

Time estimation and attentional perspective

MARILYN BOLTZ

Haverford College, Haverford, Pennsylvania

The purpose of the present experiment was to investigate whether the filled interval effect is influenced by different attentional perspectives of the same event. Musically sophisticated listeners were asked to compare the duration of paired melodies that varied both in lower order contour changes and the relative timing of higher order phrase periodicities. Results indicated that a given pair of melodies could produce over- or underestimations, depending on which hierarchical level of structure was selectively tracked. These findings are discussed in terms of a framework that links event coherence and dynamic attending modes to time-estimation activities.

The study of time estimation has a long and diverse history within the psychological literature, ranging from studies addressing psychophysical problems (e.g., Weber's Law) to those assessing the relative impact of task demands, constitutional variables (e.g., age, mood, arousal, drugs), and temporal perspective (see Allan, 1979; Block, 1979; and Fraisse, 1984, for excellent reviews). However, one of the more puzzling phenomena that has confronted researchers in recent years is *the filled interval effect*. Within the context of a comparative judgment task, two equivalent intervals of time are not always judged to have the same duration. Instead, one might seem longer or shorter than another, depending on the nature of items filling these intervals. For example, the variables of stimulus familiarity (Avant & Lyman, 1975; Thomas & Weaver, 1975), task difficulty (Burnside, 1971; Hicks, Miller, Gaes, & Bierman, 1977), and the number and arrangement of stimulus items (Adams, 1977; Buffardi, 1971; Ornstein, 1969) have all been reported to influence comparative duration estimates. The intent here is to extend this previous literature by investigating whether selective attending toward different aspects of an event's structure yields different time estimates of the same event.¹

Some Models of Time Estimation

Over the past 30 years, several different theoretical perspectives have been developed to explain the filled interval effect. Many of these models adhere to the common assumption that experienced duration lengthens with an increased amount of information-processing activity. The variable of stimulus complexity is therefore assumed to be critically important since a more complex interval will require more cognitive processing than a less complex one and thus seem longer in relative duration. Although

all models share this underlying assumption, individual models differ in their particular definition of complexity and the presumed referent for duration judgments. Some researchers have suggested that time estimates depend on the number of chunks stored in memory (Frankenhauser, 1959; Ornstein, 1969), whereas others have proposed that it depends on the amount of selective attention expended during a given time interval (Thomas & Weaver, 1975; Underwood & Swain, 1973). Still a third approach claims that judged duration is mediated by the number of changes within a particular interval (Block, 1979, 1985, 1989)—changes that can arise from either environmental sources (e.g., task activities, stimulus structure, experimental context) or organismic sources (e.g., cognitive processing activities, emotional context).

Although these models have been useful for guiding a substantial amount of research, the research itself is often characterized by divergent and conflicting results (see, e.g., Avant & Lyman, 1975, vs. Devane, 1974; Ornstein, 1969, vs. Poynter, 1983). An interval defined as more complex is sometimes judged longer, but on other occasions shorter, than a less complex interval. In part, this may be due to vague and imprecise definitions of complexity in which it is not always clear which structural characteristics of an event serve as chunk boundaries, cognitive change points, or drains of attentional resources. A musical composition, for example, displays an interwoven arrangement of pitch change, melodic phrasing, timbre, and relative timing, yet it is not clear which of these structural features are important for duration judgments.

Jones and Boltz (1989) have recently proposed an alternative model in which they suggest that some problems may be resolved by acknowledging the full range of events encountered in the natural environment. Natural events differ in the coherence of their internal structure, which may encourage different modes of attending that ultimately influence time-estimation strategies. Some of these strategies involve preparatory attending and the generation of structurally based expectancies, while others involve constructs that have been offered by previous models—namely, memorial chunks, cognitive change, or attentional effort.

This research was supported by a Faculty Research Grant from Haverford College. The author owes a special thanks to Ron Joyner for the collection of data and technical assistance skills. Requests for reprints should be addressed to Marilyn Boltz, Department of Psychology, Haverford College, Haverford, PA 19041. Electronic mail address: M_BOLTZ@HVRFORD.

Dynamic Attending and Time Estimation

Many events within the natural environment can be considered to be highly coherent in that they transpire over time spans marked by nonarbitrary beginnings and ends. Within these time spans, information tends to display a high degree of predictability with multiple levels of structure interrelated through a hierarchical scheme. Such structural arrangements are commonly found within speech utterances and conversations (see, e.g., Grosjean, Grosjean, & Lane, 1979; Jaffe & Feldstein, 1970; Martin, 1972), locomotion patterns (see, e.g., Carlsoo, 1972; Pierson, 1976; Winter, 1983), and musical compositions (see, e.g., Boltz & Jones, 1986; Clarke, 1988; Deutsch & Feroe, 1981; Lerdahl & Jackendoff, 1983; Monahan & Carterette, 1985; Narmour, 1977), as well as narrative discourse (see, e.g., Mandler & Johnson, 1977; Rumelhart, 1977; Thorndyke, 1977), routine behavioral schemes (see, e.g., Bower, Black, & Turner, 1979; Schank, 1982), and action sequences (see, e.g., Newton, 1973, 1976; Newton, Hairfield, Bloomingdale, & Cutino, 1987). For example, consider the simple folk tune depicted in Figure 1. This tune nonarbitrarily begins and resolves on the tonic note (G) drawn from the melody's underlying tonality scheme of the G major diatonic scale. Within this total time span, sequences of notes are nested within melodic phrases marked by intervals bearing special harmonic relationships to the tonic. These tonal intervals constitute members of the tonic triad chord (tonic = G, dominant = D, mediant = B) and function as harmonic accents (denoted as h) by marking points of resolution within the melody (Berry, 1976; Cone, 1968; Kramer, 1982; Meyer, 1956; Piston, 1978; Todd, 1985). Finally, pairs of melodic phrases are nested within higher level phrases of a longer time span (here, 16 beats), and an overall hierarchical arrangement of melodic structure results.²

Another feature of coherent events is that the temporal structure of information within an event's total time span typically coincides with the arrangement of nontemporal

structure. Within music, temporal structure arises from the joint interplay of meter and rhythm. Meter refers to the referent beat of a melody and the number of beats within successive measures. The melody in Figure 1, for example, is based on a 4/4 meter in which a quarter note (♩) acts as the referent beat and there are 4 beats per measure. Rhythm, on the other hand, refers to the pattern of durational changes within a melody. These changes are often characterized in terms of a hierarchical tree structure in which notes longer (e.g., ♩) or shorter (e.g., ♪) than the referent beat form nested groupings within the underlying metrical framework (Longuet-Higgins, 1976, 1978). Rhythm is considered to be psychologically important in that pauses and prolonged tonal durations (e.g., the half note, ♪, in Figure 1) create temporal accents (t) that "stand out" and capture a listener's attention. In much Western music, temporal accents regularly recur in time and coincide with melodic phrase endings to yield a common hierarchy of temporal/melodic structure (Boltz & Jones, 1986; Clarke, 1982, 1988; Gabrielsson, 1988; Lerdahl & Jackendoff, 1983; Monahan & Carterette, 1985; Sundberg, 1988; Yeston, 1976). A similar structural arrangement is found in other natural events. Speech, for example, displays prolonged pauses that reinforce the grammatical phrase structure of an utterance (Butterworth, 1975; Goldman-Eisler, 1968; Grosjean et al., 1979) as well as stressed intonations that coincide with high-content words (Martin, 1972, 1975; Shields, McHugh, & Martin, 1974). Similarly, the boundaries of constituent action units are often marked by reductions in amplitude and overall rate of movement (Stetson, 1905; Todd, 1985). One consequence of joint accent structures is that the event's organization is attentionally highlighted so that it outlines the interrelationships among individual but nonadjacent items. Others have suggested that rhythm also affords a more dynamic function of guiding attending activities over time (Jones, 1976, 1981a, 1981b, 1987; Jones & Boltz, 1989; Meyer, 1956; Narmour, 1977). Since temporal accents coincide with nontemporal ones, and since

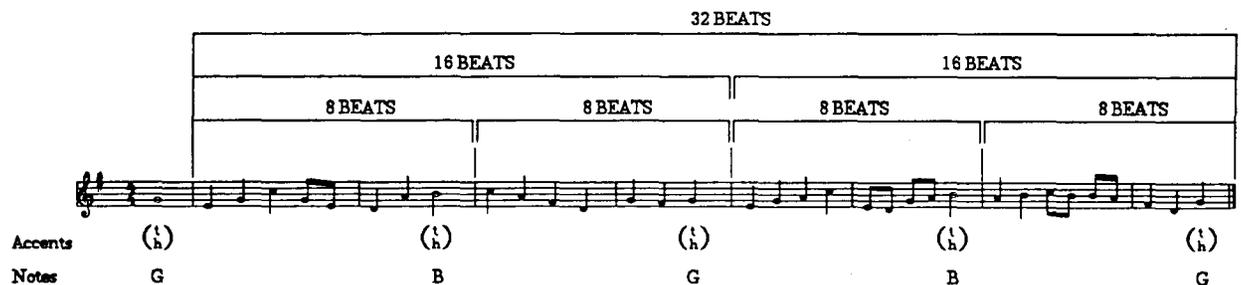


Figure 1. Example of a musical composition (excerpt from "There's Music in the Air," by G. Root) displaying a hierarchical arrangement of melodic/temporal structure (adapted from *The International Library of Music: Album of the World's Best Home Songs* [p. 124], by The Editorial Board of The University Society, New York: The University Society, Inc., 1964). This melody displays tonal resolution (Kramer, 1982; Meyer, 1956; Piston, 1978) in that it nonarbitrarily begins and ends on the tonic note, G. Within this total time span, sequences of notes are nested within melodic phrases marked by tonic triad members (here the tonic, G, and the mediant, B). These harmonic accent points (h) are attentionally highlighted by temporal accents (t); both regularly recur in time to outline higher order phrase relationships.

both recur with a predictable periodicity, a perceiver is therefore able to extrapolate the event's preceding context and generate expectancies about the "what" and "when" of future accents. This type of attending is primarily targeted toward higher order periodicities, and, given its anticipatory nature, it allows an organism to initiate preparatory responses.

In the model offered by Jones and Boltz (1989), it is suggested that the structural coherence displayed in music and other natural events allows a perceiver to regard a given event from different attentional perspectives. This in turn is assumed to influence time-estimation activities. When attending to an unfolding event, most perceivers will be captured by the pattern of temporal accents and the corresponding nontemporal information that is highlighted. Attending will therefore be directed toward higher order phrase periodicities and will generate expectancies about phrase ending points. In such situations, time estimation is assumed to be determined by the confirmation or violation of expectancies. With respect to the filled interval effect, if two events have equivalent total durations and both end "when" expected, people will correctly judge them to be the "same." However, if one event violates an expectancy by seeming to end later than anticipated, it will be incorrectly judged longer. Similarly, events appearing to end "too early" in time will be judged relatively short. In sum, when listeners are relying on an anticipatory mode of attending, duration estimates will depend on a temporal contrast that arises from a temporal disparity between an event's actual and expected ending.

Although coherent events often encourage anticipatory attending over higher order structural relationships, they can also afford a very different attentional perspective. Depending on an individual's familiarity with an event or the goals of perceiving, attending may be shifted downward toward the level of adjacent items within an event's time span. For example, during the course of a conversation, a listener may become enamored of the particular dialect of a speaker and focus attending on the enunciation of successive phonemes and words. Or when listening to a musical composition, a music critic or piano teacher might focus on the precision with which individual notes are played. Information within these lower levels of hierarchical structure often displays a high degree of predictability that supports the generation of expectancies. Narmour (1989), for example, has outlined numerous ways in which lower order expectancies are generated from musical structure. In other cases, however, structural predictability is less apparent, and perceivers are unable to reliably anticipate the event's future course. Hence, people will be more attuned to changes among adjacent notes (e.g., changes in pitch direction, tonal duration, or timbre) and local groupings of notes based on a common structural attribute. When inferring duration estimates, people may then simply count and compare the total number of changes or tonal groups within melodies, thus relying on the strategies proposed within Block's (1979, 1985,

1989) contextual change and Ornstein's (1969) storage-size hypotheses. Time estimates will be more variable here because different subjects may attend to different lower order relations, but within subjects, experienced duration is predicted to increase with more changes or chunks.

A similar process is also assumed to mediate time judgments of less coherent events. Many natural events do display less internal predictability and nested levels of inter-related structure are often lacking (see Jones & Boltz, 1989, for a more detailed analysis). This is most evident in the productions of unskilled performers. Consider, for example, the "music" of a novice pianist in which pauses and prolonged tonal durations (i.e., temporal accents) often fail to coincide with important melodic information (i.e., h accents). The toddling walk of a young child and the speech of a foreigner learning a new language demonstrate a similar lack of coherence, as does the overall chatter of a lively cocktail party or the bustle of a crowded street. These sorts of events necessarily enforce attending toward lower order information because higher level periodicities arising from the joint relationship of nontemporal and temporal information are missing. When asked to make time judgments of such events, people again will rely on the number of changes or local groupings to infer experienced duration.

To summarize, the model offered by Jones and Boltz (1989) provides a more inclusive framework for the investigation of time-estimation activities. By acknowledging the full range of events within the natural environment, an analysis of event structure can predict what attentional perspective may be adopted by an individual to determine time-estimation strategies. Highly coherent events afford anticipatory attending over higher level periodicities within the structural hierarchy. Given their lawful internal organization, however, such events also allow an individual to flexibly shift attending from higher level periodicities to lower level information encompassing "details" of the event's structure. Structural relations at this level may or may not support the generation of expectancies. Conversely, the lack of higher order structural predictability within less coherent events confines attending to lower level periodicities and does not allow for shifts in attentional activities. Time judgments resulting from anticipatory modes of attending rely on the confirmation/violation of expectancies; those arising from an attentional focus on incoherent relations reflect counting strategies.

Empirical support for this view comes from a set of studies by both Boltz (1985) and Jones and Boltz (1989). In these experiments, subjects were asked to compare the relative duration of paired folk tunes that varied in the placement of temporal accents. These melodies were designed to promote anticipatory attending in that temporal accents always coincided with melodic phrases (h accents), and, at least initially, these coupled accents recurred after an invariant number of beats. However, in some versions of the melodies, the metrical timing of the final accent

phrases was varied such that the last note occurred "too early" or "too late" in time relative to preceding phrase periodicities. The results indicated that relative durations were under- and overestimated, respectively, even though all melodies were equivalent in total duration and number of beats. Boltz (1989) has also examined the effects of expectancies involving the very last note of a musical composition. When subjects are asked to compare the durations of tunes varying in tonal resolution, melodies ending on the leading tone interval are judged shorter than those ending on the conventional tonic. Since the leading tone interval leaves a listener "hanging" before the expected return to the tonic, such endings seem too early in time and thus yield underestimations.

While this research supports the idea that anticipatory attending can lead to temporal contrast and systematic biases in judged duration, time-estimation strategies based on incoherent structural relations have been addressed less directly. The evidence that Jones and Boltz (1989) used to support this claim was drawn from previous research relying on relatively incoherent events, such as lists of nonsense or unrelated words (Block, 1974; Poynter, 1983; Thomas & Weaver, 1975), auditory clicks (Adams, 1977), abstract drawings (Ornstein, 1969), or sequences of lights (Poynter & Homa, 1983). These studies, however, were not explicitly designed to assess the effects of differential attending modes, and more direct tests are therefore needed.

Manipulations of Attentional Perspective

The present experiment was designed to investigate whether different attentional perspectives of the same event may result in different estimates of relative duration. Subjects were asked to compare the duration of paired melodies that displayed structural predictability at higher level phrase periodicities and thus supported the generation of listeners' expectancies. At lower level periodicities, however, local changes in pitch direction occurred irregularly in time to yield a relative lack of predictability at this level. By forcing listeners to attentionally track either higher or lower order information, the relative impact of differential attending modes could be assessed.

The simultaneous manipulation of both higher and lower order information demands a rigorous experimental control of melodic structure, and to this end, all melodies within this study were generated from the *next* rule within the dihedral symmetry group D_7 (see Boltz & Jones, 1986, for a more detailed discussion).³ As can be seen in Figure 2, all melodies were based on the C major diatonic scale and contained 36 notes arranged into 12 sets of three tone groups. Notice that the tones within a group are separated by either adjacent scale steps or semitones (N^1 to N^2), and that notes between groups are separated by a significant pitch skip (N^2 to N^3). To establish higher order phrase markings that were perceptually salient and could be quickly detected in an ongoing monitoring task, tonal mistakes were introduced at the end of certain tonal groupings.⁴ These mistakes consisted of notes outside the

C major scale (e.g., A#, D#, F#) and were selected so that no key other than C major was implied by the melody's sequence of pitch intervals.

Temporal contrast was manipulated by varying the relative placement of mistakes within a melody. Previous manipulations of this variable (Boltz, 1985, 1989; Jones & Boltz, 1989) were achieved by imposing a given meter on the first half of a melody and then changing the meter in the second half so that a fewer or greater number of beats occurred between final phrase periodicities. This, however, resulted in a confound of tempo in which a fewer number of beats was always correlated with a relatively faster ending, and a greater number of beats was correlated with a relatively slower ending. Hence, to avoid this type of confound in the present experiment, overall meter, rhythm, and tempo were held constant both within and between all melodies. Temporal contrast was then manipulated by varying the number of tonal groups between tonal mistakes. Specifically, the standard melody of all pairs always contained tonal mistakes that recurred after an invariant periodicity. For example, notice that the melody depicted in Figure 2A displays tonal mistakes after every third tonal grouping. In this type of melody, the listener is allowed to generate expectancies about when future mistakes should occur, and these expectancies are always confirmed. Comparison melodies, however, were designed to sometimes violate listener expectancies. For example, consider the melody labeled 3324 in Figure 2B. This tune contains tonal mistakes that initially recur with an invariant periodicity (i.e., after every third tonal group), but the placement of mistakes in the latter half of the melody violates this predictability by occurring after the second and fourth grouping, respectively. Given the melody's preceding context, a listener should expect the last mistake to occur after either the third or the second tonal grouping. But since the final mistake actually occurs at the end of the fourth grouping, the melody ends later than expected and results in positive temporal contrast.⁵ It should therefore seem "long" in relative duration. The melody shown in Figure 2C should yield an even greater magnitude of positive temporal contrast, since the final mistake here occurs significantly later than would be expected on the basis of the preceding phrase periodicities (i.e., after five tonal groupings).

Other comparison melodies were designed to create conditions of negative temporal contrast. The melody labeled 3342 in Figure 2D, for example, should lead a listener to expect the upcoming final mistake to occur after three or four additional tonal groups. However, since it actually occurs after only two additional groups, the melody ends earlier than would be expected, and it should seem "short" in relative duration. Again, the tune shown in Figure 2E should show a greater magnitude of this effect, since the end is significantly earlier than the preceding periodicities would suggest. Last, a final set of comparison melodies displayed "null contrast" in that tonal mistakes always recurred after an invariant periodicity. Since expectancies about the relative placement of mis-

A. 3333 Phrase Periodicities (Five Contour Changes)

C4 D4 E4 \uparrow A4 B4 C5 \uparrow F4 E4 D#4 \uparrow G4 A4 B4 \uparrow E5 F5 G5 \uparrow C5 B4 A#4 \uparrow G4 F4 E4 \uparrow B3 A3 G3 \uparrow E4 F4 F#4 \uparrow C5 D5 E5 \uparrow B4 A4 G4 \uparrow E4 D4 C#4

N⁺⁵ N⁻⁷ N⁺⁴ N⁺⁵ N⁻⁷ N⁻³ N⁻⁵ N⁺⁹ N⁺⁴ N⁻⁵ N⁻³

B. 3324 Phrase Periodicities (Seven Contour Changes)

C4 D4 E4 \uparrow A4 B4 C5 \uparrow F4 E4 D#4 \uparrow B4 C5 D5 \uparrow G4 F4 E4 \uparrow C4 B3 A#3 \uparrow E4 F4 G4 \uparrow D4 C4 A#3 \uparrow F4 G4 A4 \uparrow C5 D5 E5 \uparrow B4 A4 G4 \uparrow E4 D4 C#4

N⁺⁵ N⁻⁷ N⁺⁸ N⁻⁷ N⁻⁴ N⁺⁶ N⁻⁵ N⁺⁷ N⁺³ N⁻⁵ N⁻³

C. 3315 Phrase Periodicities (Three Contour Changes)

C4 D4 E4 \uparrow A4 B4 C5 \uparrow E5 F5 F#5 \uparrow A5 B5 C6 \uparrow G5 F5 E5 \uparrow C5 B4 A#4 \uparrow F4 E4 D#4 \uparrow B4 C5 D5 \uparrow C5 A5 B5 \uparrow F5 E5 D5 \uparrow B4 A4 G4 \uparrow E4 D4 C#4

N⁺⁵ N⁺⁴ N⁺³ N⁻⁵ N⁻⁴ N⁻⁵ N⁺² N⁺⁵ N⁻⁶ N⁻³ N⁻³

D. 3342 Phrase Periodicities (Nine Contour Changes)

C4 D4 E4 \uparrow B3 A3 G3 \uparrow E4 F4 F#4 \uparrow A4 B4 C5 \uparrow G4 F4 E4 \uparrow B4 C5 C#5 \uparrow B4 A4 G4 \uparrow E4 D4 C4 \uparrow F4 G4 A4 \uparrow D4 C4 A#3 \uparrow E4 F4 G4 \uparrow E4 D4 C#4

N⁻⁵ N⁺⁹ N⁺³ N⁻⁵ N⁺⁷ N⁻² N⁻³ N⁺³ N⁻⁷ N⁺⁶ N⁻³

E. 3351 Phrase Periodicities (Seven Contour Changes)

C4 D4 E4 \uparrow A4 B4 C5 \uparrow F4 E4 D#4 \uparrow B4 C5 D5 \uparrow G4 F4 E4 \uparrow C4 B3 A#3 \uparrow F4 G4 A4 \uparrow E4 D4 C4 \uparrow E4 F4 G4 \uparrow C5 D5 E5 \uparrow B4 A4 G#4 \uparrow E4 D4 C#4

N⁺⁵ N⁻⁷ N⁺⁸ N⁻⁷ N⁻⁴ N⁺⁷ N⁻⁵ N⁺⁴ N⁺⁵ N⁻⁵ N⁻⁴

Figure 2. An exemplar set of the melodies that were used as experimental stimuli. These melodies, constructed from the set of *next* rules, differ in both the number of lower order contour changes and the timing of higher order phrases. The melody depicted in panel A displays tonal mistakes (coupled with temporal accents, \uparrow) that regularly recur after every third tonal grouping. The temporal predictability of these atonal notes allows listeners to generate expectancies about *when* future mistakes will occur, which here are reliably confirmed. The melodies within panels B and C illustrate the concept of positive temporal contrast, in that the timing of mistakes in the third and fourth periods has been manipulated so that the last mistake occurs later than expected. Similarly, the timing of atonal notes within panels D and E creates negative temporal contrast because the last mistake occurs earlier than is suggested by the preceding phrase periodicities. Melodies within panels C and E show a greater magnitude of temporal contrast than those within panels B and D.

takes are always confirmed, these pairs of tunes should be correctly judged as equivalent in total duration.

These predictions of temporal contrast were assessed in the present study by asking a group of subjects to attentively track, via buttonpresses, the occurrence of all tonal mistakes in both the standard and the comparison melodies of a pair. This mode of anticipatory attending should yield systematic over- and underestimations as a function of positive and negative temporal contrast, respectively, and thus converge with other manipulations of structurally based expectancies (Boltz, 1985, 1989; Jones & Boltz, 1989).

A second manipulation of higher order phrase structure was designed to contrast certain theoretical predictions of time-estimation activities. In the models that rely on constructs of cognitive change (Block, 1979, 1985, 1989) and memorial chunks (Ornstein, 1969), it is assumed that a greater number of changes (or chunks) will result in longer estimates of relative duration, regardless

of where these changes occur or how they are arranged in time. In the model proposed by Jones and Boltz (1989), however, it is assumed that the number of change points/chunks is important only in cases of low structural coherence in which such changes occur unpredictably in time. When attending is directed toward predictably recurrent periodicities, the confirmation/violation of expectancies is the more critical determinant of duration judgments. These two positions were contrasted here through presenting subjects with standard and comparison melodies that could vary in their respective numbers of higher order phrases. This manipulation is shown in the upper portion of Table 1 which depicts all pairings of higher order phrase periodicities used in the present experiment. Notice that for each condition of temporal contrast, standard and comparison melodies within the first two pairings both contain four tonal mistakes (which can act as either change points or chunk boundaries). In the latter two pairings, however, standards and comparisons differ in their relative

Table 1
The Pairing of Standard and Comparison Melodies That Give Rise to Conditions of Higher Order Temporal Contrast and Differences in the Number of Lower Order Contour Changes

Pairings of Higher Order Phrase Periodicities		
Positive Contrast	Negative Contrast	Null Contrast
3333-3324	3333-3342	3333-3333
3333-3315	3333-3351	3333-3333
22222-3324	444-3342	444-3333
22222-3315	444-3351	22222-3333

Pairings Based on the Number of Lower Order Contour Changes		
Positive Difference	Negative Difference	Null Difference
5-9	9-5	3-3
3-7	7-3	5-5
3-5	5-3	7-7
7-9	9-7	9-9

Note—Numbers within the temporal contrast conditions indicate the number of tonal groups within phrases marked by tonal mistakes.

number of tonal mistakes.⁶ In the pairing of 22222-444, for example, the standard contains atonal notes after every second tonal grouping to yield a total of six mistakes, whereas the comparison melody contains only three mistakes, each one occurring after every fourth tonal group. This manipulation allowed subjects to rely on one of two strategies. If subjects use the arrangement of tonal mistakes to generate expectancies about event endings, as predicted by Jones and Boltz (1989), then the relative number of tonal mistakes should have no impact within a given condition of temporal contrast. On the other hand, if subjects do infer experienced duration from the relative number of changes/chunks within paired melodies, then systematic differences as a function of this variable should emerge. This finding would support models in which it is claimed that judged duration depends on the amount of processing activity.

A final manipulation was designed to assess time-estimation strategies when subjects were attending to lower level information within paired melodies. As can be seen in Figure 2, all melodies displayed contour changes (shifts from a rise in pitch to a fall in pitch, or vice versa) at the end of certain tonal groupings, and these changes occurred irregularly in time to yield an overall contour that was asymmetrical in shape. Melodies varied in relative number of contour changes (i.e., three, five, seven, or nine changes), and for each of the 12 pairs differing in temporal contrast, six different versions were presented. In two of these, the standard and comparison shared the same number of contour changes. However, in the remaining four versions, the comparison contained relatively more changes than the standard did (two or four more changes), or relatively fewer changes (two or four less changes). In the bottom half of Table 1, all pairings of contour changes used in the present experiment are presented. The prediction here was that subjects who are attentionally tracking this level of information will engage in

counting strategies and thereby base their judgments on the relative number of contour changes. This prediction is consistent with both Ornstein's (1969) storage-size and Block's (1979, 1985, 1989) contextual change hypotheses.

In sum, this experiment was designed for the purpose of examining whether a given event would be over- or underestimated in relative duration, depending on the level of structure to which a perceiver was attentionally attuned. A comparison melody may seem relatively longer than its standard if one is attending to higher level periodicities, but on other occasions, it will seem relatively shorter if one is attuned to lower order information involving contour change points. To evaluate these various predictions, the following experiment was conducted.

METHOD

Design and Subjects

The design was a $3 \times 2 \times 2 \times 3 \times 2 \times 2$ mixed factorial. Subjects received a series of paired melodies in which the standard and comparison could differ in temporal contrast of structurally based expectancies (positive, negative, null); magnitude of temporal contrast (high, low); number of melodic phrases (*same*, *different*); and magnitude of contrast in the relative number of contour changes (high, low, zero). Attending mode (trackers of higher order phrase periodicities or lower order contour changes) and counterbalance order (I, II) were the two between-subjects factors.

Thirty-two subjects from an introductory psychology course at Haverford College participated in the experiment in return for course credit. Each had had at least 4 years of musical experience within the past 6 years and was currently playing a musical instrument. Eight subjects were randomly assigned to one of the four between-subjects conditions.

Stimulus Materials

A total of 28 musical patterns were constructed, each containing 36 notes primarily drawn from the C major diatonic scale. These patterns, which were generated from the *next* rule within the dihedral symmetry group D_7 , consisted of 12 sets of 3-tone groups. As can be seen in Figure 2, notes within a tonal group were related by the N^1 or N^2 rule (in semitone units; see note 3), and notes between groups were separated by a pitch skip that could range from two to nine semitones (i.e., N^2 to N^9). To establish the C major scale as the underlying tonality scheme, each melody opened with the major triad chord (i.e., $C_4 E_4 G_4$), followed by the three-tone argument of $C_4 D_4 E_4$ (the subscripts indicate the octave). Across all melodies, there was an attempt to equate octave range as well as the relative size distribution of all pitch skips. As described below, these 28 melodies varied in melodic phrase structure and contour arrangement.

Melodic Phrase Structure

A higher order phrase structure was created for each melody by inserting *tonal mistakes* at the end of certain tonal groupings. These tonal mistakes, which consisted of chromatic notes outside the C major diatonic scale (e.g., $F\#$, $D\#$, $A\#$), were carefully selected so that no key other than C major was suggested. In three pattern sequences, the tonal mistakes were predictably arranged and they regularly recurred at the end of every second, third, or fourth tonal group (i.e., 22222; 3333; 444). In the remaining sequences, the tonal mistakes were placed at more irregular intervals to yield conditions of temporal contrast. Here, the tonal mistakes initially recurred with an invariant periodicity (i.e., after every third tonal group),

but the placement of mistakes within the latter half of the melody was arranged so that the third mistake occurred after the fourth (or fifth) tonal group, placing the fourth mistake at the end of the final tonal group ending a tune (i.e., 3342 and 3351). Since the last mistake seems to occur "too early" relative to the preceding context, these conditions should result in negative contrast. Similarly, conditions of positive contrast were created by placing the last mistake at an interval that seemed "too late" in time relative to preceding phrases (i.e., 3324 and 3315).

Together, these variations of phrase structure formed a set of 12 pattern pairs, and tunes within a pair were always equivalent in total duration and total number of beats. As can be seen in Table 1, the standard melody of each pair always confirmed expectancies about "when" each tonal mistake would occur, including the final one that ended each melody. The comparison melody for 4 of the pairs displayed a similar structural arrangement (i.e., null contrast) or an array that was predicted to result in either positive or negative temporal contrast—this contrast being of a greater magnitude (i.e., 3315 and 3351) or a lesser one (i.e., 3324 and 3342). Finally, half of the pattern pairs within each condition displayed standards and comparisons that either shared the same number or different numbers of melodic phrases.

Contour Arrangement

Standard and comparison melodies within a given pair could differ not only in melodic phrase structure, but also in the relative number of contour changes. For each of the seven melodic phrase sequences described above, four different versions were constructed that varied in number of contour changes (three, five, seven, or nine changes), thus yielding a total of 28 experimental melodies. These changes in pitch direction always occurred at the end of certain tonal groups, and, as can be seen in Figure 2, they appeared irregularly in time to yield an overall contour that was asymmetrical in shape.

During the actual presentation of experimental trials, six different contour pairings occurred for each of the 12 pattern pairs varying in melodic phrase structure. In two of these six pairings, the comparison contained the same number of contour changes as the standard did, more changes than the standard did (two or four more changes), or fewer changes than the standard did (two or four fewer changes). Each of the 12 contour pairings shown in Table 1 occurred six times throughout an entire experimental session, and these were evenly distributed across the temporal contrast conditions. A total of 72 experimental trials thus resulted; these were randomized into two blocks of 36 trials.

Finally, all melodies displayed the same metric and rhythmic structure, namely a repetitive pattern of ♩ ♩ ♩, with the same stimulus onset asynchrony for all quarter notes ($\text{♩} = 450$ msec), eighth notes ($\text{♪} = 225$ msec), and dotted quarter notes ($\text{♪} = 675$ msec). The dotted quarter note served as an agogic accent coinciding with all tonal mistakes and contour change points, thereby facilitating the tasks of attentional tracking.

Apparatus

All melodies were constructed and generated with the MIDILAB software system (Todd, Boltz, & Jones, 1989). During each experimental session, melodies were presented on line by 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-A4010 receiver and played over Koss-Pro 4AAA Plus headphones at a comfortable listening level (60–65 dB). The timbre for all melodies approximated that of a flute.

Procedure

Recorded instructions informed the subjects of pattern presentation details and task requirements. On each trial, a 1-sec warning tone (C, pitch) preceded the standard melody of a pair by 2 sec.

Three seconds after the offset of this standard sequence, a comparison melody was played. Those subjects tracking higher order phrases were told that for each melody within a pair, they should listen for any tonal mistakes—that is, "notes that sound wrong and inappropriate to the melody's tonality scheme." They were then told that they should try to synchronize with these mistakes by pressing a button on the response console in front of them as quickly as possible. Those subjects in the contour-tracking condition were asked to press a button corresponding to any changes in pitch direction within both melodies of a pair—that is, "when a melody shifts from a rise in pitch to a fall in pitch or vice versa." They were also told that all melodies would contain some tonal mistakes, but that they should try to ignore them and simply listen for any changes in pitch direction. After tracking their respective levels of structural information, all subjects were asked to compare the total duration of the comparison melody relative to its standard. To indicate this judgment, they were provided with a response sheet that contained two horizontal lines for each trial. Subjects were informed that the length of the top line represented the total duration of the standard melody. A second line directly underneath but indented halfway relative to the first line contained values ranging from -50% to 0 to $+50\%$. Subjects were asked to mark this second line to reflect the relative total duration of the comparison melody as a proportion of the standard. It was suggested that if the comparison melody seemed much longer (or shorter) than the first, they might mark values of $+40\%$ to $+50\%$ (or -40% to -50%), and that if the comparison seemed only slightly longer (shorter) than the standard, values of $+10\%$ to $+20\%$ (or -10% to -20%) might be appropriate. A value of zero in the middle of the scale indicated that the two melodies seemed equivalent in total duration. Subjects were allowed 5 sec to make this duration judgment, after which the warning tone recurred for the next trial.

Subjects were tested in small groups of 2–4 individuals and were asked to remove their watches during the experimental session. After receiving six practice trials, subjects were then presented with two blocks of 36 experimental trials, with a brief rest in between. Each entire experimental session was approximately 45 min in duration.

RESULTS

Separate analyses of variance were conducted for each mode of selective attending. Each is discussed in turn.

Trackers of Higher Order Phrase Periodicities

In this condition, subjects were asked to track tonal mistakes marking higher order phrases within paired melodies before indicating a relative time judgment. Since counterbalance order and differences in the relative number of contour changes both contributed negligible effects ($F < 1.00$), mean responses were collapsed over these two variables. The mean duration judgments in Table 2 are therefore presented as a function of temporal contrast conditions.

The overall analysis of variance (ANOVA) indicated that time estimates did vary with the relative timing of tonal mistakes [$F(2,28) = 704.56$, $MS_e = 100.40$, $p < .0001$]. Within the null contrast condition, in which mistakes recurred with an invariant periodicity, melodies were not judged significantly different from a value of zero (representing equivalent relative duration). However, relative overestimations arose when the final mistake within a tune occurred later than was expected, thus yielding positive temporal contrast. A set of Bonferroni post

Table 2
Mean Duration Judgments for Subjects Tracking
Higher Order Phrase Periodicities Marked by Tonal Mistakes

Phrase Periodicity Pairings	<i>M</i>	<i>SD</i>
Positive Contrast		
3333-3324	13.96	5.36
3333-3315	21.66	6.64
222222-3324	14.12	5.86
222222-3315	22.81	5.98
Negative Contrast		
3333-3342	-13.49	7.56
3333-3351	-4.85	16.91
444-3342	-10.94	8.47
444-3351	-6.01	15.26
Null Contrast		
3333-3333	.73	2.95
3333-3333	.87	3.08
222222-3333	1.09	3.49
444-3333	.94	3.81

Note—Values of zero indicate equivalent judged duration; values < 0 indicate the comparison was underestimated relative to the standard; and values > 0 indicate the comparison was overestimated.

hoc comparisons (Dunn, 1961) confirmed that all judgments within this condition were significantly different from zero ($p < .01$). Conversely, mistakes occurring earlier than was expected created negative temporal contrast and relative underestimations of judged duration (all Bonferroni post hocs, $p < .01$).

Some evidence also suggested that time judgments were influenced by the magnitude of temporal contrast, although this interacted with the particular direction of temporal disparity [$F(2,28) = 17.49, MS_e = 142.46, p < .0001$]. Within the positive contrast conditions of Table 2, notice that final phrase periodicities showing a greater deviation from their preceding context (i.e., 3315) produced longer time estimates than did those with less deviation (i.e., 3324) ($p < .01$). The negative contrast condition, however, failed to reveal this magnitude effect and in fact yielded an opposite pattern of results. Here, final phrase periodicities showing a greater deviance from the preceding context (i.e., 3351) were less underestimated and more variable ($p < .01$).

Finally, the overall ANOVA showed no effect for the relative number of higher order phrases within standard and comparison melodies. Standards containing three (i.e., 444) or six tonal mistakes (i.e., 222222) yielded the same pattern of temporal contrast effects as pairings in which the standard and comparison both shared the same number of atonal notes.

Overall, the present set of results supports the idea that attending directed toward coherent structural relations produces time estimates dependent on the confirmation/violation of event-specific expectancies.

Trackers of Lower Order Contour Changes

Mean duration estimates are presented in Table 3 as a function of the number of relative contour changes within

standard and comparison melodies. Means are collapsed over higher order phrase periodicities and counterbalance order since both factors were nonsignificant ($F < 1.00$).

The results here indicate that time judgments seem to reflect a counting strategy based on the relative number of contour changes [$F(2,28) = 1238.92, MS_e = 62.38, p < .0001$]. Comparison melodies containing more changes than the standard yielded overestimations of judged duration, whereas those containing fewer changes than the standard yielded relative underestimations. A set of Bonferroni post hoc comparisons confirmed that all pairings within these conditions differed significantly from zero, the value of equivalent judged duration ($p < .01$). Melodies sharing the same number of contour changes were in fact judged to have the same duration.

The overall ANOVA also revealed that the magnitude of over- and underestimations depended on the magnitude of difference in contour change [$F(2,28) = 54.88, MS_e = 59.15, p < .0001$]. Within Table 3, notice that standard and comparison melodies displaying a difference of four contour changes yielded higher values than those based on a difference of two changes. This was true for both conditions of directional change (all Bonferroni post hocs, $p < .01$).

DISCUSSION

The results of this experiment support the idea that different attentional perspectives of the same event can yield different relative duration judgments. Many natural events, including music, contain multiple levels of structure nested within a hierarchical arrangement. Some structural levels will display predictably recurrent phrase periodicities that support anticipatory attending and the

Table 3
Mean Duration Judgments for Subjects Tracking the Number
of Lower Order Contour Changes

Pairings Based on the Number of Contour Changes	<i>M</i>	<i>SD</i>
Positive Difference		
5-9	21.30	9.77
3-7	19.79	8.56
3-5	13.70	7.64
7-9	12.71	6.90
Negative Difference		
9-5	-17.66	10.64
7-3	-19.17	7.67
5-3	-10.21	9.96
9-7	-8.65	9.91
Null Difference		
3-3	.52	4.18
5-5	.68	3.05
7-7	1.15	3.75
9-9	1.25	3.49

Note—Values of zero indicate equivalent judged duration; values < 0 indicate the comparison was underestimated relative to the standard; and values > 0 indicate the comparison was overestimated.

generation of event-specific expectancies. Within the same event, however, other structural levels may display less predictability over the unfolding time course and thereby direct attending toward adjacent items. The results obtained here suggest that these two different modes of selective attending engage different time-estimation strategies such that the relative duration of the same event may be over- or underestimated.

Within the present experiment, the basis for anticipatory attending was established through the temporal patterning of tonal mistakes within paired melodies. Standard melodies contained mistakes that always recurred with an invariant periodicity, and expectancies concerning the timing of future mistakes could then be carried over to a comparison melody. The eventual confirmation or violation of these expectancies, in turn, influenced judged duration. As illustrated in the null contrast conditions, duration judgments were quite accurate when the final mistake ending a tune coincided with its expected periodicity. However, biases emerged with temporal contrasts between an event's actual and expected ending; endings occurring later than had been expected led to overestimations of relative duration, whereas endings occurring earlier than had been expected produced underestimations.

These findings generalized across melodic pairs containing different patterns of contour change as well as a different number of higher order phrase periodicities. The opportunity to base judgments on the relative number of tonal mistakes did, in fact, exist within the present experimental design. This effectively allowed subjects to use the strategies prescribed by the contextual change (Block, 1979, 1985, 1989) and storage-size hypotheses (Ornstein, 1969). But this was not the case. Subjects appeared to rely on the *relative timing* of mistakes and base their judgments on the temporal disparity between expected and actual ending times. Such findings converge with previous research reported by Boltz (1985, 1989) and Jones and Boltz (1989).

Although no evidence for a counting strategy emerged with attending toward higher order phrase periodicities, it did appear when attending was directed toward lower order contour changes. At this level, time estimates directly reflected the relative number of contour changes within standard and comparison melodies. These changes in pitch direction occurred irregularly in time and offered no structural basis from which to reliably generate expectancies about upcoming future changes. Hence, subjects were forced to attend on a note-by-note basis and use the number of salient change points to infer duration estimates. This strategy is consistent with other research in which the number of chunks or changes within relatively incoherent events were manipulated and yielded the finding that experienced duration lengthens with an increased amount of information (e.g., Adams, 1977; Block, 1978; Ornstein, 1969; Schiffman & Bobko, 1974).

An additional finding within the comparative judgment data deserves mention. It involves the assumption that du-

ration judgments are determined by both the direction and the amount of contrast between standard and comparison melodies. This assumption was in fact supported when subjects were attending to lower order relations. Melodic pairings showing a greater discrepancy in the relative number of contour changes (i.e., four changes) produced greater under- and overestimations of judged duration than pairings showing less of a discrepancy (i.e., two changes). However, subjects using an anticipatory mode of attending provided only modest evidence suggesting that the degree of judgment bias is directly proportional to the magnitude of temporal contrast. As is indicated in Table 2, magnitude effects did obtain in the positive contrast conditions in which comparisons showing greater contrast in their final phrase periodicity (i.e., 3315) produced greater overestimations than those showing less contrast (i.e., 3324). But this pattern of results did not generalize to the negative contrast conditions. Here, pairings displaying greater temporal contrast were actually less underestimated in relative duration. An explanation for the latter effect is suggested by the higher variability accompanying these pairings and the fact that some subjects ($n = 5$) reported that the 3351 melody seemed longer than its standard. This may be due to the relative timing of the third phrase periodicity (encompassing five tonal groups), which spans to the near end of the melody. This significant lengthening relative to preceding phrase periodicities may have seemed so long to subjects that it became the more salient contrast used in duration estimates, as opposed to the last phrase ending a melody. This suggests that the function relating temporal contrast to judgment bias may be nonlinear, and the model offered by Jones and Boltz (1989) in fact allows for this possibility.

In sum, the overall pattern of results observed here indicates that the filled interval effect may be influenced by different attentional perspectives of the same event. Anticipatory attending generates expectancies targeted toward the beginnings and ends of nested periodicities and elicits time estimates dependent on temporal contrast. Conversely, attending focused on incoherent structural relations yields time estimates that reflect counting strategies. These two modes of attending appear to be independent of one another in that variations of relative contour change exerted no impact on judgments based on temporal contrast, and vice versa. This created situations in which the same event could produce relative over- or underestimations of duration or even judgments of equivalent duration, depending on which level of information was selectively tracked.

The musical stimuli used in this experiment displayed coherent relations at relatively high levels in the structural hierarchy but incoherent ones at the lower level of contour change. Although this type of arrangement frequently occurs in music, it does not exclude other kinds of structural formats. As noted by Narmour (1989), lower level predictability can arise from many sources, including regularly recurrent changes in contour and interval-

lic progressions. Thus, a listener attending to these lower order relations is able to generate structurally based expectancies whose confirmation/violation may later affect relative duration estimates. Similarly, the same melody may show incoherent relations at higher levels (i.e., unpredictable melodic phrasing), which in turn promote counting strategies. In sum, the type of time-estimation strategy used for a given situation does not strictly depend on the hierarchical level to which one is attending but on the degree of predictability at that particular level.

Many factors, both environmental and organismic, can determine the mode of attending adopted for a given event. Incoherent events, as mentioned earlier, encourage lower order attending because higher order predictability is lacking. Thus, the music of a novice pianist or the speech of a foreigner will typically direct attending toward local items within the event's time span. Of more interest are coherent events that allow a perceiver to flexibly shift attending from predictably recurrent phrase periodicities to lower level details.

Here, attending will often be captured by the array of temporal accents and the higher order level of nontemporal information that is perceptually highlighted and hierarchically outlined. This nontemporal information often appears to be the structural level that embodies the primary meaning or "gist" of an event. In speech, for example, primary accents mark grammatical phrase boundaries (Butterworth, 1975; Goldman-Eisler, 1968; Grosjean et al., 1979) and high-content words, whereas in behavioral schemes, reductions in rate and amplitude of movement mark the beginnings and ends of constituent action units (Stetson, 1905; Todd, 1985). Analyses of musical performances reveal similar findings. Numerous studies have shown that pianists display salient changes in amplitude, tempo, and tonal prolongation at melodic phrase boundaries nested within the underlying tonality scheme (e.g., Clarke, 1982, 1988; Gabrielsson, 1988; Shaffer, 1981, 1984; Sundberg, 1988; Todd, 1985). Such findings suggest that the individual producing an event may rely on a hierarchical motor scheme that temporally highlights the intended meaning. To a perceiver, it is adaptive to focus on this higher level of structure, not only to discern the underlying meaning and organization of the event, but to generate expectancies about the future course of an event and initiate preparatory responses. This in turn facilitates interactive behavior and may underlie turn-taking in conversational discourse (Duncan, 1972, 1973) and the exchange of lead instruments during a concert performance (Rasch, 1979, 1988).

Other circumstances, however, may encourage a perceiver to lock into lower level details within an event. A lack of familiarity or perceptual learning may initially encourage this mode of attending because the intrinsic organization of the event has not yet been discerned. This frequently happens, for example, when one hears a new style of music for the very first time. Attending may also shift downward because of "attention-grabbing" devia-

tions at this level (e.g., mistuned notes, speech slurring) or to derive a subsidiary meaning from the event (e.g., determining the speaker's dialect or state of sobriety; the virtuosity of a musical performance). Finally, it may be that higher levels of internal arousal shift one's preferred tempo so that attending is directed downward to relatively faster moving periodicities of a smaller time span. This may underlie the time distortions that often accompany states of high arousal due to stress, stimulant drugs, or certain moods (Orme, 1969).

In closing, the perspective offered here provides a framework from which to link event coherence and dynamic attending modes to time-estimation activities. It incorporates the strategies suggested by previous models into an approach that also recognizes a wider range of natural event structure and the attentional perspectives they afford. Given that attending modes can be induced through both internal and environmental factors, this framework may ultimately offer some applications in other areas of time-estimation research.

REFERENCES

- ADAMS, R. D. (1977). Intervening stimulus effects of category judgments of duration. *Perception & Psychophysics*, *21*, 527-534.
- ALLAN, L. G. (1979). The perception of time. *Perception & Psychophysics*, *26*, 340-354.
- AVANT, L. L., & LYMAN, P. J. (1975). Stimulus familiarity modifies perceived duration in prerecognition visual processing. *Journal of Experimental Psychology: Human Perception & Performance*, *1*, 205-213.
- BAKER, M. (1989). An artificial intelligence approach to musical grouping analysis. *Contemporary Music Review*, *3*, 43-68.
- BERRY, W. (1976). *Structural functions in music*. Englewood Cliffs, NJ: Prentice Hall.
- BLOCK, R. A. (1974). Memory and the experience of duration in retrospect. *Memory & Cognition*, *2*, 153-160.
- BLOCK, R. A. (1978). Remembered duration: Effects of event and sequence complexity. *Memory & Cognition*, *6*, 320-326.
- BLOCK, R. A. (1979). Time and consciousness. In G. Underwood & R. Stevens (Eds.), *Aspects of consciousness: Vol. 1. Psychological issues* (pp. 179-217). London: Academic Press.
- BLOCK, R. A. (1985). Contextual coding in memory: Studies of remembered duration. In J. A. Michon & J. L. Jackson (Eds.), *Time, mind, and behavior* (pp. 169-178). Heidelberg: Springer.
- BLOCK, R. A. (1989). Experiencing and remembering time: Affordances, context, and cognition. In I. Levin & D. Zakay (Eds.), *Time and human cognition: A life-span perspective* (pp. 333-363). Amsterdam: North-Holland.
- BOLTZ, M. (1985). *An expectancy model of judged duration: An ecological perspective*. Unpublished doctoral dissertation, Ohio State University.
- BOLTZ, M. (1989). Time judgments of musical endings: Effects of expectancies on the "filled interval effect." *Perception & Psychophysics*, *46*, 409-418.
- BOLTZ, M., & JONES, M. R. (1986). Does rule recursion make melodies easier to reproduce? If not, what does? *Cognitive Psychology*, *18*, 389-431.
- BOWER, G., BLACK, J., & TURNER, J. (1979). Scripts in memory for text. *Cognitive Psychology*, *11*, 177-220.
- BUFFARDI, L. (1971). Factors affecting the filled-duration illusion in the auditory, tactual, and visual modalities. *Perception & Psychophysics*, *10*, 292-294.
- BURNSIDE, W. (1971). Judgment of short time intervals while performing mathematical tasks. *Perception & Psychophysics*, *9*, 404-406.

- BUTTERWORTH, B. (1975). Hesitation and semantic planning in speech. *Journal of Psycholinguistic Research*, 4, 75-87.
- CARLSOO, S. (1972). *How man moves* (W. P. Michael, Trans.). New York: Cran, Russak, & Co.
- CLARKE, E. (1982). Timing in the performance of Erik Satie's "Vexations." *Acta Psychologica*, 50, 1-19.
- CLARKE, E. (1988). Generative principles in music performance. In J. Sloboda (Ed.), *Generative processes in music: The psychology of performance, improvisation, and composition* (pp. 1-26). Oxford, U.K.: Clarendon Press.
- CONE, E. (1968). *Musical form and musical performance*. New York: Norton.
- DEUTSCH, D., & FEROE, J. (1981). The internal representation of pitch sequences in tonal music. *Psychological Review*, 85, 341-354.
- DEVANE, J. (1974). Word characteristics and judged duration for two response sequences. *Perceptual & Motor Skills*, 38, 525-526.
- DUNCAN, S. (1972). Some signals and rules for taking speaking turns in conversation. *Journal of Personality & Social Psychology*, 23, 283-292.
- DUNCAN, S. (1973). Toward a grammar for dyadic communication. *Semiotica*, 9, 29-47.
- DUNN, O. J. (1961). Multiple comparisons among means. *Journal of the American Statistical Association*, 56, 52-64.
- FRAISSE, P. (1984). Perception and estimation of time. *Annual Review of Psychology*, 35, 1-36.
- FRANKENHAUSER, M. (1959). *Estimation of time: An experimental study*. Stockholm: Almqvist & Wiksell.
- GABRIELSSON, A. (1988). Timing in music performance and its relations to music experience. In J. Sloboda (Ed.), *Generative processes in music: The psychology of performance, improvisation, and composition* (pp. 27-51). Oxford, U.K.: Clarendon Press.
- GIBSON, J. (1966). *The senses considered as a perceptual system*. Boston: Houghton Mifflin.
- GIBSON, J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.
- GOLDMAN-EISLER, F. (1968). *Psycholinguistics: Experiments in spontaneous speech*. London: Academic Press.
- GROSJEAN, F., GROSJEAN, L., & LANE, H. (1979). The patterns of silence: Performance structures in sentence production. *Cognitive Psychology*, 11, 58-81.
- HICKS, R., MILLER, G., GAES, J., & BIERMAN, K. (1977). Concurrent processing demands and the experience of time-in-passing. *American Journal of Psychology*, 90, 431-446.
- JAFFE, J., & FELDSTEIN, S. (1970). *Rhythms of dialogue*. New York: Academic Press.
- JONES, M. R. (1976). Time, our lost dimension: Toward a new theory of perception, attention, and memory. *Psychological Review*, 83, 323-335.
- JONES, M. R. (1981a). Only time can tell: On the topology of mental space and time. *Critical Inquiry*, 7, 557-576.
- JONES, M. R. (1981b). A tutorial on some issues and methods in serial pattern research. *Perception & Psychophysics*, 30, 492-504.
- JONES, M. R. (1987). Perspectives on musical time. In A. Gabrielsson (Ed.), *Action and perception in rhythm and music* (pp. 153-175). Stockholm: Royal Swedish Academy of Music.
- JONES, M. R., & BOLTZ, M. (1989). Dynamic attending and responses to time. *Psychological Review*, 96, 459-491.
- KRAMER, J. (1982). Beginnings and endings in Western art music. *Canadian University Music Review*, 3, 1-14.
- 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.
- MANDLER, J., & JOHNSON, N. (1977). Remembrance of things parsed: Story structure and recall. *Cognitive Psychology*, 9, 111-151.
- MARTIN, J. (1972). Rhythmic (hierarchical) versus serial structure in speech and other behavior. *Psychological Review*, 79, 487-509.
- MARTIN, J. (1975). Rhythmic expectancy in continuous speech perception. In A. Cohen & S. Nooteboom (Eds.), *Structure and process in speech perception* (pp. 161-176). New York: Springer.
- MEYER, L. (1956). *Emotion and meaning in music*. Chicago: University of Chicago Press.
- MONAHAN, C., & CARTERETTE, E. (1985). Pitch and duration as determinants of musical space. *Music Perception*, 3, 1-32.
- NARMOUR, E. (1977). *Beyond Schenkerism: The need for alternatives in music analysis*. Chicago: University of Chicago Press.
- NARMOUR, E. (1989). The "genetic code" of melody: Cognitive structures generated by the implication-realization model. *Contemporary Music Review*, 4, 45-63.
- NEWTSON, D. (1973). Attribution and the unit of perception of ongoing behavior. *Journal of Personality & Social Psychology*, 28, 28-38.
- NEWTSON, D. (1976). Foundations of attribution: The perception of ongoing behavior. In J. Harvey, R. Ickes, & R. Kidd (Eds.), *New directions in attribution research* (Vol. 1, pp. 223-248). Hillsdale, NJ: Erlbaum.
- NEWTSON, D., HAIRFIELD, J., BLOOMINGDALE, J., & CUTINO, S. (1987). The structure of action and interaction. *Social Cognition*, 5, 191-237.
- ORME, J. (1969). *Time, experience, and behavior*. New York: Elsevier.
- ORNSTEIN, R. (1969). *On the experience of time*. Baltimore, MD: Penguin.
- PIERSON, K. (1976). Control of walking. *Scientific American*, 235, 72-86.
- PISTON, W. (1978). *Harmony* (4th ed., revised and expanded by M. DeVoto). New York: Norton.
- POYNTER, W. D. (1983). Duration judgment and the segmentation of experience. *Memory & Cognition*, 11, 77-82.
- POYNTER, W. D., & HOMA, D. (1983). Duration judgment and the experience of change. *Perception & Psychophysics*, 33, 548-560.
- RASCH, R. (1979). Synchronization in performed ensemble music. *Acustica*, 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, U.K.: Clarendon Press.
- RUMELHART, D. (1977). Understanding and summarizing brief stories. In D. LaBerge & S. J. Samuels (Eds.), *Basic processes in reading: Perception and comprehension* (pp. 265-303). Hillsdale, NJ: Erlbaum.
- SCHANK, R. (1982). *Dynamic memory*. New York: Cambridge University Press.
- SCHIFFMAN, H. R., & BOBKO, D. J. (1974). Effects of stimulus complexity on the perception of brief temporal intervals. *Journal of Experimental Psychology*, 103, 156-159.
- SHAFFER, L. H. (1981). Performances of Chopin, Bach, and Bartok: Studies in motor programming. *Cognitive Psychology*, 13, 326-376.
- SHAFFER, L. H. (1984). Timing in solo and duet piano performances. *Quarterly Journal of Experimental Psychology*, 36A, 577-595.
- SHIELDS, J., MCHUGH, A., & MARTIN, J. (1974). Reaction time to phoneme targets as a function of rhythmic cues in continuous speech. *Journal of Experimental Psychology: Human Perception & Performance*, 102, 250-255.
- STETSON, R. (1905). A motor theory of rhythm and discrete succession. *Psychological Review*, 12, 250-350.
- SUNDBERG, J. (1988). Computer synthesis of music performance. In J. Sloboda (Ed.), *Generative processes in music: The psychology of performance, improvisation, and composition* (pp. 52-69). Oxford, U.K.: Clarendon Press.
- THOMAS, E. A. C., & WEAVER, W. B. (1975). Cognitive processing and time perception. *Perception & Psychophysics*, 17, 363-367.
- THORNDYKE, P. (1977). Cognitive structures in comprehension and memory of narrative discourse. *Cognitive Psychology*, 9, 77-110.
- TODD, N. (1985). A model of expressive timing in tonal music. *Music Perception*, 3, 33-58.
- TODD, R., BOLTZ, M., & JONES, M. R. (1989). The MIDILAB auditory research system. *Psychomusicology*, 8, 83-96.
- UNDERWOOD, G., & SWAIN, R. (1973). Selectivity of attention and the perception of duration. *Perception*, 2, 101-105.

- WINTER, D. (1983). Biomechanical motor patterns in normal walking. *Journal of Motor Behavior*, 15, 302-330.
- YESTON, M. (1976). *The stratification of musical rhythm*. New Haven, CT: Yale University Press.

NOTES

1. The term *event* has multiple meanings in the psychological literature. From the tradition of psychophysics, *event* often refers to a singular item (e.g., a note or a word) within a larger sequence. Here, however, the term refers to an ecologically valid object undergoing change or physical motion (Gibson, 1966, 1979). In this context, then, *event* refers to any environmental object or activity (e.g., a melody or a speech utterance) that unfolds over a given time span and displays an intrinsic arrangement of structural relations.

2. Melodic phrases that conform to a strict hierarchical structure are most commonly found within folk tunes and nursery rhymes. Nevertheless, melodic phrasing in other genres of music also tends to display a very lawful arrangement in which subsequent phrases are related to earlier ones through ratio relations (see, e.g., Baker, 1989; Clarke, 1988; Lerdahl & Jackendoff, 1983). Thus, even though the strict hierarchical nesting of structure shown in Figure 1 is somewhat idealized, it illustrates the general notion that music contains a high degree of predictability in which temporal and melodic structure are lawfully related to one another.

3. The values of *next* rules are normally indicated in terms of scale steps so that the relationship between (e.g.) C₄ and D₄ would be described as N¹. However, given that melodies within the present experiment always contain some atonal notes outside the C major diatonic scale, rule values referring to scale steps are no longer appropriate. Hence, the value of *next* rules was altered to refer to semitone units so that the relationship between C₄ and D₄ is now described as N².

4. The melodies originally constructed for this experiment displayed more conventional phrase endings marked by tonic triad members (i.e., CEG). However, a preliminary pilot study revealed that subjects had difficulty tracking these types of tonal markers during an ongoing monitoring task. This is probably due to the intrinsic nature of rule-generated melodies and their relative lack of harmony and interesting melodic lines. However, rule-generated melodies were deemed necessary to the rigorous control of melodic structure required by the present experimental design. Hence, tonal mistakes were introduced as markers of phrase endings, since listeners reported these to be quite noticeable and easy to detect during an unfolding melody.

5. An exact definition of pairwise temporal contrast for use in comparative judgment tasks is provided by Jones and Boltz (1989).

6. In the upper portion of Table 2, notice that standard melodies containing three tonal mistakes (i.e., phrase periodicities of 444) are specific to conditions of negative temporal contrast, and that standards containing six tonal mistakes (i.e., 222222 phrase periodicity) are unique to conditions of positive temporal contrast. A more extensive design involving pairings of all possible pattern combinations not only would result in an extraordinary number of experimental trials but also would yield pairings open to alternative explanations. Within a pairing of 444-3324, for example, the comparison melody might seem longer than its standard because (1) the comparison contains more tonal mistakes marking phrase periodicities (as in Block, 1979, or Ornstein, 1969) or (2) violation of expectancies results in positive temporal contrast (as in Jones & Boltz, 1989). Since these types of pairings fail to distinguish among alternative theoretical predictions, they were excluded from the present experimental design.

(Manuscript received July 2, 1990;
revision accepted for publication December 20, 1990.)