain unrelated fragments of information. Moreover, incoherent melodies fail to support an anticipatory mode of attending because there are no lawfully recurrent relations that can be extrapolated into future time.

These predicted effects of event coherence upon cognitive processing abilities have received some empirical support. In a recent study by Boltz (1991a), subjects were asked to recall a set of folk tunes in which melodic relationships either recurred with an invariant periodicity to reveal an underlying hierarchical organization of temporal/melodic structure or were irregularly timed to yield a more incoherent melody. The results showed that coherent melodies yielded a much higher level of performance than did incoherent ones. In addition, subsequent error analyses indicated that recall was markedly superior at phrase ending points, suggesting that these served as anchors for remembering and facilitated the retrieval of lower order relations among notes. In contrast, incoherent patterns of structure yielded a very high incidence of missing notes and fragmented relations among adjacent intervals. Similar effects have also been observed for the remembering of auditory event durations (Boltz, 1992). When subjects were asked to either recognize or extrapolate the duration of melodic patterns that had been learned in an incidental fashion, those displaying a coherent arrangement of structure yielded high levels of accuracy that increased with greater learning experience. The duration of incoherent patterns, on the other hand, were remembered at only a chance level of performance, regardless of the amount of learning experience. This research, then, suggests that a coherent array of temporal/melodic information offers a particularly efficient scheme for perceiving and remembering, because the arrangement of temporal accents can be used to guide attending toward the hierarchical organization of nested relations. As a set, these studies also suggest that temporal and melodic relations are inextricably entwined on both a structural and a behavioral level in such a way that manipulations of one dimension exert an impact on the other.

Evidence for expectancy generation within coherent patterns is less straightforward and has primarily received indirect support. Some evidence is found in the study of time estimation behavior. Jones and Boltz (1989), for example, assessed the experienced duration of folk tunes in which the unfolding array of temporal accents was intended to create expectancies about when in time the last

note of a melody would occur. The results suggested the mediation of temporal expectancies, in that tunes ending earlier than expected yielded underestimations of a tune's total duration whereas those ending later than expected produced systematic overestimations. In contrast, relatively accurate time judgments were observed for tunes ending as expected (i.e., on time). These findings have been corroborated in subsequent studies (e.g., Boltz, 1989c, 1991b; Jones, Boltz, & Klein, 1992) in which there were more rigorous controls for alternative explanations of the data, and they have been extended by research suggesting the role of temporally based expectancies in the timed extrapolation of a melody's final note (Boltz, 1992; Jones & Boltz, 1989; Jones et al., 1992). Other research has also provided indirect support for expectancy generation in the contexts of perceptual rating tasks (Boltz, 1989a, 1989b; Povel, 1985), the detection of hidden melodies from a background of distractor tones (Dowling, 1990; Dowling et al., 1987), and selective attention to musical relationships (Jones et al., 1981).

Although these studies do imply the mediation of expectancies in cognitive behavior, they are in fact indirect tests of this construct in that the experimental methodology does not include a means for ensuring whether listeners are *actively* extrapolating temporal relations from a melody's preceding context. A more direct approach might rely on production tasks, such as those used in the study of melodic expectancies (Carlson, 1981; Schmuckler, 1990; Unyk & Carlson, 1987), which would allow one to examine rhythmic accent structure in sung or keyboard continuations. Alternatively, one might manipulate the pattern of temporal accentuation in a melody's preceding context to determine whether this differentially gives rise to expectancies that carry over to an isochronous sequence of notes. Unfortunately, such research is scarce within the literature and has yielded inconclusive findings. Schmuckler (1990) found that listeners were able to continue the underlying metrical structure of a tune but did not examine the extent to which a melody's rhythmic and temporal accent structure was extrapolated from the preceding context. Using the second type of method, Jones, Boltz, and Kidd (1982) investigated the influence of temporally based expectancies in a melody recognition task. In this study, nine-tone melodies were generated from a set of mathematical symmetry rules in which one of four types of rhythms (isochronous, irregular, dactylic, anapestic) was imposed on the initial three-tone argument of each tune. This manipulation was designed to promote temporal expectancies that then carried over to the following three tones, which were always isochronously timed. The results indicated that relative to the isochronous and irregular rhythmic contexts, listeners were more apt to detect a deviant pitch change on the fourth note if the preceding rhythm was dactylic (i.e., AUU, where A and U refer to accented and unaccented notes) and a deviation on the sixth note if the preceding rhythm was anapestic (i.e., UUA). Although these findings do suggest the generation of temporal expectancies, a later study

by Monahan, Kendall, and Carterette (1987) cast doubt on this interpretation, because other structural relations within the melodies of Jones et al. (1982) may have contributed to recognition performance.

In sum, there is a need for systematic examination of the process of temporal expectancy generation and how anticipatory attending is directed over the course of an unfolding melody. In addition, it is necessary to determine what structural relations within the initial context of a tune support the formation of temporal expectancies, as well as the extent to which these are entwined with melodic structure and expectancies concerning the upcoming *what* of tonal relations.

Rationale of the Present Research

The purpose of the present experiments was to investigate these issues within the context of a melody recognition task. In many ways, the experimental method of this research is reminiscent of that of Jones et al. (1982) but with several important refinements. In lieu of presenting subjects with rule-constructed melodies of a limited duration and melodic context, the subjects in the present experiments heard a set of unfamiliar folk tunes, an ex-

ample of which is shown in Figure 2A. In the initial context of each, there were always three melodic phrases, the ending points of which were marked by a tonic triad member. This was then followed by a final variation, consisting of two melodic phrases, in which all notes had an isochronous rhythm. The basic strategy was to vary the pattern of temporal accentuation within the melody's initial context in order to establish temporal expectancies that would carry over to the final variation of isochronous notes.

These manipulations were designed to assess whether the temporal and melodic structure of a melody exert an independent or interactive influence on the generation of temporal expectancies. An interactive relationship indicates that manipulations of one dimension affect the cognitive processing of the other. This has previously been found for the remembering (Boltz, 1991a, 1992; Boltz & Jones, 1986; Deutsch, 1980; Drake, Dowling, & Palmer, 1991), experienced duration (Boltz, 1989c, 1991b; Jones & Boltz, 1989), and perceptual judgments (Boltz, 1989a, 1989b) of musical events. If a similar principle also applies to expectancy generation, the formation of temporal expectancies may not simply depend on a lawful pattern

LOOK OUT, HOW IT'S RAINING by L. Speyer (Austria)

A. Accentuation of Tonic Triad Intervals with an Invariant 10 Beat Periodicity



B. Accentuation of Tonic Triad Intervals with a Variant (8,10,12) Beat Periodicity



C. Accentuation of Non-Tonic Triad Intervals with an Invariant 10 Beat Periodicity



D. Accentuation of Non-Tonic Triad Intervals with a Variant (8,10,12) Beat Periodicity



Figure 2. An exemplar melody from Experiment 1. Panels A and C illustrate conditions in which tonic and nontonic triad intervals, respectively, were temporally accentuated (\downarrow) with an invariant periodicity that might support the generation of expectancies (shown as arrows) that carry over to the final variation of isochronous notes. Conversely, panels B and D depict the accentuation of tonic and nontonic triad intervals that occur with a variant periodicity and are therefore less likely to support temporal expectancies. TT, NT, D1, and D2 indicate the loci of pitch deviations, and the asterisk above each indicates the new note that occurred on the *different* trials of the experiment.

of temporal structure (e.g., invariant periodicity of accents) but on the kind of tonal interval that temporal accents coincide with. In particular, one strong candidate is the set of tonic triad members marking phrase ending points, since these highlight the primary meaning and organization of a tune and have previously been found to act as cognitive referents in music perception and remembering (Boltz, 1989b, 1991a; Schmuckler, 1990; Sloboda & Gregory, 1980). The other possibility is that the temporal and melodic structure of a tune maintain an independent relationship in the generation of expectancies. That is, each dimension may contribute to expectancy formation, but their interactive influence is not required. If this is the case, the generation of temporal expectancies may simply arise from the invariant timing of temporal accents, regardless of the type of tonal interval that is accentuated.

These two alternatives were assessed in the present experiment by imposing two types of manipulations on the pattern of temporal accentuation within a melody's initial context. The first manipulation involved the kind of tonal interval that was accentuated. In half the tunes that all subjects heard, temporal accents always coincided with tonic triad members marking phrase ending points, whereas in the remaining half, temporal accents always coincided with nontonic triad members occurring within a phrase. These are illustrated in Figures 2A-2B and 2C-2D, respectively. The second manipulation involved the regularity of temporal accentuation. In one condition, temporal accents always recurred after an invariant number of beats, whereas in a second, the number of beats between temporal accents varied in such a way that there was an overall irregular periodicity. The basic prediction was that an invariant periodicity of temporal accents should support the generation of temporal expectancies that would carry over to the isochronous notes in the final variation. In Figure 2A, for example, given that the periodicity between accent onsets is always 10 beats, attention should be targeted toward the 9th beat in the final variation where the *onset* of a temporal accent would occur if it was actually to appear. In contrast, a variant periodicity of beats should not support expectancy generation, because there is no invariant temporal relation that can be extrapolated to the melody's upcoming notes. The primary question, however, was whether the generation of temporal expectancies would depend on melodic structure and only occur when phrase ending points were accented, or whether melodic structure would be irrelevant and simply occur whenever there was an invariant number of beats between accents in the preceding context. This was assessed by manipulating the locus of the pitch deviation in a melody recognition task. In this task, subjects heard a standard melody followed by a comparison, and their task was to decide whether the comparison tune had exactly the same pitch intervals as the standard or whether there was a pitch deviation (i.e., *same/different* judgment). On half the trials, the comparison was in fact the same as the standard, but on the other half, there was a single pitch deviation that could occur in one of four locations.

As can be seen in Figure 2, two were in the final variation containing the sequence of isochronous notes, one being a tonic triad interval marking a phrase ending point that occurred with the same temporal periodicity as that of the accentuation of the preceding tonic triad intervals (TT locus). The other was a nontonic triad interval that occurred with the same periodicity as that of the preceding accentuation of nontonic triad intervals (NT locus). The remaining two deviations (D1, D2) occurred in the melody's initial context and were included to prevent listeners from simply attending to the final variation. Again, the prediction was that an invariant periodicity of accents should support expectancies that would carry over to attentionally highlight certain tonal intervals in the final variation, and thereby facilitate the detection of any deviant pitch changes that might occur at these locations. If temporal and melodic structure maintain an independent relationship, this facilitating effect should occur for any deviation that was temporally predictable from the preceding context, regardless of what tonal interval was regularly accentuated. Hence, pitch changes at the TT locus should be detected more accurately when tonic triad members were accentuated, and the NT locus should produce better performance when nontonic triad members were accentuated. An interactive relationship, on the other hand, would reveal that performance was only facilitated in the condition in which meaningful pitch intervals were accented—namely, tonic triad members marking phrase ending points.

EXPERIMENT 1

Method

Subjects and Design. The design was a $2 \times 5 \times 2$ mixed factorial. The subjects listened to a series of folk tunes that varied in the type of tonal interval that was temporally accentuated (tonic triad or nontonic triad) and the locus of a pitch deviation point (TT, NT, D1, D2, *same*). The periodicity between accent onsets (invariant, variant) was the single between-subject factor.

Thirty-two subjects from an introductory psychology course at Haverford College, all with normal hearing, participated in the experiment in return for course credit. Each had at least 4 years of musical experience within the past 6 years and was currently playing a musical instrument.

Stimulus materials. A set of folk tunes from other countries was selected from a musical composition book.¹ To ensure that these melodies were unfamiliar to listeners, a preliminary rating study was conducted with an independent group of 12 experienced musicians. Any tune that was judged even slightly familiar to the subjects was eliminated from the experiment. The six melodies selected as experimental stimuli are depicted in Appendix A.

These melodies represent a fairly diverse range of music in that they vary in key (C, F, G, D major), meter (2/4, 3/4), and overall pitch contour. In their original composition, however, all melodies displayed a similar structural arrangement in that they opened with three melodic phrases in which the ending of each was marked by a tonic triad member (i.e., tonic, mediant, and dominant) that was prolonged in duration (by a half note, ♭). This was then followed by a final variation in which an isochronous sequence of notes was grouped into two melodic phrases.

This set of melodies was then manipulated along three different dimensions. The first involved the periodicity of temporal accen-

tuation. In their original versions, all melodies displayed an *invariant* number of beats (i.e., 8, 10, or 12 beats) between temporal accents within the tunes' initial context. Figure 2A, for example, depicts an exemplar melody from the experiment in which temporal accents always recurred after 10 beats. A second version of each melody was then constructed so that the number of beats between accents *varied* to yield an overall irregular periodicity of temporal accentuation. This was achieved by manipulating the rhythmic patterning within a melodic phrase while preserving the melody's original meter. Figure 2B provides an illustration of this variation, wherein the periodicity between temporal accents is 8, 10, and 12 beats, respectively. For all melodies, tonal duration (on-time) and the intertone interval (off-time) were held constant so that all notes corresponded to the following values: quarter notes (♩), 200 msec on, 100 msec off; eighth notes (♪), 100 msec on, 50 msec off; and half notes (♭), 400 msec on, 200 msec off.

A second manipulation involved the kind of tonal interval that was temporally accentuated in a melody's initial context. For the original version of each melody, in which temporal accents always coincided with *tonic triad* members marking phrase ending points, a second version was created in which temporal accents coincided with a *nontonic triad* note within a melodic phrase. This interval was always two to four notes away from the true phrase ending point. The temporal periodicity between tonic triad and nontonic triad accents was either an invariant or variant one, and once again, this required manipulating the rhythmic patterning of notes within a phrase while preserving the melody's original meter.

The final variation involved the pitch deviation loci for the melody recognition task. For each of the 24 tunes described, four types of pitch deviations were successively applied to yield a total of 96 experimental melodies. As can be seen in Figure 2, two of these deviations occurred in a melody's final variation of isochronous notes. One, termed the TT locus, coincided with the tonic triad member marking the first phrase ending point and occurred with the same temporal periodicity as did the *accent onsets* of the preceding tonic triad intervals. The other was a nontonic triad interval that occurred with the same periodicity as did the accentuation of the preceding nontonic triad intervals (NT locus). The remaining two deviations (D1, D2) occurred in a melody's initial context and were included to encourage attending to an entire melody and prevent listeners from simply focusing on the final variation of notes. Given the lengthy duration of each melody, all pitch deviations were designed to be fairly salient ones that always broke the melody's original contour and consisted of a four-scale step change. All deviations, however, were tonal and preserved the melody's original key.

Melodies were randomized into four sets (A_i, A_v, B_i, B_v), each of which contained 96 trials grouped into two blocks of 48. Set A melodies contained six melodies, half in which tonic triad members were accentuated in a melody's preceding context ("Reapers on the Mountain"; "A Love Song"; and "Why, Oh Mother") and half in which nontonic triad intervals were accentuated ("Spite"; "I'll Have No Other"; and "Look Out, How It's Raining"). Set B melodies contained the same set of six melodies, but the tunes that accentuated tonic triad members in Set A now accentuated nontonic triad members, and vice versa. The melodies within Sets A_i and B_i all displayed an invariant periodicity of temporal accentuation, while those within the A_v and B_v sets displayed a variant, irregular periodicity of accentuation. Within each set, each experimental trial consisted of a standard melody followed by a comparison. On half of the trials, the comparison melody contained exactly the same sequence of notes as did the standard. On the remaining half of the trials, however, the comparison contained a single pitch deviation, occurring twice at each of the four locations (i.e., TT, NT, D1, D2). Last, two different counterbalance orders were created for each of the four melody sets, so that the 96 trials within each set occurred in a different random order.

Apparatus. All melodies were constructed and generated with the MIDILAB software system (R. E. Todd, Boltz, & Jones, 1989).

During each experimental session, the melodies were presented on line with a Yamaha TX81Z FM tone generator controlled by an IBM AT computer with a Roland MPU-401 Midi interface unit. Sequences of tones were amplified by a Kenwood KR-4010 receiver and played over Koss-Pro 4AAA Plus headphones at a comfortable listening level. The Midi voice (i.e., timbre) of all the melodies was that of a percussive flute.

Procedure. The subjects were randomly assigned to one of the four melody sets and one of the two counterbalance orders. Recorded instructions informed the subjects of the details of pattern presentation and the task requirements. On each trial, a 1-sec warning tone (C8) preceded a standard melody by 2 sec. Two seconds later, the comparison melody of the pair was presented and immediately followed by a 5-sec response period. As they listened to the comparison melody, the subjects were asked to decide whether it contained the same sequence of notes as did the standard or whether it contained a pitch deviation. They were asked to indicate this judgment on a response console, as quickly as possible, by pressing a button on the right or left for *same* and *different*, respectively. The computer automatically recorded the reaction time and *same/different* response on each trial.

The subjects were tested in small groups of 2-4 individuals. Before each experimental session, the subjects received 6 practice trials that consisted of the six experimental melodies presented in a *same* or *different* pairing. Each experimental session was approximately 1.5 h long, with a 5-min rest break provided after the first 48 trials.

Results and Discussion

The recognition accuracy and reaction time data were analyzed by separate analyses of variance. Each will be discussed in turn.

Recognition accuracy. Figure 3 depicts mean percent-correct responses as a function of the experimental manipulations. Means are collapsed over melody set and counterbalance order, since these contributed negligible effects ($F < 1.0$).

The primary finding was a three-way interaction between accent periodicity, type of tonal interval that was accentuated, and the pitch deviation loci [$F(4,120) = 17.84$, $MS_e = 189.42$, $p < .001$]. Let us first consider the conditions in which tonic triad members, marking phrase ending points, were accentuated in a melody's preceding context. When these intervals were accentuated with an invariant periodicity, performance was quite high for *sames*, and more importantly, listeners were also quite accurate at detecting a pitch deviation at the location that was temporally predictable from the preceding context—namely, the TT locus, which was a phrase ending point itself. A set of Bonferroni post hoc comparisons indicated that the remaining types of deviation were detected at a significantly lower level of accuracy and, in fact, hovered around chance performance ($p < .01$). The facilitating effects observed in this condition were not simply due to the attentional highlighting and subsequent attunement to phrase ending points. If true, a similar pattern of findings would have emerged in the variant condition, in which tonic triad members were also accentuated. As can be seen in Figure 3, however, overall accuracy was not only much lower here, but performance at the TT locus was no different from that of the other deviation points. In sum, these results suggest that the recognition of pitch deviations does not simply depend on expectancies about

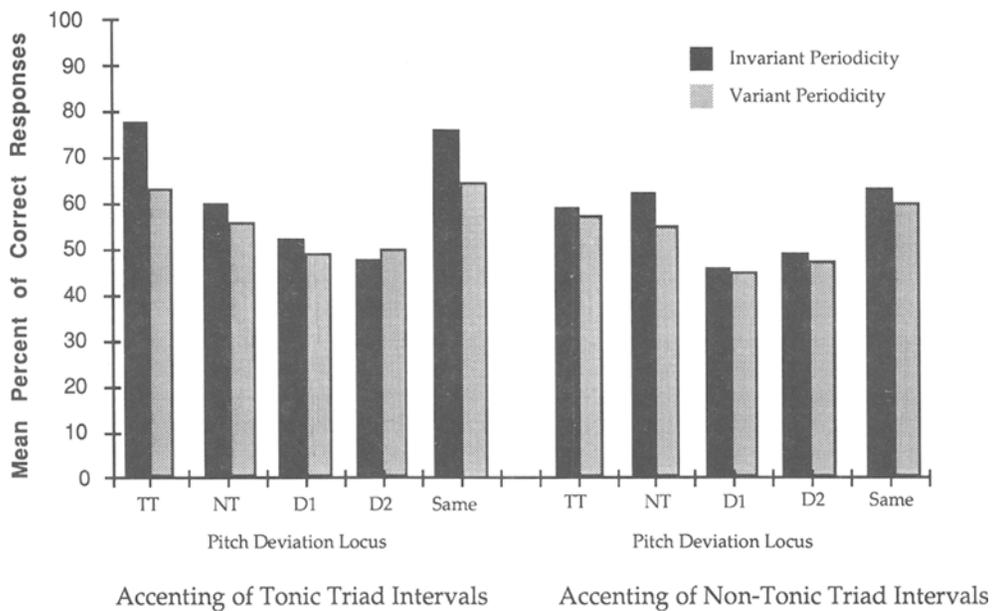


Figure 3. Mean percent-correct responses in Experiment 1, as a function of accent periodicity, type of tonal interval that was accentuated, and locus of the pitch deviation.

what tonal interval might be deviated, but instead is facilitated by expectancies about *when* in time this deviation might occur.

The next issue is whether these effects also generalized to situations in which nontonic triad intervals were accented with an invariant periodicity. As can be seen in Figure 3, they did not. Performance at the NT locus, which was temporally predictable from the preceding context, was not enhanced in relation to the TT locus in this condition or the NT locus when tonic triad intervals were accentuated. In addition, a set of Bonferroni post hoc comparisons revealed that performance in the invariant condition was indistinguishable from that in the variant condition, and this was true for both *sames* and *differents*.

The second finding to emerge from this initial analysis was a significant main effect for the pitch deviation loci [$F(4,60) = 7.89$, $MS_e = 74.57$, $p < .001$]. As indicated in Figure 3, deviations appearing in the final variation of a melody (i.e., TT and NT loci) were more likely to be detected than those appearing in a melody's initial context (D1 and D2). Given that there was an equal probability for the occurrence of each deviation, this most likely reflects a recency effect such that those deviations in the latter part of a melody were simply more salient in memory.

To confirm that these overall findings were not simply due to response bias effects, a d' analysis was also conducted from a procedure reported by Hochhaus (1972). In this analysis, hit rate (proportion of *same* trials that were correctly recognized) is evaluated in relation to false recognition rate (proportion of TT, NT, D1, D2 trials that were *each* misrecognized as *same*), and these are transformed into d' and bias values with the use of the abscissa and ordinate values of the standardized normal distribu-

tion. Values of 0 represent random guessing, while values of 4.65 reflect perfect discrimination accuracy. Bias scores independently estimate whether subjects are differentially inclined toward *same* or *different* responses. Thus, values of 1.00 reflect no bias effects, whereas values > 1.00 reveal a bias toward *different* responses, and values < 1.00 reveal a bias toward *same* responses. The results of this analysis converged with the percent-correct response data and revealed the same pattern of findings.

Reaction time. Figure 4 depicts the mean reaction time for correct responses in each experimental condition. An overall analysis of variance (ANOVA) revealed that the results of this analysis closely converged with the recognition accuracy data. As before, there was a significant three-way interaction between accent periodicity, type of tonal interval that was accentuated, and the pitch deviation loci [$F(4,120) = 14.09$, $MS_e = 215,132.81$, $p < .001$]. The source of this interaction can be traced to the fact that reaction time was fastest when accents recurred with an invariant periodicity and coincided with tonic triad members marking phrase ending points. In particular, this is the only condition in which reaction time was facilitated by temporal expectancies. A set of Bonferroni post hoc comparisons indicated that the TT locus, which was temporally predictable from a melody's preceding context, yielded faster response times than the NT locus, which also occurred in the final variation, and the D1 and D2 loci occurring in a tune's initial context. This facilitating effect did not occur in the variant condition, nor in those conditions in which accents coincided with nontonic triad intervals.

The second result from this data involved a significant main effect for pitch deviation loci [$F(4,60) = 18.02$, MS_e

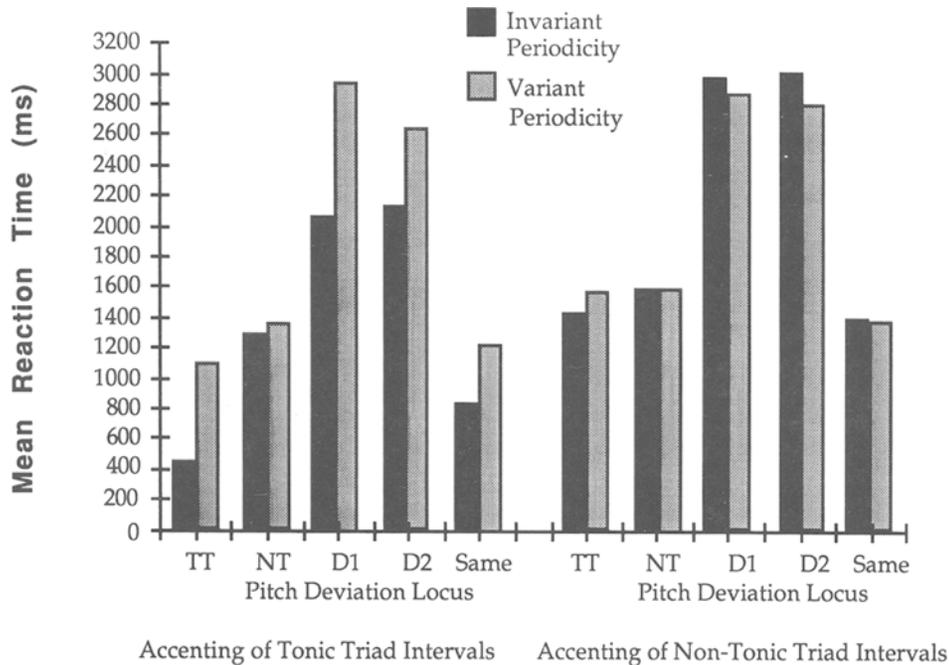


Figure 4. Mean reaction time (in milliseconds) for correct responses in Experiment 1, as a function of accent periodicity, type of tonal interval that was accentuated, and locus of the pitch deviation.

= 243,764.31, $p < .001$]. Across all conditions of the experiment, deviations appearing in a melody's final variation (i.e., TT, NT loci) were correctly detected at a faster rate than were those appearing in a melody's preceding context (i.e., D1, D2). This finding also converges with the accuracy data and further suggests that the most recent deviations were more accessible in memory.

In sum, the results of this experiment illustrate that melody recognition performance can be facilitated by the mediation of temporal expectancies. Given that the isochronous sequence of notes within a melody's final variation always remained constant and it was only the preceding rhythmic context that varied, this suggests that the unfolding structural relations within a tune can establish a course of attending that is future-oriented in nature. A lawfully recurrent pattern of temporal accentuation allows listeners to extrapolate a melody's preceding context so that attention is directed toward particular loci in future time and the corresponding nontemporal information that is there (i.e., notes). The present set of results, however, also indicates that the generation of expectancies is dependent on the interactive influence of temporal and melodic structure. Expectancies do not arise from the mere accentuation of meaningful tonal intervals (i.e., phrase ending points) or from an invariant periodicity of accentuation. Instead, a joint relationship between these two structural dimensions must exist in order to support a future-oriented mode of attending.

EXPERIMENT 2

One issue to consider is whether a joint relationship between temporal and melodic structure is simply necessary

for the formation and development of expectancies, or whether this joint influence continues to exist *after* expectancies have been generated. That is, are the temporal and melodic components of expectancies equally weighted so that expectancies concerning the upcoming what of a melody are as strong as those concerning the upcoming when? Or is it the case that one type of expectancy is stronger than the other, so that greater attention is directed toward a particular kind of tonal interval versus a particular point in future time (or vice versa)? Experiment 1 does not allow us to address this issue, because the only deviation point that was facilitated by the preceding rhythmic context of a melody (i.e., the TT locus) was one that jointly confirmed temporal and melodic expectancies—it not only appeared with the same temporal periodicity as in the preceding context but was a tonic triad interval itself that marked a phrase ending point. However, suppose that melodic and temporal expectancies were dissociated from one another so that the pitch deviation points confirmed one type of expectancy but not the other. Is one type of deviation detected faster and more accurately than the other, or do they produce a comparable level of accuracy? Experiment 2 was designed to answer this question, once again with the use of a melody recognition task.

Method

The method of this experiment was identical to that of Experiment 1, but for the following exceptions.

Subjects and Design. The design was a 2×6 repeated measures factorial. The subjects were presented with a set of folk tunes that varied in their type of expectancy confirmation (joint, dissociated) and the locus of a pitch deviation point (T, M, C, D1, D2, *same*).

Twenty-four musically sophisticated subjects from an introductory psychology course at Haverford College participated in the

A. DISSOCIATED EXPECTANCY CONFIRMATION

Near Krakow by W.A. Roberts - 8 Beat Periodicity



B. JOINT EXPECTANCY CONFIRMATION

Lithuanian National Hymn by W. Griffith - 9 Beat Periodicity



Figure 5. Exemplar melodies from Experiment 2. In panel A, temporal and melodic expectancies that carry over from the initial context are dissociated from one another. The T locus is temporally predictable from the preceding context but does not coincide with a phrase ending point. It thereby confirms temporal expectancies but not melodic ones. In contrast, the M locus confirms melodic expectancies but not temporal ones, because it coincides with a phrase ending point but is not temporally predictable from the preceding context. Deviations at C, D1, and D2 confirm neither type of expectancy and thereby act as controls. In panel B, expectancies concerning the upcoming what and when of tonal intervals are jointly confirmed by the TM locus, which is temporally predictable from the preceding context and coincides with a phrase ending point. Deviations at C1, C2, D1, and D2 all act as controls by confirming neither type of expectancy. The asterisk at each deviation point indicates the pitch change on *different* trials.

experiment in return for course credit. Each had normal hearing and had not participated in the previous experiment.

Stimulus materials. In addition to the six melodies used in Experiment 1, two additional folk tunes were selected from the same musical composition book (see Note 1) and preterated as unfamiliar to a group of 12 experienced listeners. In the initial context of each, temporal accents (i.e., half notes) always recurred with an invariant periodicity and coincided with tonic triad members marking phrase ending points. Once again, this was followed by a final variation of isochronous notes that were grouped into two melodic phrases.

This total set of stimuli was divided into two melodic sets, each containing four tunes, that varied in their type of expectancy confirmation. In one set, expectancies about the upcoming what and when of tonal intervals that should carry over from the initial context were *dissociated* and failed to coincide with one another. This is apparent from the manipulation of the pitch deviation points shown in Figure 5A. One, the T locus, was a tonal interval that appeared with the same temporal periodicity as that in the preceding context but was not a tonic triad interval marking a phrase ending point. It therefore confirmed temporal expectancies but disconfirmed melodic ones. A second, the M locus, did coincide with a tonic triad interval marking a phrase ending point but did not recur with the same temporal periodicity as that in the preceding context. Hence, this locus confirmed melodic expectancies but not temporal ones. The location of these deviations were counterbalanced across the set of tunes so that the M locus occurred before the T locus for two of the folk tunes and after the T deviation for the others. The remaining three deviation loci were controls: one, the C locus, also occurred in the final variation of a tune but corresponded to neither a phrase ending point nor a temporally expected locus. As before, D1 and D2 occurred in a melody's preceding context to

simply enforce attending to an entire tune. The second set of melodies differed in that the tonal interval that was temporally expected was also a tonic triad member marking a phrase ending point (TM locus)—it thereby *jointly* confirmed both temporal and melodic expectancies. As can be seen in Figure 5B, a pitch deviation could occur at this point as well as at four other loci that all acted as controls by confirming neither type of expectancy. Two of these occurred in the final variation (C1, C2) whereas the two remaining appeared in the preceding context of a melody (D1, D2). The C1 deviation always occurred before the TM locus, whereas the C2 deviation always occurred after the TM locus.

Three of the four melodies that occurred in the joint expectancy set were identical to those of Experiment 1: "Why, Oh Mother"; "Spite"; and "A Love Song." They were written, respectively, in the keys of C, D, and F major, and they displayed a 12-, 10-, and 8-beat periodicity between temporal accents (see Appendix A). The fourth, the "Lithuanian National Hymn," was based on the G major diatonic scale and showed a 9-beat periodicity between temporal accents (see Figure 5A). As is depicted in Appendix B, three of the four tunes in the dissociated expectancy set were also adapted from Experiment 1 ("I'll Have No Other"; "Look Out, How It's Raining"; and "Reapers on the Mountain"). Although these melodies contained the same sequence of notes, and were in the same key and meter as in Experiment 1, their original rhythm was altered so that there were 10, 12, and 9 beats, respectively, between temporal accents. The fourth tune, "Near Krakow," was based on the key of D major and showed an 8-beat periodicity between accents.

Within each of the two expectancy sets, the melodies were randomized into a series of 80 trials that consisted of two blocks of 40 standard/comparison pairs. There were equivalent numbers of

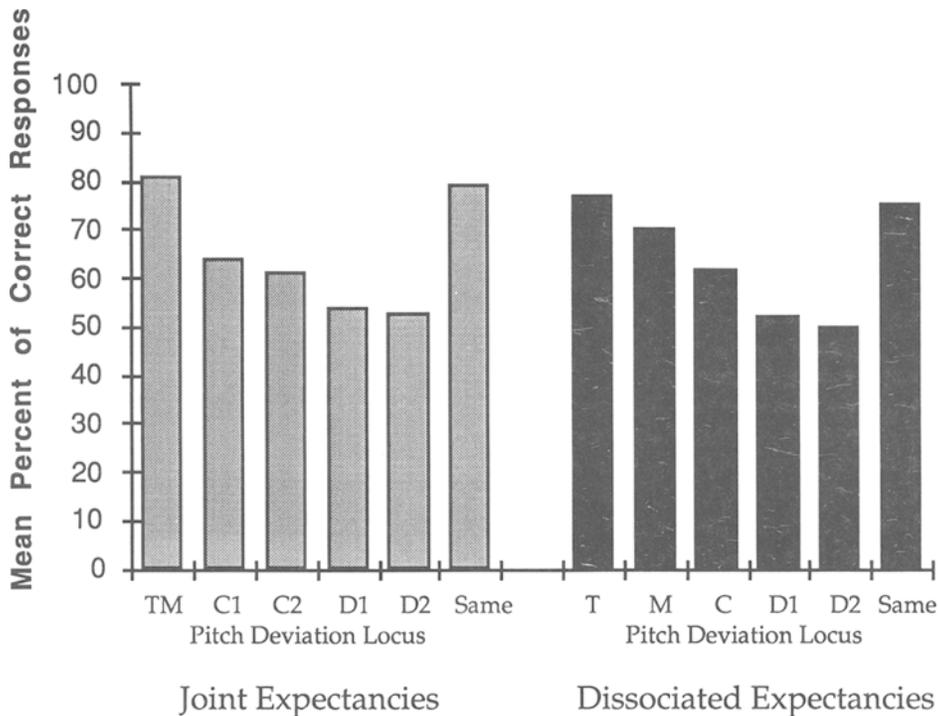


Figure 6. Mean percent-correct responses in Experiment 2, as a function of the pitch deviation locus and type of expectancy confirmation.

same and *different* trials that resulted from each of the four folk tunes being presented twice at each of their five deviation points and 10 times with the *same* version of their comparison melody. As before, two different counterbalance orders were prepared for each of the two melody sets, so that the 80 trials within each appeared in a different random order.

Procedure. As in Experiment 1, the subject's task on each trial was to make a *same/different* judgment as quickly as possible. Each experimental session was approximately 1.25 h long, with a 5-min rest break after the first 40 trials. This study was conducted over the course of 2 days, with melody set and counterbalance order blocked by day. Before each daily session, the subjects received 4 practice trials that consisted of the four different tunes within a set being presented in a *same* or *different* melodic pairing.

Results

Recognition accuracy. Figure 6 depicts the mean percent of correct responses as a function of the type of expectancy confirmation and the locus of the pitch deviation. An overall ANOVA revealed a significant interaction between these two variables [$F(5,115) = 23.09$, $MS_e = 174.03$, $p < .001$]. In the joint expectancy condition, the results from Experiment 1 were replicated, and they indicated that listeners were most accurate at detecting a deviation that was both temporally and melodically predictable from a melody's preceding context (i.e., TM locus). A set of Bonferroni post hoc comparisons confirmed that performance at this location was significantly superior to that at the remaining pitch deviation loci ($p < .01$). The more interesting results were found within the dissociated condition. As can be seen in Figure 6, listeners

were most accurate at detecting a pitch deviation that occurred with the same temporal periodicity as in the preceding context and thereby confirmed temporal expectancies (i.e., T locus). A set of post hoc comparisons revealed that performance at this location was comparable to the TM locus in the joint expectancy condition, and significantly higher than detectability at the M locus, which corresponded to a phrase ending point and thereby confirmed melodic expectancies ($p < .05$). Nonetheless, performance at the latter location was significantly higher than that at the control loci, which confirmed neither type of expectancy. These results, then, suggest that expectancies about the what and when of upcoming pitch intervals facilitate melody recognition, but that temporal expectancies are the stronger of the two.

Once again, there was a significant main effect for the pitch deviation loci [$F(5,22) = 14.31$, $MS_e = 162.81$, $p < .001$] such that listeners were more accurate at detecting a pitch change in a melody's final variation than in the initial context. As in Experiment 1, d' and bias analyses were also conducted and revealed the same pattern of results as the percent-correct response data.

Reaction time. The mean response time for correct responses is shown in Figure 7. The overall ANOVA indicated that these data were consistent with the recognition accuracy scores and revealed the same pattern of results. The interaction between pitch deviation loci and type of expectancy confirmation was significant [$F(5,115) = 11.92$, $MS_e = 1,182.71$, $p < .001$], as was the main effect for

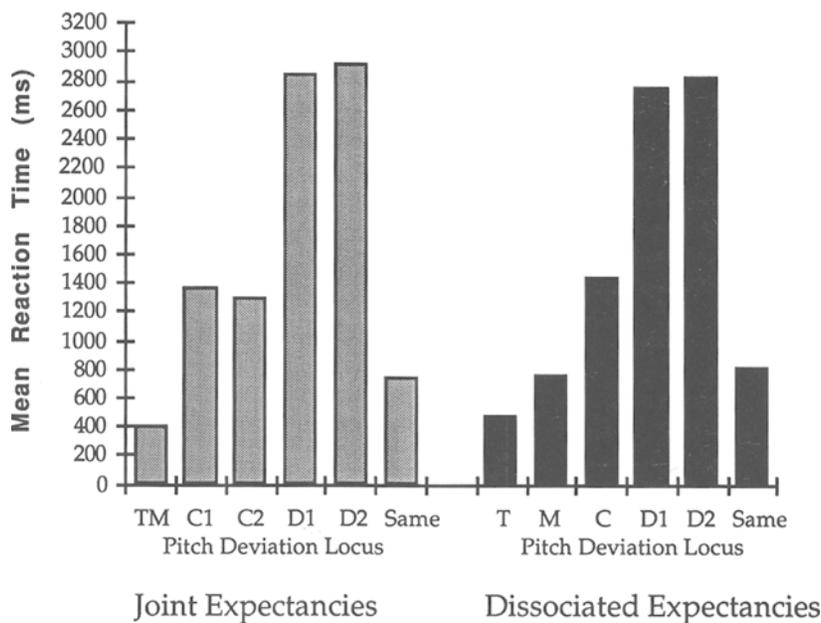


Figure 7. Mean reaction time (in milliseconds) for correct responses in Experiment 2, as a function of the pitch deviation locus and type of expectancy confirmation.

the pitch deviation loci [$F(5,22) = 16.84$, $MS_e = 1,159.87$, $p < .001$].

GENERAL DISCUSSION

The results from the present set of experiments reveal evidence for the generation of temporal expectancies during musical listening. In contrast to previous studies, in which the mediation of this construct has been tacitly assumed, the present research relied on a technique that enables listeners to actively extrapolate temporal invariants into future time. By systematically manipulating the temporal accent structure of a melody's initial context, one can then determine the structural relations that initiate anticipatory attending and cast attending toward certain upcoming intervals. Here, these future points in time were all isochronous, so that any differences in the ability to detect deviant pitch intervals could be solely attributed to the preceding rhythmic context. The pattern of recognition accuracy data did in fact reveal systematic differences in performance, and, in conjunction with reaction time responses, indicated that listeners were differentially prepared for their *same/different* judgments.

The most important findings to emerge involve the set of structural invariants that support the generation of temporal expectancies. These invariants are defined by a joint relationship between temporal and melodic structure that is characteristic of coherent melodies. Although the consistent accentuation of phrase ending points did afford the opportunity to anticipate what kind of tonal interval might be deviated in a melody's final variation, this alone was not sufficient for the initiation of anticipatory attending.

Similarly, a simple, invariant periodicity of accented notes was also insufficient. Instead, anticipatory attending only arose when these two structural relations jointly co-occurred—namely, with an invariant periodicity of temporal accents that coincided with phrase ending points. This, then, suggests that listeners are better able to anticipate *when* in time an event will occur if they have a specific idea about *what* to expect. This phenomenon is consistent with the interactionist perspective of the Jones and Boltz (1989) model. Within this framework, the joint influence that pitch and time exert on behavior can be primarily explained through the structural description of a melody. At the most fundamental level of analysis, each acoustical dimension is dependent on the other for a melody's characterization; temporal structure cannot arise unless there is a serial sequence of notes that mark different points in time, and pitch relations themselves cannot emerge until they unfold in and over time. The level of greatest psychological relevance, however, is that which examines the extent to which temporal and melodic relations mutually reinforce one another. Although a tune may display a hierarchical framework of nested melodic relations, it is the concomitant pattern of temporal structure that will either highlight or obscure this organizational array. The attentional highlighting of phrase ending points is particularly important, because they anchor a melody to the underlying tonality scheme that both constrains the set of tonal intervals that can occur in a tune and determines their particular tonal function. In addition, phrase endings provide what one might call the "basic level" of melodic organization (Rosch, 1975)—the level that integrates the sequence of adjacent notes into nested

periodicities and lays the foundation for higher order nestings arising from interrelations among successive melodic phrases. It is therefore adaptive for a composer to temporally accentuate phrase endings to ensure that listeners' attention will be directed toward the points that maximally convey the underlying meaning and organization of musical structure. The psychological validity of phrase ending points has, in fact, been demonstrated in the previous literature. Sloboda and Gregory (1980), for example, found that when a series of clicks was superimposed on a melody, the clicks were consistently misperceived as occurring at phrase boundaries. A melody recall study by Boltz (1991a) also revealed that subjects used phrase endings as cognitive referents in that recall not only was selectively higher at these points but facilitated the retrieval of lower order relations among adjacent notes.

Although the attentional highlighting of phrase ending points is in itself beneficial to cognitive processing activities, an invariant periodicity between accents will further aid perceptual pickup because a listener can use these lawfully recurrent accents to attentionally track the tune with a minimal degree of effort. That is, if the pattern of temporal accentuation reliably guides one to stable reference points (i.e., phrase endings), one doesn't need to expend attentional effort to determine where the melody is going; instead, one can focus more on the sequence of information as it dynamically unfolds. In this sense, then, the process of perceptual pickup is anticipatory in nature. When listening to a coherent melody, one implicit aspect of attentional tracking is to cast attending forward toward the next upcoming accent point and to use this as a referent toward which current perceiving and attending activities are directed. This suggests that attending is goal-directed and aimed toward phrase ending points within the unfolding course of a melody. Some support for this notion comes from the Schmuckler (1990) study. When subjects were asked to continue a melody through keyboard responses, it was found that the ending of most extrapolations coincided with a tonic triad interval marking a phrase ending and, more interestingly, occurred at the points in time at which the actual phrase endings themselves would have appeared if played in real time. This anticipatory attending process was also reflected in the present study with the use of a melody recognition task. During the initial presentation of a tune (i.e., the standard), the extrapolation of temporal/melodic structure from the preceding context directed attending toward certain notes in a melody's final variation so that any subsequent deviations were recognized with a high degree of accuracy and speed. One could argue, as have Bharucha and Stoeckig (1986, 1987), that reaction time responses are one of the most informative measures of expectancy generation because they reflect the extent to which subjects are prepared and "primed" for an upcoming decision and motor response. Here, these responses were in fact fastest when the preceding context temporally accentuated phrase boundaries.

In sum, it can be suggested that the joint relationship between temporal and melodic structure that is found in

many forms of Western music is adaptive for cognitive processing activities and performs several important functions. An invariant pattern of temporal accent structure serves to (1) attentionally highlight the primary meaning (i.e., tonality scheme) and organization of a melody, (2) facilitate attentional tracking, and (3) support anticipatory attending and the generation of expectancies. In turn, expectancy generation enables a listener to initiate preparatory responses that may be required in a cognition experiment or in the context of a more ecological setting, the synchronization and smooth exchange of instruments during a concert performance (Rasch, 1979, 1988). At the same time, however, temporal accent structure can also disrupt cognitive performance. This is apparent with incoherent melodies in which temporal accents fail to coincide with phrase ending points and/or recur with a variant periodicity. When tonal intervals within a phrase are accented, as was true for the accentuation of non-tonic triad intervals in the present research, attending is misguided away from the underlying tonality scheme and the hierarchical arrangement of melodic relations. The overall organization of the melody is thereby obscured and there are no structural referents that can be used for attentional tracking and anticipatory attending. This results not only in a minimal pickup of information but also in a seeming array of disjointed items in memory (Boltz, 1991a). Similarly, a variant periodicity of temporal accents also disrupts cognitive performance, because a listener cannot anticipate when in time these will recur and they therefore no longer provide stable and reliable referents for attending activities. As the present set of results suggests, either source of structural incoherence is sufficient to produce a marked decrement in melody recognition performance, a decrement that is comparable to the case in which both incoherent relations appear in a melody.

Although the results of Experiment 1 indicate that a joint relationship between temporal and melodic structure is necessary for the formation and development of expectancies, this no longer applies *after* expectancies have been generated to guide the course of anticipatory attending. As seen in Experiment 2, notes in the final variation that confirmed temporal *or* melodic expectancies were both detected faster and more accurately than control loci, which confirmed neither type of expectancy. This, then, suggests that anticipatory attending can be split and independently directed toward the upcoming what and when of a melody. The results, however, also indicated a greater facilitating effect for the confirmation of temporal expectancies, suggesting that this dimension was the stronger of the two. The dominating influence of time over pitch has been observed in other empirical contexts, including the judged similarity of melodies (Monahan & Carterette, 1985) and the detection of hidden melodies (Dowling et al., 1987). From the present perspective, the temporal structure of a melody is the more important dimension for cognitive processing activities because it is the vehicle that drives the course of attending over a melody's

total time span. That is, temporal accent structure initially determines whether the underlying hierarchical arrangement of melodic relations will be attentionally highlighted or obscured for a listener, which in turn determines the success of perceptual pickup and attentional tracking. In addition, when anticipatory attending does occur, it is assumed to be guided by temporal accent structure, which takes a listener to particular points in time so that the melodic information that is there can be processed for its underlying tonality and organizational status. The results of Experiment 2 lend support to this notion with the superior performance engendered by temporally based expectancies.

These findings may also help to clarify the current divergence in the literature concerning the independent versus interactive influence of temporal and melodic structure. To a large extent, this may be resolved by considering the structure of tunes that are presented to subjects and the types of cognitive processing activities that are required for a given task. In particular, if subjects are required to learn and subsequently remember a melody, a joint relationship between temporal and melodic structure is critical, because this determines the overall coherence of the tune and the ability to extract the underlying organizational framework and tonality scheme for perceiving and retrieval purposes (Boltz, 1991a, 1992). Tasks that require or are facilitated by the formation of temporal expectancies will also reflect an interactive relationship like that found here in the context of a melody recognition task. Similarly, the joint impact of temporal and melodic structure has been observed in the perceptual ratings (Boltz, 1989b) and judged duration (Boltz, 1989c, 1991b; Jones & Boltz, 1989) of melodies in which the formation of temporal expectancies was initially supported through structural relations and subsequently violated in systematic ways. Independent relationships, on the other hand, have primarily emerged when subjects have been asked to judge the similarity among melodies (Monahan & Carterette, 1985) or perform perceptual ratings of tunes that display a less coherent structure (i.e., temporal accents fail to coincide with phrase endings) (Palmer & Krumhansl, 1987). In these situations, the formation of temporal expectancies is either not necessary for task performance or unable to arise, because of the structure of the events themselves.

In closing, the present experiments have revealed a set of structural invariants that support anticipatory attending during an unfolding melody. One issue worthy of future investigation concerns the overall generality of these results. There is in fact a need for converging operations to determine whether temporal expectancy generation is also observed in production and rating tasks. In addition, there are at least three other issues that involve the generality of structural relations. First, the types of expectancies explored here were relatively global ones that spanned over several measures and involved the level of phrase ending points. There are, however, lower order rhythmic relations that can also remain invariant over the course

of a tune. For example, notes within measures and phrases may display a consistent and recurrent pattern of short and long notes that thereby afford anticipatory attending. If empirically demonstrated, this then raises the possibility that there may be multiple levels of anticipatory attending that span over different hierarchical levels of structure and are nested within one another. Both Narmour (1989) and Jones (1990) have suggested that this may, in fact, be the case and have described the types of melodic and temporal invariants that can arise at different hierarchical levels of musical structure. A second issue concerns whether temporal expectancies are based on melodic relations other than phrase ending points. For example, certain chordal progressions may generate stronger temporal expectancies than others, and it may be that other such relations also afford anticipatory attending. Last, a question concerns the extent to which the generation of melodic expectancies is dependent on temporal relations. Given that both dimensions jointly contribute to the structural description of a tune, this phenomenon may also be a valid one that merits further investigation.

REFERENCES

- ABE, J., & HOSHINO, E. (1990). Schema driven properties in melody cognition: Experiments on final tone extrapolation by music experts. *Psychomusicology*, *9*, 161-172.
- BARTLETT, F. (1932). *Remembering: A study in experimental and social psychology*. Cambridge: Cambridge University Press.
- BELMORE, S., & HUBBARD, M. (1987). The role of advance expectancies in person memory. *Journal of Personality & Social Psychology*, *53*, 61-70.
- BHARUCHA, J. J., & STOECKIG, K. (1986). Reaction time and musical expectancy. *Journal of Experimental Psychology: Human Perception & Performance*, *12*, 403-410.
- BHARUCHA, J., & STOECKIG, K. (1987). Priming of chords: Spreading activation or overlapping frequency spectra? *Perception & Psychophysics*, *41*, 519-524.
- BOLTZ, M. (1989a). Perceiving the end: Effects of tonal relationships on melodic completion. *Journal of Experimental Psychology: Human Perception & Performance*, *15*, 749-761.
- BOLTZ, M. (1989b). Rhythm and "good endings": Effects of temporal structure on tonality judgments. *Perception & Psychophysics*, *46*, 9-17.
- BOLTZ, M. (1989c). Time judgments of musical endings: Effects of expectancies on the "filled interval effect." *Perception & Psychophysics*, *46*, 409-418.
- BOLTZ, M. (1991a). Some structural determinants of melody recall. *Memory & Cognition*, *19*, 239-251.
- BOLTZ, M. (1991b). Time estimation and attentional perspective. *Perception & Psychophysics*, *49*, 422-433.
- BOLTZ, M. (1992). The remembering of auditory event durations. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *18*, 938-956.
- BOLTZ, M., & JONES, M. R. (1986). Does rule recursion make melodies easier to reproduce? If not, what does? *Cognitive Psychology*, *18*, 389-431.
- CARLSON, J. (1981). Some factors which influence melodic expectancy. *Psychomusicology*, *1*, 12-29.
- CLARKE, E. (1985). Structure and expression in rhythmic performance. In P. Howell, I. Cross, & R. West (Eds.), *Musical structure and cognition* (pp. 209-236). London: Academic Press.
- DEUTSCH, D. (1980). The processing of structured and unstructured tonal sequences. *Perception & Psychophysics*, *28*, 381-389.
- DEWITT, L., & SAMUEL, A. (1990). The role of knowledge-based expectancies in music perception: Evidence from musical restoration. *Journal of Experimental Psychology: General*, *119*, 123-144.

- DOWLING, W. J. (1990). Expectancy and attention in melody perception. *Psychomusicology*, **9**, 148-160.
- DOWLING, W. J., LUNG, K. M.-T., & HERRBOLD, S. (1987). Aiming attention in pitch and time in the perception of interleaved melodies. *Perception & Psychophysics*, **41**, 642-656.
- DRAKE, C., DOWLING, W. J., & PALMER, C. (1991). Accent structures in the reproduction of simple tunes by children and adult pianists. *Music Perception*, **8**, 315-334.
- DUFFY, S. (1986). Role of expectations in sentence integration. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **12**, 208-219.
- EHRlich, S., & RAYNER, K. (1981). Contextual effects of word perception and eye movements during reading. *Journal of Verbal Learning & Verbal Behavior*, **20**, 641-655.
- HASTIE, R. (1980). Memory for information that confirms or contradicts a personality impression. In R. Hastie, T. Ostrom, E. Ebbesen, R. Wyer, D. Hamilton, & D. Carlston (Eds.), *Person memory: The cognitive basis of social perception* (pp. 155-177). Hillsdale, NJ: Erlbaum.
- HOCHHAUS, L. (1972). A table for the calculation of d' and B' . *Psychological Bulletin*, **77**, 375-376.
- JONES, M. R. (1981). A tutorial on some issues and methods in serial pattern research. *Perception & Psychophysics*, **30**, 492-504.
- JONES, M. R. (1990). Learning and the development of expectancies: An interactionist approach. *Psychomusicology*, **9**, 193-228.
- JONES, M. R., & BOLTZ, M. (1989). Dynamic attending and responses to time. *Psychological Review*, **96**, 459-491.
- JONES, M. R., BOLTZ, M., & KIDD, G. (1982). Controlled attending as a function of melodic and temporal context. *Perception & Psychophysics*, **32**, 211-218.
- JONES, M. R., BOLTZ, M., & KLEIN, J. (1992). *Judged duration and expected ending times*. Manuscript submitted for publication.
- JONES, M. R., KIDD, G., & WETZEL, R. (1981). Evidence for rhythmic attention. *Journal of Experimental Psychology: Human Perception & Performance*, **7**, 1059-1073.
- KRAMER, J. (1982). Beginnings and endings of Western art music. *Canadian University Music Review*, **3**, 1-14.
- KRUMHANSL, C., & KESSLER, E. (1982). Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychological Review*, **89**, 334-368.
- LERDAHL, F., & JACKENDOFF, R. (1983). *A generative theory of tonal music*. Cambridge, MA: MIT Press.
- LONGUET-HIGGINS, H. C. (1976). Perception of melodies. *Nature*, **263**, 646-653.
- LONGUET-HIGGINS, H. C. (1978). The perception of music. *Interdisciplinary Science Reviews*, **3**, 148-156.
- MEYER, L. B. (1956). *Emotion and meaning in music*. Chicago: University of Chicago Press.
- MEYER, L. B. (1967). *Music, the arts, and ideas*. Chicago: University of Chicago Press.
- MEYER, L. B. (1973). *Explaining music: Essays and exploration*. Chicago: University of Chicago Press.
- MILLS, C. (1980). Effects of the match between listener expectancies and coarticulatory cues on the perception of speech. *Journal of Experimental Psychology: Human Perception & Performance*, **6**, 528-535.
- MONAHAN, C. B., & CARTERETTE, E. C. (1985). Pitch and duration as determinants of musical space. *Music Perception*, **3**, 1-32.
- MONAHAN, C. B., KENDALL, R. A., & CARTERETTE, E. C. (1987). The effect of melodic and temporal contour on recognition memory for pitch change. *Perception & Psychophysics*, **41**, 576-600.
- NARMOUR, E. (1989). The "genetic code" of melody: Cognitive structures generated by the implication-realization model. *Contemporary Music Review*, **4**, 45-63.
- PACHELLA, R. G., & MILLER, J. O. (1976). Stimulus probability and same-different classification. *Perception & Psychophysics*, **19**, 29-34.
- PALMER, C., & KRUMHANSL, C. (1987). Independent temporal and pitch structure in determination of musical phrases. *Journal of Experimental Psychology: Human Perception & Performance*, **13**, 116-126.
- PALMER, C., & KRUMHANSL, C. (1990). Mental representations for musical meter. *Journal of Experimental Psychology: Human Perception & Performance*, **16**, 728-741.
- PISTON, W. (1978). *Harmony* (4th ed., revised and expanded by M. DeVoto). New York: Norton.
- POSNER, M., & SNYDER, C. (1975). Attention and cognitive control. In R. Solso (Ed.), *Information processing and cognition: The Loyola symposium* (pp. 55-85). Potomac, MD: Erlbaum.
- POVEL, D. (1985). Time, rhythms, and tension: In search of the determinants of rhythmicity. In J. Michon & J. Jackson (Eds.), *Time, mind, and behavior* (pp. 215-225). New York: Springer-Verlag.
- RASCH, R. (1979). Synchronization in performed ensemble music. *Acoustica*, **43**, 121-131.
- RASCH, R. (1988). Timing and synchronization in ensemble performance. In J. Sloboda (Ed.), *Generative processes in music: The psychology of performance, improvisation, and composition* (pp. 70-90). Oxford: Oxford University Press/Clarendon Press.
- ROSCHE, E. (1975). Cognitive reference points. *Cognitive Psychology*, **7**, 532-547.
- ROSNER, B., & MEYER, L. B. (1982). Melodic processes and the perception of music. In D. Deutsch (Ed.), *The psychology of music* (pp. 317-341). New York: Academic Press.
- ROSNER, B., & MEYER, L. B. (1986). The perceptual roles of melodic process, contour, and form. *Music Perception*, **4**, 1-40.
- SCHMUCKLER, M. (1989). Expectation in music: Investigation of melodic and harmonic processes. *Music Perception*, **7**, 109-150.
- SCHMUCKLER, M. (1990). The performance of global expectancies. *Psychomusicology*, **9**, 122-147.
- SLOBODA, J., & GREGORY, A. (1980). The psychological reality of musical segments. *Canadian Journal of Psychology*, **34**, 274-280.
- SRULL, T., & WYER, R. (1989). Person memory and judgment. *Psychological Review*, **96**, 58-83.
- TODD, N. (1985). A model of expressive timing in tonal music. *Music Perception*, **3**, 33-57.
- TODD, R. E., BOLTZ, M., & JONES, M. R. (1989). The MIDILAB research system. *Psychomusicology*, **8**, 83-96.
- UNYK, A., & CARLSON, J. (1987). The influence of expectancy on melodic perception. *Psychomusicology*, **7**, 3-23.
- YESTON, M. (1976). *The stratification of musical rhythm*. New Haven, CT: Yale University Press.

NOTE

1. The original versions of these folk tunes were selected from *Folk Songs of Many Peoples* (Vol. 1), edited by Florence Hudson Botsford (New York: Woman's Press, 1921).

APPENDIX A
The Set of Folk Tunes Used in Experiment 1

A Love Song by W. Bynner - 8 Beat Periodicity

China

Musical notation for 'A Love Song' in 2/4 time. The melody consists of 8 beats. Above the staff, 'N' is written above the first, third, and fifth notes. Below the staff, 'D1' is under the first note, 'D2' is under the fifth note, and 'NT TT' is under the eighth note.

Why, Oh Mother by E. St. Vincent Millay - 12 Beat Periodicity

Poland

Musical notation for 'Why, Oh Mother' in 3/4 time. The melody consists of 12 beats. Above the staff, 'N' is written above the first, fifth, and ninth notes. Below the staff, 'D1' is under the first note, 'D2' is under the fifth note, and 'NT TT' is under the twelfth note.

Spite by A. Mathewson - 10 Beat Periodicity

Italy

Musical notation for 'Spite' in 3/4 time. The melody consists of 10 beats. Above the staff, 'N' is written above the first, fifth, and ninth notes. Below the staff, 'D1' is under the first note, 'D2' is under the fifth note, and 'NT TT' is under the tenth note.

I'll Have No Other by J. Mokrejs - 8 Beat Periodicity

Czechia

Musical notation for 'I'll Have No Other' in 3/4 time. The melody consists of 8 beats. Above the staff, 'N' is written above the first, fifth, and seventh notes. Below the staff, 'D1' is under the first note, 'D2' is under the fifth note, and 'NT TT' is under the eighth note.

Look Out, How It's Raining by L. Speyer - 10 Beat Periodicity

Austria

Musical notation for 'Look Out, How It's Raining' in 2/4 time. The melody consists of 10 beats. Above the staff, 'N' is written above the first, fifth, and ninth notes. Below the staff, 'D1' is under the first note, 'D2' is under the fifth note, and 'NT TT' is under the tenth note.

Reapers On the Mountain by E. Leamy - 12 Beat Periodicity

Ukraine

Musical notation for 'Reapers On the Mountain' in 3/4 time. The melody consists of 12 beats. Above the staff, 'N' is written above the first, fifth, and ninth notes. Below the staff, 'D1' is under the first note, 'D2' is under the fifth note, and 'NT TT' is under the twelfth note.

Note—In the initial context of each melody, the temporal accentuation (i.e., ♩) of tonic triad members marking phrase ending points is depicted. N indicates nontonic triad members that were accentuated in a second condition of the experiment. Pitch deviation points are indicated by the D1, D2, NT, and TT loci, and the asterisk above each indicates the pitch change on the *different* trials of the experiment.

APPENDIX B
Melodies From Experiment 2 That Display a Dissociated Confirmation of Melodic and Temporal Expectancies

Near Krakow by W.A. Roberts - 8 Beat Periodicity

Poland

Musical notation for 'Near Krakow' in 3/4 time. The melody consists of 8 beats. Below the staff, 'D1' is under the first note, 'D2' is under the fifth note, and 'T M C' is under the eighth note.

I'll Have No Other by J. Mokrejs - 10 Beat Periodicity

Czechia

Musical notation for 'I'll Have No Other' in 3/4 time. The melody consists of 10 beats. Below the staff, 'D1' is under the first note, 'D2' is under the fifth note, and 'C M T' is under the tenth note.

Look Out, How It's Raining by L. Speyer - 12 Beat Periodicity

Austria

Musical notation for 'Look Out, How It's Raining' in 2/4 time. The melody consists of 12 beats. Below the staff, 'D1' is under the first note, 'D2' is under the fifth note, and 'C M T' is under the twelfth note.

Reapers On The Mountain by E. Leamy - 9 Beat Periodicity

Ukraine

Musical notation for 'Reapers On The Mountain' in 3/4 time. The melody consists of 9 beats. Below the staff, 'D1' is under the first note, 'D2' is under the fifth note, and 'T M C' is under the ninth note.