

## Dynamic Attending and Responses to Time

Mari Riess Jones  
Ohio State University

Marilyn Boltz  
Haverford College

A temporally based theory of attending is proposed that assumes that the structure of world events affords different attending modes. *Future-oriented attending* supports anticipatory behaviors and occurs with highly coherent temporal events. Time judgments, given this attending mode, are influenced by the way an event's ending confirms or violates temporal expectancies. *Analytic attending* supports other activities (e.g., grouping, counting), and if it occurs with events of low temporal coherence, then time judgments depend on the attending levels involved. A weighted contrast model describes over- and underestimations of event durations. The model applies to comparative duration judgments of equal and unequal time intervals; its rationale extends to temporal productions/extrapolations. Two experiments compare predictions of the contrast model with those derived from other traditional approaches.

One characteristic of modern society is a preoccupation with fixed time schedules and standardized timekeepers. We maintain appointments at hourly intervals, rush to meet the 5:00 p.m. bus, and dine at predetermined hours. Yet our natural ability to judge time remains poorly understood. How often do we estimate the time elapsed since last glancing at a clock and discover with surprise that we were fairly accurate? Surprise is understandable because at least as often we lose track of time and err. The validity of these impressions is confirmed by laboratory research showing that duration judgments depend not only on actual physical duration but also on a variety of nontemporal factors. These include the spatial layout and complexity of an event as well as the attentional set, skill, affect, and constitutional state of the judge (Allan, 1979; Fraisse, 1984; Kristofferson, 1984).

Researchers have addressed many of these issues that include both psychophysical problems (e.g., Weber's Law for time discrimination) and organismic variables (e.g., age, drugs, and arousal effects). Of recent interest is the influence of nontemporal information on time judgments, due largely to a fascination with such problems as the *filled interval effect*. This phenomenon reveals that two equivalent time intervals may not be judged as such because of the nontemporal information that fills them. Although the most popular models of judged duration attempt to explain this effect (e.g., Block, 1978; Ornstein, 1969), the effect itself raises problems for a general theory of time estimation (Allan, 1979).

In this article we focus on some problems raised by the filled

interval effect. We consider these and related issues concerning responses to time from a more general perspective, one based on an analysis of event time itself and dynamic aspects of attending. We suggest that events define time intervals and that their inherent rhythmic patternings will affect the way people attend to them and judge their durations. The general framework leads to hypotheses about duration judgments of both equal and unequal time intervals as well as temporal extrapolations. Although these hypotheses cannot explain all of the diverse facts of time estimation, they do suggest ways of linking a general theory of attending to specific models of time judgment and discrimination. In this, our approach is not intended to usurp contemporary time models but to incorporate some of their assumptions into a more inclusive framework.

This article has five parts. Part 1 introduces some contemporary time models. Part 2 outlines a theoretical approach to event time and dynamic attending, and Part 3 returns to time estimation and poses specific hypotheses about dynamic attending in various tasks. Part 4 describes two experiments relevant to these hypotheses that present difficulties for contemporary time models. Part 5 concludes with additional theoretical implications of this approach.

### Part 1: Contemporary Theories and Issues

In Part 1, several contemporary models of judged duration are presented along with relevant empirical support. These models primarily address the case in which people judge identical durations and errors give rise to the filled interval effect. One influential model is Ornstein's (1969).

#### *Ornstein's Storage Size Hypothesis*

Influenced by Frankenhauser's (1959) original model, Ornstein (1969) further developed the *storage size hypothesis* to explain the estimation of time periods lasting about 10 s or more. According to this view, two equal time intervals will seem to have different durations if one is more complex and therefore requires more storage space in memory: "The central metaphor is that the experience of duration of an interval is a construction

---

This research was supported by Grant BNS-8204811 from the National Science Foundation and by a fellowship from the Netherlands Institute for Advanced Study awarded to the senior author (1986-1987).

The authors thank Chris Antons, David Butler, Walter Johnson, Gary Kidd, Kerri Marsh, Elizabeth Marshburn, John Michon, Mitch Pratt, Ken Pugh, Jackie Ralston, and Wicher van Vreden. Special thanks are due to Steve Handel and two anonymous reviewers for their excellent comments on an earlier version of this article.

Correspondence concerning this article should be addressed to Mari Riess Jones, Department of Psychology, The Ohio State University, 142 Townshend Hall, Columbus, Ohio 43210.

formed from its storage size. As storage size increases, duration experience lengthens" (Ornstein, 1969, p. 42).

Experiments supporting this view have relied on various definitions of stimulus complexity (e.g., number of angles in visual figures, stimulus arrangement in time and space) and found that the experienced duration of intervals between 30 s and 10 min does, in fact, lengthen with increasing stimulus complexity (Hogan, 1975; Ornstein, 1969; Schiffman & Bobko, 1974). Others, however, have reported conflicting results. For example, both Block (1974) and Poynter (1983) have found that a sequence of words is judged longer when words are grouped by semantic category than when randomly arranged. Because the latter presumably contains more chunks, this contradicts Ornstein's model. As a result of these discrepancies, models assuming a different referent for judged duration have been proposed.

### *Attentional Effort Models*

A model offered by Underwood and Swain (1973) posits that duration judgments are mediated by attentional effort. They tested Ornstein's (1969) prediction that increased attention leads to more stored information and hence to longer time estimates. Attentional demands were varied independently of information content in a vigilance task in which subjects detected target digits embedded in prose passages that were partially masked by various white noise intensities. When unexpectedly asked to judge the relative duration of each passage, subjects reported those masked by a high-intensity noise (i.e., requiring more attention) to be longer than those masked by a low-intensity noise. However, contrary to Ornstein's storage analysis, detection levels indicated that *less* information was encoded in highly masked passages.

Such findings support the idea that experienced duration is less dependent on memory load than on attentional effort or arousal associated with presented information. Others, using very brief stimuli, have reached similar conclusions (Thomas & Cantor, 1978; Thomas & Weaver, 1975).

### *Contextual Change Model*

Another challenge to Ornstein's (1969) proposal comes from Block's (1978, 1985, in press) *contextual change hypothesis*. Judged duration is hypothesized to increase as a linear function of the number of contextual changes occurring in both the environmental situation (e.g., changes in stimulus properties, task demands) and in the organism (e.g., mnemonic activities). Changes are monitored by an internal cognitive device that later outputs a complexity index based on the total number of changes within a time interval.

In one test of this model, Block and Reed (1978) required subjects to encode word lists at different levels of processing (à la Craik & Lockhart, 1972). For example, some people judged the typing style of words (i.e., a shallow task) or categorized words into semantic categories (i.e., a deep task), whereas others alternated between both tasks. Afterward, all of the subjects were unexpectedly asked to judge which activity seemed longer. Both the storage size and attentional effort models predict that deep processing (more information, more effort) should produce longer time estimates. This was not the case. When people

alternated shallow-deep strategies (i.e., more changes), duration seemed longer. They interpreted such findings as support for Block's change hypothesis and as problematic for both the storage size and attentional effort models.

### *Evaluation of Current Models*

How successful are these models? Each is quite successful within certain contexts. However, all share certain empirical and theoretical limitations.

Empirical difficulties stem from seemingly conflicting results that emerge from the literature as a whole. An interval defined as more complex is sometimes judged longer, but on other occasions is judged shorter than a less complex one. For example, divergent results have been observed with the variables of stimulus familiarity (e.g., Avant & Lyman, 1975, vs. Devane, 1974), task difficulty (e.g., Burnside, 1971, vs. Underwood & Swain, 1973), and stimulus arrangement (e.g., Poynter, 1983, vs. Schiffman & Bobko, 1974).

Some conflicting findings may be due to methodological differences. Others, however, suggest a need for re-evaluating certain tacit assumptions of the models themselves (also see Block, in press). In particular, three issues are relevant.

The first involves the presumed referent for duration judgments. Each model assumes that time judgments are inferred from the amount of some processing activity. However, this processing activity strictly refers to the *nontemporal* information that fills the stimulus interval: the number of spatial angles, the arrangement of word lists, the amount of background noise, and so on. But the *temporal* information within an event and its impact on behavior is ignored.

A second problem concerns complexity. Complexity is assumed to increase the amount of processing activity and thereby lengthen experienced duration. But it isn't always clear why. The relationship of psychological complexity to the stimulus or its duration is rarely fully developed. For example, the storage size hypothesis claims that complexity is determined by the number of memory chunks. Although recent coding theories add formalization by equating complexity with memory code length (e.g., Deutsch & Feroe, 1981; Leeuwenberg, 1969; Simon, 1972), it remains a difficult construct. What exactly is a chunk, and what stimulus or task characteristics determine chunk boundaries? Such questions have never been satisfactorily answered. Complexity determinants of attentional effort and cognitive change are equally elusive. In a speech utterance, for example, there are changes in sound, grammar, meaning, and intonation, and yet we don't know which kind or how much of a change is required to affect judged duration. Because contemporary time models lack precise definitions of complexity, this may partially explain the conflicting nature of experimental results.

A final problem concerns the choice of experimental stimuli. During the course of a day, we interact with friends, drive, listen to music, and so on. Yet, these kinds of events are rarely selected for study. Instead, subjects must compare intervals filled with static abstract drawings (Ornstein, 1969), clicks (Adams, 1977), and lists of unrelated words or nonsense words (Poynter, 1983). Although the latter may offer tight experimental control and may in fact represent certain everyday experiences, they fail to reflect the full range of stimuli we routinely experience.

Commonplace events differ from those typical of traditional time studies in several ways, notably in their structure and function. With respect to structure, events such as speech utterances, musical patterns, and body movements contain much more structural coherence in time than those of current research. Typically, these events display multiple levels of interrelated structure that unfold predictably over a given time span. Often there is distinctiveness as natural time patterns, including special beginnings and endings, characteristic tempi, and rhythms. All of this contributes to temporal predictability. And predictability allows perceivers to anticipate an event's future course, including when in time it should end. In interactive speech, for example, the smooth exchange of speaking roles and turn-taking behavior suggests that people anticipate ends of utterances. In other cases, event structure can communicate mood or intention. It is possible that the structure and function of events systematically affect time estimation.

In sum, contemporary research on time estimation is characterized by conflicting experimental findings. At the same time, questions can be raised about the structure, function, and representativeness of stimuli used in this research. We suggest that these are related problems. Divergent experimental findings arise because variations in the structure of temporal events and the ways people respond to them have not been considered.

### *Overview of an Alternative Hypothesis of Judged Duration*

An alternative perspective takes its cue from the idea that events are, by definition, temporal and that their structure in time is critical. That is, the temporal patterning of nontemporal information (e.g., words, tonal pitches, lights, and even haphazard items) within any interval is critical in determining *how* one attends to the event itself. Societal and individual needs ensure that all kinds of events are encountered, but we claim that people attend differently to events with high and low structural coherence, and that this, in turn, differentially affects time estimates. In brief, we propose a distinction between two different modes of dynamic attending, future-oriented and analytic, that can bias time estimates of events with high and low coherence, respectively.

Events<sup>1</sup> differ in terms of their structural coherence and predictability. And although we assume a continuum of coherence, for convenience we distinguish between events with high and low temporal coherence. Highly coherent events, such as those of speech, tonal music, and body gestures, offer structural predictability and display characteristic rhythmic patterns that occur over nonarbitrary time spans. Events of low coherence, such as a list of unrelated words or simultaneous cocktail-party chatter, unfold over arbitrary time spans that contain little structural predictability.

Highly coherent events afford future-oriented attending. Because they offer high temporal predictability, people can track and use higher order time patternings to generate expectancies about how and when they will end. In Western music, for example, notes within an unfolding melody occur in a temporally ordered fashion, often with such coherence that listeners can anticipate not only *what* notes are likely but also *when* in time they "should" occur. Thus, future-oriented attending exploits the global time structure of such events. In these situations, we

propose that time estimates are determined by the confirmation or violation of expected ending times. When two events of equivalent duration both end when expected, people will correctly judge them to be the same duration. However, if one violates an expectancy by seeming to end later than anticipated, then it will be incorrectly judged as longer. Similarly, an event appearing to end *too early* will be judged as relatively short. This reasoning extends to judgments of events that actually do differ in duration. According to this view then, duration estimates of coherent events are biased by temporal contrast where contrast involves an apparent temporal disparity between an event's actual and expected ending.

Analytic attending occurs with less coherent events. These events have low temporal predictability, and so people cannot anticipate their future course. Instead, they are forced to attend locally to adjacent elements in an attempt to organize the unstructured information. Depending on the task, this kind of attending supports strategies directed toward lower level relationships (e.g., grouping or counting the number of items or changes within the event). Finally, to estimate the duration of these events, people will be biased by their attention to local details and will judge events filled with more items to be longer.

Both future-oriented and analytic attending are dynamic ways of interacting with temporal events. In this article the emphasis will be on future-oriented attending and its influence on time estimates, inasmuch as this topic has received less attention than the memory-oriented approach found in contemporary time models.

### *Summary of Part 1*

Contemporary time models have focused on the filled interval effect in which nontemporal information distorts judgments about equal time intervals. Their explanations rely on memory-based processes that gauge the total amount of a nontemporal construct (i.e., complexity, effort, change). But these approaches encounter problems of divergent empirical support, imprecise definitions of complexity, and lack of stimulus representativeness. An alternative view is presented that is more explicitly temporal. It proposes two modes of dynamic attending (future-oriented, analytic) that occur with events of high and low temporal coherence, respectively. This model is designed to accommodate diverse experimental findings by supplementing memory-oriented attending (analytic) with future-oriented attending.

### **Part 2: Temporal Environment and Attunements**

In the remainder of this article, we have three goals. The first is to show that analyses of event structure are critical to theories

<sup>1</sup> Cognitive approaches have traditionally referred to world objects and experimental materials as *stimuli*. In protest to implications regarding the impoverished nature of such materials, some recent approaches have adopted the term *event* to refer to ecologically valid objects undergoing physical motion or change (Gibson, 1966, 1979). Our framework reflects both notions in that an event refers to any environmental object or activity that varies along a continuum of structural coherence. Some natural events are very predictable in that embedded temporal relationships are lawfully interrelated with the event's nontemporal information. Other natural events are less coherent and predictable in this regard and, in fact, may contain little internal structure (e.g., silences).

of attending. The second is to describe distinct ways in which people can attend to events that vary in temporal coherence. The third is to demonstrate that both event structure and attending mode must be considered when building theories of time estimation.

Part 2 addresses the first goal. Any interpretation of responses to time should be based on a theory about what time means in events and to people who perceive these events. In this approach, time means relative time. Our formalization leads to dynamic conceptions of event structure, temporal predictability, and complexity. It also leads to hypotheses about the ease with which people can use dynamic structure to attend in various ways, some of which operate in time estimations.

### *Theoretical Background*

A relativistic approach to time implies that points in time and absolute time intervals are less important than time periods determined relative to other time periods (rhythmic structure) and time periods determined relative to spatial extents (velocity or flow structure; Jones, 1976). These time relations are inseparable from the event itself. This idea is at the heart of our assertion that people may be unreliable when judging absolute lengths of arbitrary or isolated time intervals.

Our theoretical emphasis on relative time is consistent with an ecological analysis of event structure in terms of transformations (changes) and invariants (nonchanges) and their functions for the organism (Freyd, 1987; Gibson, 1979; Jones, 1976; Shaw, McIntyre, & Mace, 1974; Shepard, 1984; Turvey & Carello, 1981). People and other living things both create and respond to temporal event structure found in conversations, dance, music, and so on. We suggest further that time transformations often support these interactions in special ways. This leads to a different approach to complexity and, ultimately, to the way people respond to dynamic event structure. To preview, in subsequent sections we propose that events characterized by certain time transformations are easier to anticipate in time (Jones, 1976, 1981a, 1982; D. N. Lee, 1980; Michon & Jackson, 1985).

### *Relative Time in Environmental Structure*

Our environment is filled with all sorts of temporal events, some based on activities of living things and others not. All lie on a continuum of temporal coherence ranging from highly arbitrary (low coherence) to nonarbitrary (high coherence). The gist of temporal coherence was presented in Part 1 where it was related to temporal predictability and attending mode. Here, we introduce greater formalization by distinguishing between temporally coherent, or hierarchical (H) time structures, and temporally incoherent, or nonhierarchical (NH) ones. Ultimately, complexity is dynamically conceived in terms of these time structures.

We begin with some examples of temporally coherent (hierarchical) events in Figure 1a and 1b. Both examples illustrate our use of the term *coherent* to imply objective accent regularities (i.e., not necessarily subjective experiences). In each, nontemporal information is distributed in ways that convey much predictability in time. In Figure 1a, a musical phrase of 3,200 ms is prominently outlined by salient opening and closing

itches ( $C_4$ ).<sup>2</sup> Lawfully embedded within this span are other time periods (e.g., 1,600, 800 ms), also significantly identified by onsets of various pitches. The smallest time period (200 ms) is determined by the smallest observable tone-to-tone onset interval. Another hierarchical event is illustrated in Figure 1b, in which cycles of a cat's locomotion sequence are shown to embed on-ground time periods (e.g., 270 ms) and smaller periods. Notice that the former are more strongly marked by co-occurring downsteps of two legs (from left hind [LH] with right front [RF] ↓ to left front [LF] with right hind [RH] ↑). Below each event is an abstraction of its nesting properties showing relevant sets of time periods ( $n = 0, 1, 2 \dots$ ). Onsets of more prominent time levels are usually marked by more salient nontemporal information (more dots at, e.g., locations x and y).

These examples highlight two points. One involves identification of distinctive nontemporal markings of various time levels. The other involves clarification of lawfulness in temporal nestings. The issue of distinctive markings is an experimental problem to which we return in later sections. The issue of lawfulness concerns time rules that we formalize as *time transformations*.

A time transformation changes one time period into others while preserving some temporal property as invariant (Jones, 1976, 1981c). Thus, if a 200-ms time period is nested within one of 400 ms, and this, in turn, is nested in an 800-ms period, as in Figure 1a, then one of the simplest time transformations results in which all time levels are related in a binary fashion via multiples of 2 (i.e., period doubling). This sort of change is a *ratio time transformation*. It is one of two types of transformations that enter into temporal coherence. The other is an *additive time transformation*. If 200 ms is changed into 250 ms and then 300 ms, an invariant additive change of +50 ms exists. Often, additive time changes enter into temporal nestings as variations within a given level. For example, in Figure 1b the level  $n = 0$  (50 ms) reflects a recurrent change of -85 ms with respect to the (missing) periodicity of level  $n = 1$  (135 ms; dashed line). In contrast to ratio time transformations, which operate vertically to change level  $n$  into  $n + 1$  (or vice versa), additive transformations operate horizontally to modulate periodicities recurring at some fixed level.

<sup>2</sup> A musical scale is defined by a set of pitch classes and relationships among these pitches. Of most relevance to Western classical music are diatonic scales, composed of seven pitch classes. Diatonic scales are formed by a series of pitch changes (musical intervals) involving whole steps (w) and half-steps (s), where  $w = 2s$ , and s (also called *semitone*) is a unit difference. The half-step, or semitone unit, is logarithmically defined as a frequency ratio:  $\Delta f/f = .05946$ . A major diatonic scale, beginning on the keynote or tonic, has the following sequence of whole steps and half-steps: w, w, s, w, w, w, s. Thus, the most familiar major diatonic scale is C major, in which successive pitch classes of CDEF-GABC are separated by w, w, s, w, w, w, s. In this scale, successive pitch classes are referred to as scale degrees with C being the first scale degree, D the second, E the third, and so on (see Hahn & Jones, 1981). Some important harmonic relationships within each scale include a major third (e.g., C-E in C-major scale), which corresponds to the pitch changes of (w, w) from an initial referent pitch, and a perfect fifth (e.g., C-G), which corresponds to the scale steps of (w, w, s, w). The mediant (E) is the note that defines a major third, whereas the dominant is the note (G) that defines a perfect fifth above the tonic (C). Together the notes C-E-G form the tonic triad chord in C major.

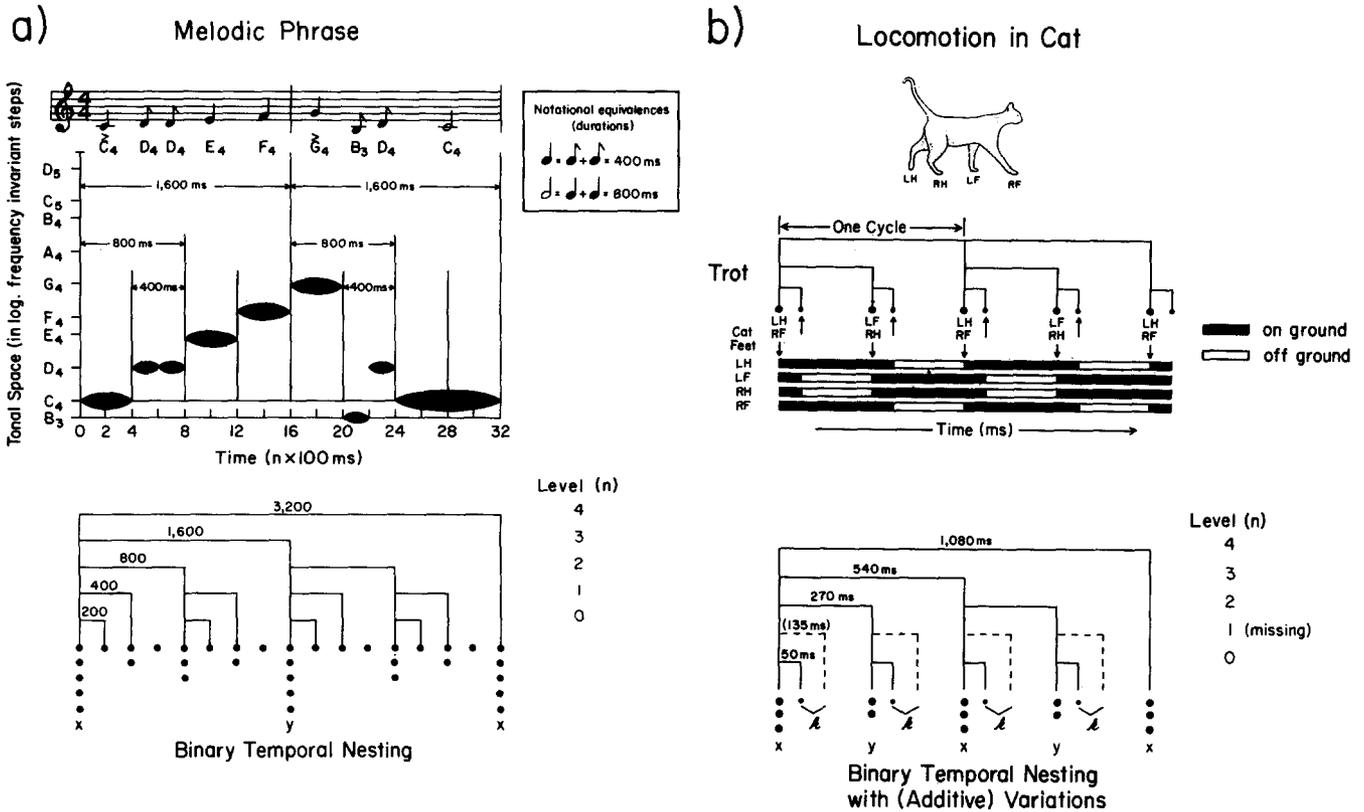


Figure 1. Two coherent events. (Figure 1a is a two-measure melodic phrase in the key of C major. The tonic [C] opens and closes this melody, whereas stress [ $>$ ] and other pitch changes mark various [ratio related] embedded levels. Figure 1b shows on and off ground times of legs in a trotting cat. Coupled down steps mark ratio-related higher order time levels, whereas lower order levels contain an additive time change. Nesting relations are outlined below each event, where more prominent contextual changes are shown by more dots [adapted from Pierson, 1976].)

Ratio and additive time transformations preserve as invariant *time ratios* and *time differences*, respectively. Ratio time transforms involve rhythmic embeddings: time relative to time. Additive time transforms often involve time periods that underlie velocity or flow structure: time change relative to space (Jones, 1987a). Both kinds of relative time contribute to temporal event structure as subsequent formalizations and examples will show.

*Hierarchical and Nonhierarchical Time Structures: Some Formalisms*

Hierarchical time structures are differentiated from nonhierarchical ones by the way nontemporal information is distributed in time to define nesting invariances. In hierarchical events, the total duration of an event is tied to its internal structure through consistent time transformations. This is not so with nonhierarchical events, in which low temporal coherence is conveyed by less regular nestings and by more obscure connections between the total time span and internal event structure.

The distinction between hierarchical and nonhierarchical structures is loose because, as Part 2 illustrates, many degrees and varieties of coherence in time structure exist. Two features

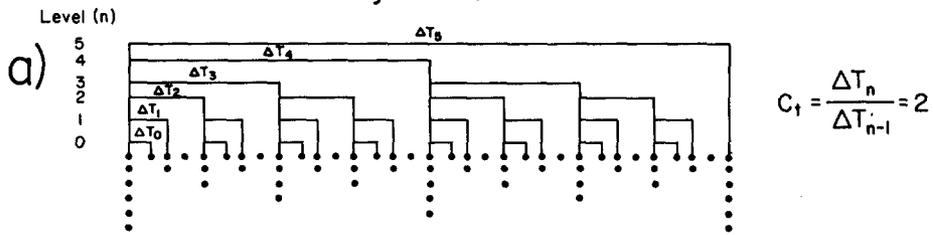
of time structures contribute to this, *time markers* and *distributions of time markers*.

*Time Markers: The Role of Nontemporal Information*

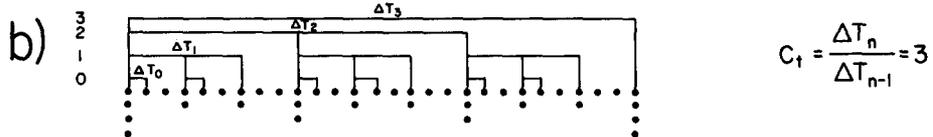
Events contain many nested time periods whose beginnings and ends are intrinsically marked by various structural changes. However, experimental isolation of these markers is tricky because their salience varies with context. That is, markers are defined relationally as contextual change. In auditory events, they involve onsets of unusual frequency or amplitude changes and are termed *accents*. Accents of the first (C<sub>4</sub>) and sixth (G<sub>4</sub>) tone in Figure 1a, for instance, are of these sorts. In addition, harmonic and temporal changes can enhance accent prominence in music (and speech). In visual events, changes in direction and velocity serve similar functions. In any context, nontemporal information has some potential for carving out meaningful time intervals within and between events.<sup>3</sup>

<sup>3</sup> Musically, this is a conservative definition of accentuation. However, it serves the function of acknowledging the importance of serial context and structural changes within a pattern. One goal of this article is to stimulate theory and research that will refine our knowledge of accents, and ways in which they are perceived as such, as a function of serial structure.

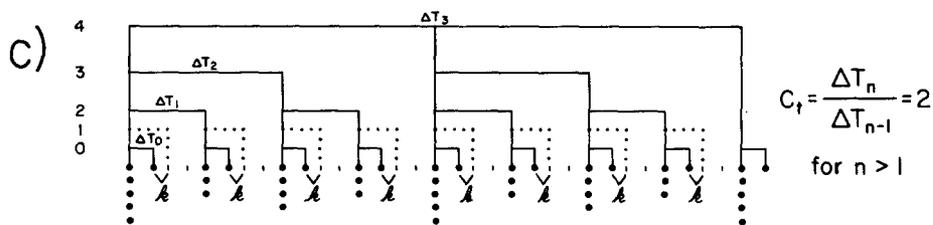
# Hierarchical (rhythmic) Time Structures



Binary Time Structure

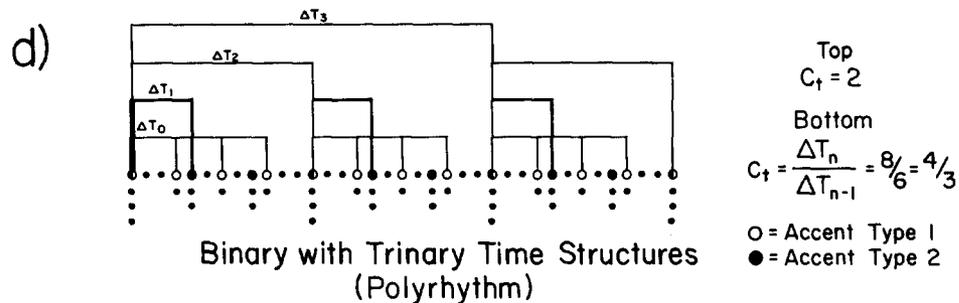


Trinary Time Structure

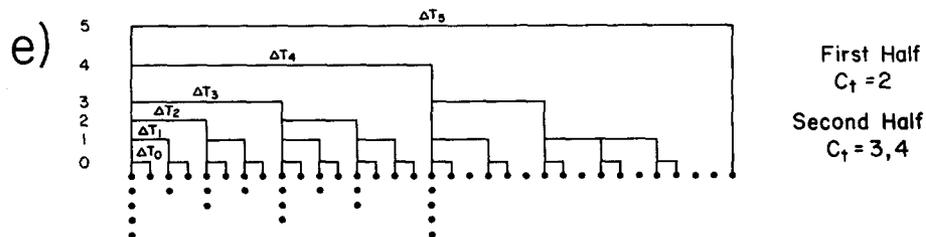


Binary Time Structure with Variations

# Non-Hierarchical Time Structures



Binary with Trinary Time Structures (Polyrhythm)



Non-hierarchical



Non-hierarchical

Accents differ in relative strength (or prominence). In Figure 1 (lower portions) more dots indicate greater accent strength that is associated with greater energy change and/or multiple accent occurrences. Thus, the simultaneous occurrence of amplitude change (i.e., the > stress notation) with certain special pitch changes ( $C_4$  or  $G_4$ ) in Figure 1a, or that of an LH downstep with an RF one in Figure 1b, renders these stronger markings. *Coupled accents* are in fact typically stronger than *decoupled* ones. Furthermore, stronger accents tend to mark higher time levels in coherent time hierarchies (Benjamin, 1984; Jones, 1976, 1987a; Lerdahl & Jackendoff, 1983; Martin, 1972). In what follows, we assume that it is possible to reliably determine functional markers at different time levels.

### Distributions of Time Markers: Role of Time Transformations

The term *hierarchy* has been variously used (Jones, 1981b). Here it refers to a time structure in which the temporal distribution of markers reveals nested time levels that are consistently related to one another at a given level by ratio or additive time transformations. Conversely, to the degree that a distribution of time markers does not reveal this consistency in time transformations, a nonhierarchical time structure of greater dynamic complexity occurs.

In hierarchical events, each nested level is associated with a recurrent time period, denoted by  $\Delta T_n$ , in which  $\Delta T$  refers to the marked time span and  $n$  refers to a level in the hierarchy ( $n = 0, 1, 2, \dots$ ). The smallest time period, denoted  $\Delta T_0$ , occurs at the lowest level and is often marked, especially initially, by onsets of certain adjacent elements. In an ideal hierarchy, other nested periods are related to  $\Delta T_0$  by sets of simple time ratios (or ratio time transformations).<sup>4</sup> Thus, the total duration of a hierarchical event is implied by the time period of any embedded level.

Figure 2 presents three examples of hypothetical time hierarchies that range in temporal coherence from very high to moderate. The first, and simplest, is shown in Figure 2a. It is merely an extension of the binary structure shown in Figure 1a. All nested time levels are related by a ratio time transform of 2. Thus, the smallest level,  $\Delta T_0$ , is related to  $\Delta T_1$  by 2, to  $\Delta T_2$  by 4, and so on, until its total duration is linked to all others. This is a binary time structure. It is hierarchical because the nesting is consistent, meaning that the same ratio (here 2) obtains over recurrent periods at each level. Such a ratio is a time rule and will be denoted as  $C_{in}$ , in which  $C_{in} = \Delta T_n / \Delta T_{n-1}$  (for levels  $n, n-1$ ). This time rule operates vertically to summarize temporal nesting. Finally, in this binary structure,  $C_{in}$  is an invariant in that it is constant over different levels,  $n$ :  $C_{in} = C_i = 2$ .

Simple time hierarchies are based on  $C_i$  values that are small and constant integers (i.e.,  $C_i = 1, 2, 3$ , etc.). When  $C_i = 3$ , a

ternary structure results as in Figure 2b. This hierarchy is less coherent than the binary one because the time ratio involves a larger integer. When different  $C_i$  values appear at different levels (e.g.,  $C_{im} = 2$ , and  $C_{ip} = 3$ ;  $n = 1 \dots m \dots p \dots$ ) more complex hierarchies are specified.<sup>5</sup>

Finally, a moderately complex hierarchy is shown in Figure 2c. This is a binary hierarchy with additive time changes (shown as  $k$ ). This change essentially shifts the (missing)  $\Delta T_1$  to form the lowest time level, much as in the cat example of Figure 1b. In other cases, if such shifts consistently apply to stronger accents (higher level), interesting syncopated versions of this hierarchical form may be specified. All such structures are hierarchical for two reasons: (a) The additive transform is invariant for all periods at a given level, and (b) other time levels (here  $n > 1$ ) are consistently related by  $C_i = 2$ .

All of these time structures exhibit consistent temporal nesting. In those with simple time symmetries (e.g.,  $C_i = 1, 2, 3 \dots$ ), this takes on a tight recursive form that draws on ratio time invariance. Any higher time period ( $\Delta T_n$ ) can be related to the unit period by a power of the constant ratio,  $C_i$ :

$$\Delta T_n = \Delta T_0 C_i^n \quad (\text{where } n = 0, 1, 2, \dots) \quad (1)$$

Temporal recursiveness is important. This expression indicates how it ties an event together dynamically by showing that any embedded time level can be consistently transformed into any other via time ratios.

In sum, hierarchical time structures display regular, ratio-based temporal nestings within an event's total duration, as well as consistent additive modulations to this ratio base. In terms of dynamic complexity, hierarchies fall into two general categories: those that do not incorporate additive time rules (Figure 2a and 2b) and those that do (Figure 2c). Degree of temporal coherence can be loosely gauged by the value and constancy of underlying ratio time rules, with more coherent events falling into the first category. Conversely, dynamic complexity is greater for hierarchies of the second category in which variable, larger, or noninteger ratio time transforms might occur. This

<sup>4</sup> One may argue that  $\Delta T_n$  reflects translatory time transformations forming an infinite symmetry group, and that  $C_i^n$  ( $n = 0, \pm 1 \pm 2 \pm 3$ ) reflects dilatory time transformations that also form an infinite symmetry group (e.g., Hahn & Jones, 1981; Coxeter & Greitzer, 1967). For a more general treatment in terms of time transformations that specify the logarithmic spiral, see Hahn and Jones.

<sup>5</sup> This analysis assumes an established time structure and functional markers. If the time structure refers to meters, for example, it is not concerned with issues of meter identification and ways in which certain markers, beat periods, and rhythmic figures may facilitate meter identification on the part of an attender (Essens & Povel, 1985; Longuet-Higgins & Lee, 1982, 1984). These issues involve attending and detection of temporal invariants and are considered in a later section.

Figure 2. Six examples of time structures that could hypothetically characterize a given environmental event. (Examples range from the binary time structure of (a) that is highly coherent (hierarchical) to the irregular one of low coherence in (f) that is nonhierarchical. The latter represents a temporal pattern of lights, in on [●] or off [○] positions, from a judged duration study by Poynter and Homa [1983]. In all patterns the stronger accents are signified by more dots. Temporal level,  $\Delta T_n$ , corresponds to the recurrent period at  $n = 0, 1, \dots$ . Nesting properties are given by ratio time transformations,  $C_{in} = \Delta T_n / \Delta T_{n-1}$ .)

reasoning extends to assessments of nonhierarchical time structures, three of which appear in Figure 2 (i.e., 2d, 2e, 2f).

Figure 2d illustrates an important kind of nonhierarchical time structure. It is a polyrhythm. A polyrhythm involves two simultaneously recurrent periodicities that form a noninteger time ratio (e.g., 4:3 and 4:5; Apel, 1972). It is literally a temporal phasing of two time hierarchies. In this case, binary and trinary time structures are involved, each identified by a different accent type (○ or ●) that becomes coupled (⊙) only to mark higher levels. Thus, Figure 2d has a hierarchical counterpart in Figure 2c: Both display a higher level binary time structure. But the polyrhythm lacks the consistent additive time changes of Figure 2c at its lower levels (in which time ratios are 4:3 or 8:6; Handel, 1984; Handel & Oshinsky, 1981; Yeston, 1976).

A different sort of violation of hierarchical time symmetry occurs in Figure 2e. Here, prominent accents reveal a temporal nesting that is initially hierarchical. However, this regularity of 8 beats is clearly violated in the second half of the sequence, in which higher order time spans of 10 and 6 beats occur. A more complex time ratio is thus needed to formalize this structure.

Finally, the most irregular time pattern (Figure 2f) embeds durations that are haphazardly marked at higher levels, yielding low temporal coherence. This time structure is one we hypothesize may correspond to the sequence of lights used by Poynter and Homa (1983) in a judged duration study. Stronger accents (more dots) are assumed to occur at (a) sequence beginnings, (b) light onsets (●), and (c) light onsets following pauses (no lights, ○).

In sum, nonhierarchical time structures do not display simple temporal recursivity (as in Equation 1). Dynamically, they are more complex because ratio and additive time transformations are inconsistently applied. Consequently, time relations between an event's embedded time levels and its total duration are obscured. Together with hierarchical arrangements, a continuum of temporal coherence is suggested in which coherence is loosely indexed by the ratio complexity of the rhythm generator,  $C_m$  (i.e., its invariance over levels and its integer value). Precise formalization of temporal coherence may come with progress in mathematics of dynamical systems (e.g., chaos theory). Meanwhile, the rule of thumb algorithm implied by this analysis involves the rhythm generator,  $C_m$ : As this parameter approximates a prominently outlined binary time structure, temporal coherence increases.

### *Examples of Environmental Coherence*

Often when one ventures a guess about how much time has elapsed, one is coping with time intervals filled with things such as action patterns, music, and speech. How do the formalizations just outlined advance our understanding of the structure and function of such events? They suggest that various kinds and degrees of temporal coherence will emerge and that assessment of this coherence will depend on identifications of appropriate time markers, characteristic time transformations, and multiple underlying time structures.

### *Body Gestures and Locomotion*

Body movements are classically rhythmical. Limb motions recur with fixed periods and phase locking in which markers of

these periods are often specified by distinctive directional changes in movement. Emerging research suggests that a range of underlying time structures exists, with some supporting a dynamic interplay of complex, co-occurring body gestures.

In some cases, the temporal coordination among limbs is very coherent and can be described by simple harmonic time ratios (e.g., Klapp, 1979, 1981; Kelso, Holt, Rubin, & Kugler, 1981). Furthermore, locomotion patterns in various species, including humans, reveal that episodes of walking are delimited by special beginning and ending phases and are characterized by highly regular timings of limb, torso, and head movement changes (Barclay, Cutting, & Koslowski, 1978; Carlsoo, 1972; Gray, 1968; Inman, 1966; Johansson, 1973, 1975; Pierson, 1976; Winter, 1983). Different locomotor styles appear at characteristic locomotion rates (tempi), yet each is, nonetheless, suggestive of a hierarchical time structure. For example, a cat's faster trot combines limbs in a different rhythm than its slower walk. Yet both gaits suggest the hierarchy of Figure 2c. They are differentiated primarily by distinctive additive time changes that capture corresponding velocity differences associated with on-ground time differences in the two gaits (Pierson, 1976). In short, one application of the previous formalisms suggests that categories of rhythmic style exist and can be formalized by particular combinations of ratio and additive time invariants. Both time transformations have meaning: The ratio base offers predictability for coordinative gestures, whereas additive time changes not only characterize an individual's style but they can also signal underlying velocity properties.

Of course, any systematic patterning in time lends predictability. And with motor gestures this not only affords a basis for individual motor coordination and self synchrony, it also means that visual action patterns created by one individual can support various interactive nonverbal communications with others, including turn-taking behavior, dance, nurturing, and prey-stalking, all of which partake of interactional synchrony (Condon & Sander, 1974; Kolata, 1985; Laws, 1985; Newton, Hairfield, Bloomingdale, & Cutino, 1987).

On the other hand, various complex gestures derive meaning from violations of temporal predictability (e.g., Figure 2e). For example, a dancer's skill can be signaled by singular and surprising stride changes, leaps, or turns based on one or more distinctive time intervals. Finally, complex polyrhythmic gestures routinely occur as when one does two things at once in two-handed tapping and typing (Gentner, 1987; Klapp, 1979; Jagacinski, Marshburn, Klapp, & Jones, 1988; Klapp et al., 1985). As single time sequences, such coordinative patterns are less coherent (hence nonhierarchical) according to the present definition.

### *Musical Structure*<sup>6</sup>

Animals produce song through body gestures that are coordinated within themselves and with others. Both motor productions and musical conventions give rise to compositions with distinctively marked time levels (Benjamin, 1984; Berry, 1976).

<sup>6</sup> The authors are indebted to Helen Brown (Purdue University, Department of Music) and David Butler (Ohio State Department of Music) for comments on an earlier version of this section. Any interpretative errors, however, are the authors'.

Music not only exemplifies meaningfulness in time structure, but it also illustrates how those formalisms outlined earlier have potential for expressing degrees of meaningfulness. This is because nontemporal markings can reveal different kinds and levels of temporal coherence through the dynamic play of multiple underlying time hierarchies (structural, metric).

*Nontemporal markings in music: Structural time hierarchies and harmonic markers.* A melody is a sequence of phrased pitch changes. Each phrase has a beginning and end and transpires over a given duration. In Western music, beginnings and endings often correspond to certain pitches associated with a tonal system (Dowling & Harwood, 1985; Piston, 1978). Tonality relies on special pitch relations (e.g., as in a diatonic set) and musical rules that provide listeners with a musical reference frame in which one pitch, the tonic, defines the meaning of others. Thus, although lower time levels may be indicated by tone onsets, phrases are more often outlined by the tonic and certain related pitches. Such higher level markings are shown in Figure 3 for an excerpt from a G-major melody. This excerpt is shown to end (resolve) on the tonic (G), and embedded phrases are marked either by the tonic itself or by harmonically related pitches such as the dominant (i.e., D) and the mediant (i.e., B) found in the tonic triad chord (G-B-D). Provided a key is clearly established, markers such as these add prominence to higher time levels. They are termed *harmonic accents* and denoted hereafter as *h* (H. Brown, 1988; Butler & Brown, 1984; Krumhansl, 1979; Krumhansl & Kessler, 1982).<sup>7</sup>

Melodic lines within a phrase typically point to certain final tonal goals or accents that are distributed quite lawfully in time (Cone, 1968; Boltz & Jones, 1986; Jones, 1987a; Lerdahl & Jackendoff, 1983). Furthermore, nesting of such phrases is often recursive and assumes binary and/or trinary time structures (Figure 2a or 2b). Following Lerdahl and Jackendoff (1983), we refer to these phrase-based hierarchies as *structural time hierarchies* and let *S* denote the class of structural accents that function as time markers of phrases.

Finally, if musical coherence simply involved the identification of harmonic markers and structural hierarchies, formalizations would be straightforward. But other kinds of time markers exist. These can lead to a different sort of time hierarchy, namely a metric hierarchy, or to other structural accents that must be figured into structural hierarchies, namely, temporal accents (*t*).

*Temporal relationships: Metric time hierarchies, rhythm, and tempo.* Meter provides an underlying time frame from which rhythm and tempo deviate on artful temporal journeys. This frame turns out to be a metric time hierarchy. That is, a meter involves a ratio relationship between at least two time levels (Yeston, 1976). One is a referent time level, the beat period, and the other is a higher order period based on a fixed number of beat periods, the measure (vertical bar lines denote measures). Metric levels are designated in a tune's time signature. For example, a 4/4 meter is shown in Figure 3a; the lower 4 indicates that a quarter note (♩) is the beat period, and the upper 4 indicates that there are four beats to a measure. In other words, the 4/4 meter is based on a 4:1 = 2<sup>2</sup>:1 time ratio. In fact, special, invariant ratios define all common meters: Duple and quadruple meters rest on binary time structures ( $C_n^n = 2^n$ ), and triple meters involve trinary ones (see Footnote 9). In short, these

characteristic time invariants define metric categories in Western music (Jones, 1981a, 1981c).

Markers of metric hierarchies are often associated with amplitude change (stress) at measure onsets (i.e., greater downbeat energies; Lerdahl & Jackendoff, 1983), but information other than stress will also convey meter (Benjamin, 1984; C. S. Lee, 1985; Sloboda, 1983, 1985). In fact, perceived metric accents do not always correspond to notated stress points, thus underscoring the need to identify psychologically correct time markers (Steedman, 1977; Vos & Handel, 1987; Vos, Leeuwenberg, & Collard, 1978). We denote by *M* the class of metrical accents that reliably mark metric time hierarchies (as in Figure 3). Note that metric hierarchies arise from durations of measures and thus differ from structural hierarchies that arise from durations of melodic phrases.

*Tempo* offers a yardstick of meter and of a measure's duration. It depends on the unit time period selected by a performer, namely, the beat period. This establishes a piece's characteristic pace<sup>8</sup> (e.g., largo, allegro). For example, all other things being equal, a beat period of an eighth note (♩) results in a faster pace (tempo) than one of a quarter note (♩).

*Rhythm* refers to patterned time changes within a metric frame. An important feature of rhythmic patterns is that they provide temporal accents in the form of unusual, attention-getting time changes, namely, longer or shorter durations (multiples or subdivisions of the beat period). Thus, the longer half notes (♩) in the tune of Figure 3 tend to be more salient. Such temporal accents form powerful time markers that often function as structural accents that, like harmonic ones, mark beginnings or endings of melodic phrases and so reinforce structural hierarchies (Cone, 1968; Kramer, 1982; Todd, 1985).

In sum, both structural and metrically marked hierarchies testify to the fact that music captures time symmetries formalized earlier by different time structures. It might be argued that this isn't so because notated time ratios don't always appear in produced music (Gabrielsson, 1985; Sternberg et al., 1978).<sup>9</sup>

<sup>7</sup> This analysis of tonality is oversimplified. The psychological salience of pitches in any musical context depends not only on prevailing tonality and grammatical considerations (H. Brown, 1988; Krumhansl, 1979; Krumhansl & Kessler, 1982) but on temporal context provided by rhythm (Monahan & Carterette, 1985) and temporal order (e.g., various notes identifying intervals such as tritone; H. Brown & Butler, 1981; Butler & Brown, 1984; Kidd, 1984). Finally, temporal ordering of harmonically significant pitches, as in Figure 3, can establish, via implied harmony, a sense of certain chordal progressions that move to resolution on the final note as a tonic triad. Thus, shifting accenting may also affect coherence in tonal melodies by creating conflicting tonalities and inappropriate implied harmonic progressions.

<sup>8</sup> For additional definitions of tempo, see Kronman and Sundberg (1987) and Clynes (1978). In particular, Clynes has proposed an organismic tempi that uniquely specifies a composer or performer's pulse. This idea, although compatible with our general approach, should not be confused with more traditional definitions of tempi (e.g., Kronman & Sundberg). In produced sound patterns, it is quite possible that characteristic tempi follow Clynes' exposition, a notion that reinforces our approach, rather than contradicting it.

<sup>9</sup> Sternberg, Knoll, & Zukofsky (1978) used a psychophysical task to study the perception and production of time ratios and found systematic distortions of responses to time ratios. However, their analyses of production times, which relied on a reaction-time correction factor,

But, the picture is more complex. There is evidence for preferred, simple time ratios in produced rhythms (Essens, 1986; Fraisse, 1956, 1964, 1982; Povel, 1981; Essens & Povel, 1985). Furthermore, skilled musicians often accurately produce higher order time spans and complex rhythmic ratios<sup>10</sup> (Clarke, 1985a, 1985b; Shaffer, 1981, 1982). We claim that many deviations that do occur are most reasonably represented as additive time changes that arise either from a performer's (a) failure to capture difficult ratio notations and/or (b) conscious attempts to communicate rhythmic style (e.g., legato, staccato) or emotion via tempo change (Clarke, 1984; Gabriellson, 1974; Gabriellson, Bengtsson, & Gabriellson, 1983; Shaffer, Clarke, & Todd, 1985). These factors relate to ways a performer communicates by serially modulating melodic timing and are discussed more fully elsewhere (Jones, 1987a). A final reason for blurring of temporal proportionalities is found in the score itself. This factor is considered next.

*Temporal coherence in music: Structural and metric hierarchies.* A musical score displays temporal coherence by the way it combines structural and metric time structures. Separately, each of these structures can form a coherent hierarchy. But together, overall coherence depends on how they combine in time.

Figure 3 shows how differences in temporal coherence can come from variations in musical phrasing. Here, all timing changes concern relative phrase durations of a structural hierarchy relative to fairly constant properties of a metrical hierarchy. Objectively, most coherence appears in Figure 3a in which both (binary) hierarchies coincide in time. Each measure opens with a metrical (stress) accent, and although the 8-beat melodic phrases are longer than measures, each resolves (*h* coupled with *t* accents) predictably at the end of alternate measures. This yields a single prominently marked time structure as in Figure 2a. By contrast, against the same metric frame, melodic phrasing (signaled again by joint *h* and *t* structural accents) forms a less coherent time pattern (Figure 3b). This piece opens with regular and strongly marked melodic phrasings, but it closes with a musical surprise created by deviant phrase lengths of 6 and 10 beats, respectively. In this, it shares the format of Figure 2e. The psychological impact of such deviations is considered in later sections. Finally, even more temporal uncertainty is found in Figure 3c, in which a subtle modulation of the structural hierarchy of Figure 3a occurs. In this case, the *t* accents remain hierarchical. In fact, they reinforce the metric time hierarchy. Notice that only harmonic accents are shifted to vary phrase durations. They alone create a nonhierarchical form.

In sum, musical coherence varies in meaningful ways that lend themselves to certain formalisms. We should underscore that we do not claim that the more coherent musical events are the more interesting or appealing ones; often they are not. Unquestionably, a certain level of artful and even systematic violation (e.g., syncopation) of very coherent structures is necessary to create aesthetic appeal. We merely claim that these variations speak to temporal predictabilities. In this regard, coherence decreases and dynamic complexity increases whenever simple time hierarchies are violated by additive time changes and/or by other temporally phased hierarchies.

failed to take account of anticipatory behaviors, and thus are not conclusive.

## Speech

Among other human songs intended to serve communication goals are those created by dynamically coordinated vocal gestures, namely speech patterns. Speech, like music, presents apparent rhythmicities that, on closer inspection, fail to fit the simplest time hierarchies. Again, the key to unlocking these timing mysteries may turn on the identification of appropriate time markers as well as clarification of different time structures and their interrelations.

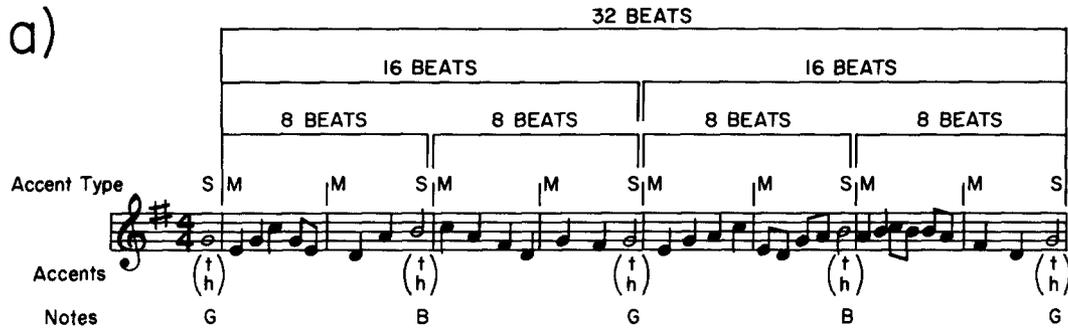
English is a stress-timed language with appropriate time intervals thus defined between stress points. The simplest rhythmic pattern proposed for the distribution of stress points is isochrony, in which accents within unbroken utterances should be temporally equidistant (Pike, 1945). However, only limited empirical support exists for the isochrony principle even in a sophisticated version based on binary time hierarchies (e.g., Lehiste, 1976; Martin, 1972; Shields, McHugh, & Martin, 1974), thereby raising questions about its sufficiency (e.g., Liberman & Prince, 1977).

Some investigators argue that deviations from isochrony function meaningfully in communication (Lehiste, 1977). For example, spoken phrases are often marked by end lengthening and proportional silences as in music (Paccia-Cooper, & Cooper, 1981). Other meaningful deviations, stemming from shifts in speaker articulation rate, marker strength, and intonational contour can communicate emotional states (Scherer, 1979), intentions (Lehiste, 1976; Pickett, 1980), and turn-taking goals of speakers (Siegman, 1978; Wiemann & Knapp, 1975). In many cases, these deviations are additive time changes that derive contextual meaning by violating the predictability conferred by underlying time symmetries.

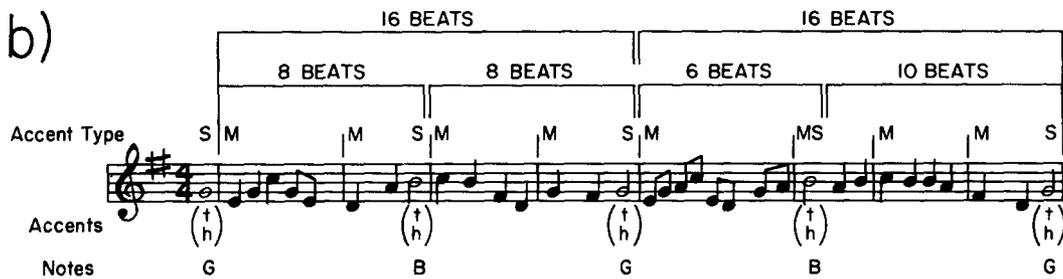
Insight into underlying time symmetries is suggested by Cutler and Isard (1980). Using careful measurements of interstress intervals, acoustical manifestations of anisochrony were found to reveal an underlying invariant rhythm. Interstress intervals within a syntactic boundary are isochronous, maintaining a ratio of 1.00; however, the relative duration of adjacent intervals separated by a syntactic boundary is consistently lengthened to a ratio of 2:1. Cutler and Isard argued: "Increased duration results not from a complete disruption of isochrony, but rather from 'skipping a beat' so that the duration of longer intervals is twice that of shorter ones" (p. 258). Other evidence supports certain ratio timing properties in speech (e.g., Kozhenikov & Christovich, 1965; Tuller, Kelso, & Harris, 1983).

Thus, it appears that both temporal regularity and systematic deviations from regularity do exist in speech patterns. Nevertheless, to align regularity simply with binary time structures based on stress markers is too restrictive for several reasons. First, even with stress-timed intervals, other time structures such as trinary and more complex hierarchical forms may be

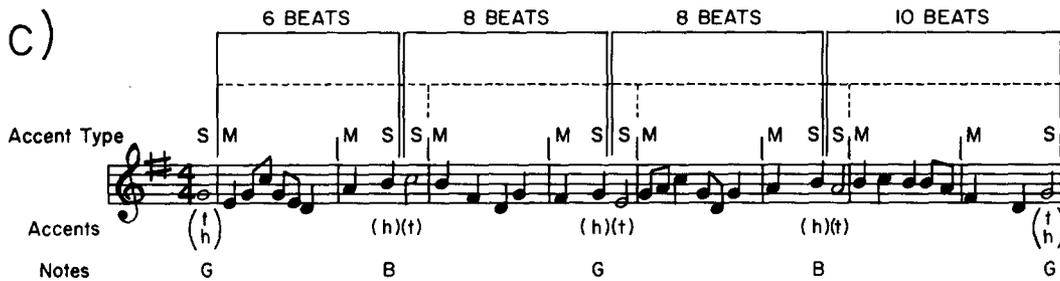
<sup>10</sup> Complex metric time ratios refer to  $C_m = \Delta T_n / \Delta T_{n-1}$ , values that form nonintegers such as 6/4 and 9/8 (or compound meters, e.g., simple meters such as 2/2 6/4 multiplied by 3). Also, ratios that differ with respect to level—as in 3/4 meter, where  $C_1 = 3$  at one level and  $C_2 = 2$  at others—are only a slightly more complex time hierarchy. Metric ratios form either simple or complex invariants in time and thus provide the basis for specifying rhythmic patterns, as Jones (1981a, 1987a) has shown.



Hierarchical Time Structure  
Accents Coupled (t h)



Non-hierarchical Time Structure  
Accents Coupled (t h)



Non-hierarchical Time Structure-h  
Accents Decoupled (h) (t)

Figure 3. Variations in temporal coherence that arise from differences in relative durations of melodic phrases. (Properties of a structural time hierarchy are changed, whereas those of the metric hierarchy are not. Most coherence occurs in (a) in which S and M hierarchies coincide because coupled S accents [t plus h] end each phrase and these also coincide with alternate measure endings. Less coherence occurs in (b) in which S accents [t plus h] are again coupled, but they create a nonhierarchical event. In (c), t and h accents are decoupled via an additive time separation; the t accents remain hierarchical [dotted lines], whereas h accents mark nonhierarchical time levels [solid lines].)

operative. Second, significant time markers other than amplitude changes (stress) are also critical in speech timing. Pitch and time accents function to outline recursively related melodic phrases and intonational patterns in speech much as they do in music (e.g., Bolinger, 1958; Fry, 1958; Ladd, 1986). Finally, just

as with metric and structural time hierarchies in music, two or more different time structures underlying speech rhythms can combine to support complex temporal accent phasings (e.g., Figure 2d). Fowler's (1983) suggestion that vowels and consonants are produced by separate but temporally phased articula-

tory streams implies this sort of nonhierarchical (polyrhythmic) structure.

In sum, the search for a single periodicity in speech patterns is unlikely to be rewarded. Instead, the characteristic temporal patternings in speech must be addressed by identifying (a) functional markers, (b) meaningful deviations from characteristic time hierarchies, and (c) temporal phase relations among various kinds of speech time hierarchies.

Finally, a theme of Part 2 is that although temporal coherence in gestures, music, and speech varies, it does so meaningfully by playing against certain predictable time bases. And because we will claim that coherence is critical to dynamic attending, some formalization is important. In the long run, a precise formalization will depend on identification of time markers and appropriate time structures. Here, we suggest that a provisional index of temporal coherence is to be found in the  $C_m$  parameter. This ratio time transformation changes in value systematically as a function of the kind of time structure(s) involved and the number of additive time changes present.

### *Attending to Dynamic Environmental Structure*

If environmental events are structured, more or less coherently in time, what does this imply for an attender? In this section we consider this question with the goal of proposing distinctions between ways people attend over time to temporal events.

We begin by assuming that attending is itself a temporal activity that serves certain goals and is guided by the way nontemporal information unfolds in time. We also assume that dynamic interactions between an attender and some environmental event occur in the form of *attunements*. One important distinction to emerge is that between different kinds of attunements, specifically between *future-oriented* and *analytic* attending modes. Finally, constraints associated with dynamic complexity are shown to affect the likelihood of these kinds of attunements and to enter into attunement failures.

### *Attunements*

The idea of attunement rests on an assumption that attending is sensitive to environmental rhythmicities (Jones, 1976). Attending to speech, body gestures, music, and many other world events is a rhythmical activity. That is, rhythmicity is inherent to living things including the activity of attending. Biologically based rhythms that are responsive to various environmental time periods underlie attending. These internal rhythms have independently variable amplitudes that provide potential ways of measuring the physical basis of attending.<sup>11</sup> Together, frequencies and amplitudes within sets of recursively related attending rhythms underlie focusing of attentional energy over selected environmental time spans, thus permitting temporally selective attending for particular goals.

Interactive attending is termed *attunement*. It involves a synchronous interplay between an attender and an event in which the former comes to partially share the event's rhythmic pattern. This involves an enlistment of organismic rhythms whose frequencies and amplitudes can mirror those of an event, thus exemplifying a dynamic type of psychophysical complementarity (Shepard, 1981). One or more attending rhythms become

phase-locked to corresponding time spans marked within the event. Attunement is most apt to occur with temporally coherent events and, in fact, we assume that organisms operate with an initial bias for high temporal coherence, as if an ideal hierarchy obtained (Essens, 1986; Fraisse, 1964; Povel, 1981; see Footnote 11). For example, because the hierarchical structure of the melody in Figure 3a has dynamic simplicity, it offers ready attunement to one or more of its embedded time levels as well as to its total time span. To illustrate these ideas, we consider the functions of referent level, focal attending, and attentional shifts.

### *Referent Level*

We assume that an attender initially entrains to some relatively regular, prominently marked time period within an event. Perhaps in an utterance, it is a person's characteristic articulation rate; in music it may be the beat period or measure span. In any case, this period functions as an anchor or referent time level for the perceiver. In the event itself, it is denoted  $\Delta T_j$ , but within the organism it denoted by  $\Delta T'_j$ , where the prime indicates an entrained rhythm and  $j$  (i.e., when  $n = j$ ) refers to a referent level. This referent period provides (a) a temporal reference frame, "locking" the attender to the event, and (b) an event-determined time scale that "calibrates" time spans. In Figure 3a, for example, the quarter note,  $\downarrow$ , is the beat period; it determines  $\Delta T'_j$  as a temporal anchor as well as specifying the length of other hierarchical time levels.

A referent period is jointly determined by context and an attender's biological make-up. It depends on environmental factors, (e.g., presentation modality, initial patterning, tempo) as well as organismic ones (e.g., species, age, arousal level; Hirsh, Bilger, & Deatherage, 1969; Kolers & Brewster, 1985; Povel & Essens, 1985). For example, this construct formalizes the idea that modal attending rhythms shift with maturation (Jones, 1976). Infants may be differentially selective to events that offer relatively small referent periods (e.g., higher tone frequencies, phonemic patterns), whereas older people may increasingly entrain with longer periods as referents. It follows that temporal scaling (calibrating) will shift with age; this has direct implications for time judgments.

### *Focal Attending*

Attunement shifts involve directing some attending energies away from the referent period to another level (Jones, 1981a). This is *focal attending*, and attended-to levels are referred to as *focal time periods* (or levels). Focal attending realizes temporally selective attending in which energy can be targeted over

<sup>11</sup> Biological energies of attending rhythms are physical, not psychological. They are directly tied to the time structure of the organism and, through attunements, to specific relative time properties of events. Furthermore, the term *recursively related sets* of attending rhythm refers to the assumption that those periods that are multiples (divisors) of simple invariant integer time ratios form a more coherently related set. In these respects, the concept of energy differs in an important way from that of, for example, attentional resource theory, in which it assumes a more metaphorical than structural-specific role (e.g., Navon & Gopher, 1979).

some time levels and not others to serve certain goals (Jones, Boltz, & Kidd, 1982; Jones, Kidd, & Wetzel, 1981; Neisser & Brecklen, 1975). For example, if a listener's goal is to "catch" the global idea of a speaker, attending will be targeted over higher focal time periods in an utterance than if the listener must concentrate on the speaker's vowel pronunciation. Because focal attending involves energy shifts, it is more effortful with dynamically complex events and with attunements to proportionately smaller time periods (higher frequencies).

How does shifting occur? It relies upon the attender's use of time transformations. Formally, let a primed interval,  $\Delta T'_n$ , reflect the attending rhythm that becomes attuned to level  $\Delta T_n$  in the event. Focal attending reflects the phase-locking of  $T_n$  with  $\Delta T'_n$ , in which  $\Delta T_n = \Delta T'_n$  for a subset of levels. The phrase *using time* refers to a reliance on an extracted time rule (e.g.,  $C'_i$ ) to shift attending energy from one level to another. Any attending level,  $\Delta T'_n$ , may therefore be a focal level if it gains energy. A focal level is related to a referent period by a time rule,  $(C'_i)^{n-j}$ . This is evident in the following recursive formula:

$$\Delta T'_n = \Delta T'_j (C'_i)^{n-j} \quad (\text{where } |n - j| \leq 5). \quad (2)$$

To the extent that  $C'_i$  and  $T'_j$  are good estimates of  $C_i$  and  $\Delta T_j$  in the event, then  $\Delta T'_n = \Delta T_n$ .

People can rarely verbalize use of a rhythm generator ( $C'_i$ ), nor can they report focal time levels ( $\Delta T'_n$ ; Jones, 1981a). Focal attending is a tacit, how-to skill that is implicitly acquired (Reber, Kassir, Lewis, & Cantor, 1980).

### Attentional Shifts

An important aspect of skilled attending and expertise is flexibility (e.g., Gentner, 1987; Keele & Hawkins, 1982). Flexible attending here refers to shifts of attending energies from the referent level to many other event levels. For example, in the event of Figure 3a, a skilled attender can use ratio time invariants to flexibly target multiple time levels, moving effortlessly from one to another to serve different needs. Thus, if a person is attending over melodic phrases (e.g., 4 and 8 beats) but must respond to some higher level, that person may easily transfer more attending energy over this level (e.g., 16 beats) or even the entire event span by relying on Equation 2. In short, attunement shifts allow one to selectively "examine" strings of notes marking a given hierarchical level.

One of us has suggested that flexible attending offers a practical solution to the problem of structural ambiguity and the fact that a given event means different things in different contexts (Jones, 1986; see also Cutting, 1982). Depending on the task, a person can rely on high or low focal levels to attune to event relationships most relevant to some function. In fact, we distinguish two general kinds of attunements. *Future-oriented* attending involves a global focal attending over time periods higher than the referent level ( $n > j$ ), whereas *analytic attending* directs attending energies to relatively low levels in a temporal hierarchy ( $n < j$ ). Thus, future-oriented attending may support one's ability to follow conversational turn-taking between two speakers, and analytic attending would underlie identifying the dialect of one of the speakers. These focal attendings rest on different temporal perspectives of the same event.

These attending modes involve different regions in a continuum of nested time levels, but with the following two qualifica-

tions: (a) that they yield different temporal perspectives on the same event relative to a given referent level associated with that event (higher for future-oriented, lower for analytic), and (b) that they may support selectively different sorts of activities, depending on task goals. That is, future-oriented attending is particularly functional when longer term preparatory attending and real time extrapolations of an unfolding event are involved. A future-oriented extrapolation is termed an *expectancy*. Analytic attending is more functional when relatively short-term activities are involved. Because relevant focal periods are usually less than that of the referent period, expectancies as such are not supported, although attending to local detail and grouping with respect to the referent span is.

Figure 4 illustrates that both attending modes are possible in a coherent event. They are shown here with different referent levels (for comparison purposes, both are standardized with respect to the same beat period). In Figure 4a, future-oriented attending involves shifts *up* to several potential focal levels higher than the referent, which here is a two beat period ( $n = j = 1$ ). The  $\Delta T_3$  focal level receives most attending energy ( $(C'_i)^{n-j} = 2^2$ ), so that future tonal onsets separated by 8 beats are readily anticipated, although final endings are also predictable (via  $C'_i = 2^4$ ). Future-oriented attending occurs here because  $C_i = 2$  may be extracted and used recursively as  $C'_i$  (where  $C'_i = C_i$ ).

In Figure 4b, the referent level becomes four beats ( $n = j = 2$ ) in order to illustrate analytic attending. With this attending mode, energy can be shifted *down* to the level of adjacent items or pulses ( $\Delta T_0$ ) from  $\Delta T_2$ . Thus, when someone attends analytically, attending is focused downward to cover fine changes that transpire over relatively small time periods (including the referent period). Such an activity might correspond, for instance, to tracking or even to counting these elementary items in groups of four. It can be effortful and will be more common in early stages of learning.

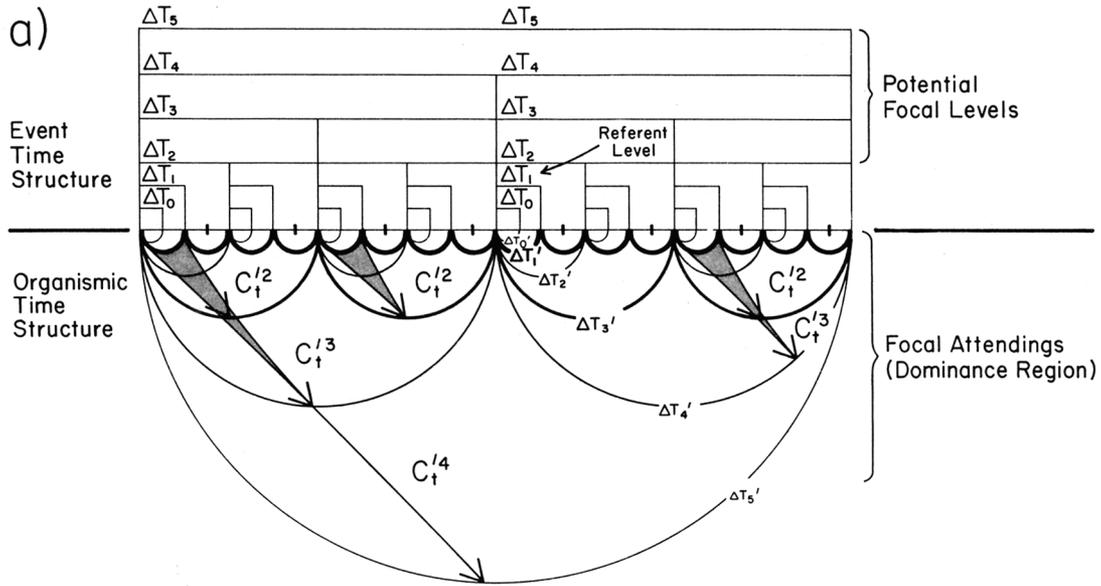
### Constraints on Attunements and Attunement Shifts

A person's goals determine time levels most relevant for attunement. But the catch is that although attunements to an event can shift to achieve different goals, the event's dynamic complexity will constrain the ease with which this is accomplished. In particular, attunement shifts are less likely in events of low temporal coherence. This idea is formalized as a constraint on attunement conferred by the event's ratio time properties (ratio complexity constraint). Other constraints on attunement exist but these concern the attender's inherent ability to shift attending (dominance region constraints).

### Ratio Complexity Constraints

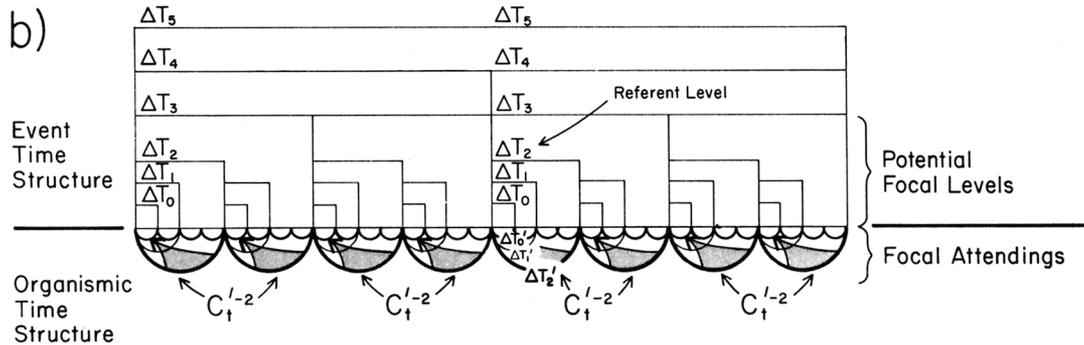
Attunements have been described with reference to highly coherent events, ones that confirm a bias for simple time ratios. The complement of this bias represents ratio complexity constraints. With less coherent events, attending becomes less flexible (at least initially) because the ratio,  $C_n$ , is complex. Attending is confined to one or two temporal levels in which biases for simple and constant ratios may be realized. In short, analytic attending is more likely with many nonhierarchical events.

To illustrate, consider the ways of listening encouraged by



Future-Oriented Attending

Focal Attending over time levels higher than Referent level ( $n > j$ ). Referent level,  $\Delta T_j$ , is  $\Delta T_1$  (2 beats). Following equation 2,  $C_t^{n-j}$  has exponents of 2, 3, 4 for potential focal levels  $\Delta T_3'$ ,  $\Delta T_4'$ ,  $\Delta T_5'$  respectively (within the Dominance Region). The most prominent focal level is  $\Delta T_3'$  of 8 beats.



Analytic Attending

Focal attending over time levels lower than (or equal to) Referent level ( $n \leq j$ ). Referent level,  $\Delta T_j'$ , here corresponds to the  $\Delta T_2$  (4 beats) level to illustrate shift of energy via  $C_t^{j-2}$  to  $\Delta T_0'$  of 1 beat.

Figure 4. Future-oriented (a) and analytic attending (b) as shifts of attending energy from a given reference level ( $\Delta T_1$ , or  $\Delta T_2$ , respectively) to higher (future-oriented) or lower (analytic) time levels. (Theoretically, dominance region constraints are a factor in both, but they are shown here only for future-oriented attending as decreasing energy [thinner lines] associated with more remote time levels.)

melodies of Figures 2 and 3. Although future-oriented attending is more apt to occur with those melodies shown in Figures 2a and 3a, where  $C_i = C_i' = 2$  facilitates anticipations of melodic phrasing, analytic attending is more likely with those of Figures 2d, 2f, or 3c, in which a constant  $C_i$  is difficult to specify. In the latter, analytic attending involves focusing on a succession of individual items because these nonhierarchical patterns involve highly complex and variable time ratios. In other words, when the big picture in time is not clear, people focus on the small picture.

Finally, ratio constraints summarize people's response to the dynamic complexity of environmental events. Attenders have biases for simple time symmetries, and this affects the ease with which they can achieve certain goals that depend, respectively, on future-oriented or analytic attending.

Dominance Region Constraints

A dominance region reflects natural attending limits that are independent of ratio-complexity constraints (see Equation 2).

Even with simple hierarchical time patterns, attenders can only shift attending energies so far before the connection between an established referent and a remote focal level breaks.

Specifically, the dominance region involves a set of related time levels around the referent beat period. It reflects the distribution of predominant attending energies over hierarchically neighboring time levels (a related concept is the serial integration region; Jones, 1976).<sup>12</sup> For example, in musical compositions in which many related time levels appear, the dominance region suggests that attending shifts from a referent to remote levels are unlikely; listeners cannot accurately anticipate the length of a symphony from its tempo. (A similar notion is *tactus*, a perceptually dominant subset of time levels surrounding the musical beat; Lerdahl & Jackendoff, 1983). Dominance region time levels ( $\Delta T_0$  to  $\Delta T_3$ ) are suggested by potential focal levels shown in Figure 4a. The key notion is that attunement shifts exceeding  $|n - j|$  of 4 or 5 become increasingly difficult.

Dominance region limits are relative not absolute time limits in that they are expressed in terms of hierarchical levels (exponents of  $C'_i$ ). Attunements are relative to an event's characteristic referent period, meaning that one can potentially attend over a limited set of related time levels, regardless of their absolute time values. But note that if attenders are strongly biased in entraining to a given referent period, then these relative constraints will seem like absolute ones. Concretely, this means that perceived coherence can be affected by a pattern's rate or duration.

### *Failed Attunements*

Given the preceding discussion of attunement constraints, it follows that attunements can fail. A failed attunement is an asynchrony between some attending rhythm ( $\Delta T'_n$ ) and the relevant focal level of an event ( $\Delta T_n$ ). The nature and consequences of such asynchronies will differ as a function of event complexity. That is, in less complex events there will be fewer attunement failures and a given temporal asynchrony can stand out as a temporal contrast based on future-oriented attending. In highly complex events, many asynchronies appear and these force attunement refinements based on analytic attending.

### *Temporal Contrast and Future-Oriented Attending*

Imagine anticipating the shape of an unfolding event and finding to your surprise that there's a discrepancy. This is what occurs whenever a nonhierarchical event begins deceptively as a hierarchical one. The melody of Figure 3b does this. Initially, it supports future-oriented attending, but as we saw earlier, it carries the potential for a musical surprise because finally it deviates from its opening regularity. Formally, this yields an expectancy violation: One expects a certain future time course (i.e.,  $C'_i = 2$ ), which does not occur. The difference between observed and expected focal periods at some level,  $n$ , represents an asynchrony that is experienced as a temporal contrast:  $\Delta T_n - \Delta T'_n$ . We assume that temporal contrast determines one's response (e.g., surprise) to systematic additive changes (asymmetries) that play against time symmetries. In short, it is contrasts of various sorts that function to communicate rhythmic styles, emotion, and time change. And something akin to temporal contrast may function in time discrimination (Kris-

tofferson, 1976).<sup>13</sup> Here, we propose that certain contrasts can influence judgments about events when the goal is time estimation.

### *Refining Attunements and Analytic Attending*

Complex events contain many irregularities, and so they inevitably invite many failed attunements. In the face of this, redoubtable attenders may turn to refining attending skills depending on the nature of the task. Indeed, if the event is based on a complex time hierarchy, this implicit learning can lead to the discovery of initially obscure temporal invariants. However, with nonhierarchical events in which no invariant time rule exists, such skill acquisition is less likely. For instance, the pattern of Figure 2f affords little opportunity for flexible attending, either initially or with practice. Here, ratio constraints on attending mean that attunements will be confined to one or two lower time levels where they may support various goal directed activities.

### *Summary of Part 2*

Dynamic attending reflects the attender's tacit use of an event's dynamic structure. It relies on the particular nature of the event's temporal invariants (e.g., in gestures, music, speech), which if present, determine its coherence or conversely its dynamic complexity. Because environmental events differ in coherence, they offer the rhythmical attender different ways to selectively synchronize with one or more nested time level(s) (focal attending) to achieve specific goals. This participatory attending is termed attunement and is temporally selective. Attunement involves entrainment to a referent time period and selective focal attending to some time level. Attunements are flexible. They can shift over focal levels to reflect either future oriented or analytic attending. However, this flexibility is limited by ratio complexity constraints related to rhythmic invariants and by dominance region constraints related to the span of an attunement shift. Consequently, future-oriented attending that permits expectancies about event endings is more likely with coherent events, whereas analytic attending that encour-

<sup>12</sup> The serial integration region addressed the way in which temporal order is preserved in world events through relationships between properties of markers at different time levels. For example, if at level  $n$ , the pitch change over that level—namely,  $\Delta P_n$ —is disproportionately great for the corresponding time change—namely,  $\Delta T_n$ —then the pattern will have low temporal coherence. A consequence is the breakdown of synchronous attending and loss of one's ability to order things in time. This is evident in phenomena such as auditory pattern streaming (Jones, 1976). The serial integration region, thus, is bounded by extreme pitch/time ( $\Delta P_n/\Delta T_n$ ) ratios. It is contained in the dominance region, which is bounded by extreme time/time ( $\Delta T_n/\Delta T_j$  where  $|n - j| = 4$  or  $5$ ) ratios. In this article we focus only on constraints affecting time levels, that is, dominance region constraints.

<sup>13</sup> Temporal contrast and lawful patterns of temporal contrasts serve special communicative functions in music. Jones (1987a) has suggested that these additive time changes arise from a horizontal component introduced into a performer's production patterns, a component that corresponds to velocity patterns of gestures. In other contexts, singular contrasts defined earlier, have been postulated to affect time discrimination (Kristofferson, 1977).

ages focusing over lower time levels is more likely with less coherent events. Finally, attunement failures can lead to temporal contrast, namely the deviation of expected from observed focal levels.

### Part 3: Dynamic Attending and Time Estimations

The conceptual framework of Part 2 has implications for attending in various contexts. Part 3 develops these with respect to time estimation. Solutions posed with regard to the filled interval effect (Part 1) are seen as consequences of constraints on a general attending activity imposed by these tasks. One goal of Part 3 is to model time estimation behaviors in comparative judgment and production tasks using general assumptions of dynamic attending.

Typically, a time estimation task requires assessment of one or several time intervals. In light of Part 2, this creates a puzzle: If people attend rhythmically, by relying tacitly on ratio time transforms, then how do they come to explicitly judge time intervals as such? Society and our own schedules require that we judge time as intervals, and we often comply with more or less accuracy. How is this puzzle resolved?

One resolution plays on the idea that interval time estimates encourage selective attending over higher order time spans that relate beginnings and endings of judged events. Clearly, this attending has a basis in relative time. Yet, the interval time estimate is not necessarily an easily achieved goal. The perspective of Part 2 suggests three assumptions about the way rhythmical attenders approach such tasks: (a) They use relative time to achieve task goals. (b) Ease of goal achievement depends on event structure, notably on the way dynamic complexity constrains flexible attending (i.e., future-oriented and analytic). (c) Time estimates themselves depend on tacit use of resultant focal levels and because awareness of these levels is poor, time estimates are liable to be biased by attunement failures (temporal contrast and attunement refinements).

According to the first assumption, interval time requests encourage attenders to target attending over one or more higher time levels (future-oriented attending). In the second assumption, their success depends on event complexity. Future-oriented attending is easier to achieve in comparative time judgments about simpler time structures. Here, expectancies about ending times of standard and comparison events are fairly effortless. Finally, according to the third assumption, if an attunement failure occurs, then it creates a temporal contrast, and this, in turn, is apt to tacitly bias time judgments especially if ending times are involved. That is, people respond to the presence or absence of expectancy violations, not to absolute values of expected intervals. This scenario predicts that people can correctly judge intervals whenever a comparison event truly ends before (or after) some expected ending time (given by the standard). By the same token, it predicts that people will err in judging truly identical intervals if one (or both) event(s) happens to violate expectancies about event endings (a filled interval effect). This occurs whenever expectancies based on *embedded*, not total, time periods lead to contrasts that implicate an event's ending. If these embeddings make an event's ending *seem* too early or too late, then time judgments can be tacitly distorted. Contrasts do not occur with expectancy confirmations nor when either the expected or observed time period is

absent. But when they do occur and impinge on ending times, then biased time estimates can result. In sum, one way rhythmical attenders cope with time intervals depends on future-oriented attending and temporal contrast.

If attenders don't succeed in attending over higher time levels, then analytic attending occurs. In this case, although interval time estimates remain influenced by the way information fills to-be-judged time intervals, biasing can take a different form. Particularly if events are incoherent, whenever task goals bring into play memory-oriented strategies, attending will be focused over lower time levels. People may engage in counting, grouping, associative labeling, and so on, in order to impose some structure on the event, and all of these activities direct attending to relationships among adjacent items. Consequently, if time estimates are required they will be biased by these activities and specifically by differences in focal levels used to attend to the events involved. For example, if two complex tunes of equal duration are filled with different numbers of tones, then the one with more tones will usually involve a denser temporal spacing. If people count tones, then attunements to the denser sequence will be based on lower focal attending. However, awareness of focal time levels is poor. Thus, whenever people tacitly (and erroneously) infer that focal levels used to attend to each tune are the same and that the one with more tones is longer, then judgment errors will appear (a filled interval effect). Again, such mistakes arise from the attenders' lack of awareness of differences in focal levels actually used. In sum, a second way in which rhythmical attenders cope with intervals of time involves focal time level differences based upon analytic attending.

In short, Part 2 suggests a broad dichotomy for the way rhythmical attenders interact with temporally coherent and noncoherent events, and ultimately this explains how they may arrive at interval time estimates. Independent evidence for this distinction comes from Povel and Essens (1985), who found that reproductions of hierarchical auditory sequences involve a reliance on higher order time relations, whereas those of nonhierarchical sequences involve the use of mnemonic strategies such as counting (see also Bamberger, cited in Povel & Essens, 1985; Hamilton & Hockey, 1974; Jackson, 1985; Maser, 1983). This distinction offers a way of explaining the notorious variability found in time estimation studies. It suggests that duration estimates of the same event will vary with attending mode.

Finally, although this framework relies on new constructs such as future-oriented and analytic attending, the analytic attending mode does incorporate notions underlying contemporary time models. It differs in linking the mnemonic activities suggested by current time models to a rhythmic base and to differences in focal time levels. Thus, it suggests that although attending to events filled with more information may involve more effort (Underwood & Swain, 1973) or complexity (Ornstein, 1969) or changes (Block, 1978), finally the decisive factors involve temporal levels of attending.

### *Time Estimation in Judgment Tasks: A Contrast Model*

A contrast model formalizes the proposal that time judgments are biased by the style of attunement failure (temporal contrast or mnemonic refinements) associated with a given attending mode (future-oriented or analytic). Two parameters of this model reflect contrast and mnemonic activities, respec-

tively. A third parameter, a coherence weight, determines the likelihood of either attending mode. It reflects ratio complexity constraints (Part 2) and so summarizes the impact of dynamic complexity on attending.

In comparative duration tasks, the temporal contrast biasing time estimates is a pairwise one based on expectancy violations in both events; for pair  $i$  it is denoted by the parameter,  $DE_i$ . The mnemonic difference parameter is based on differences in analytic attending levels over both events; it is denoted  $DM_i$ . Here, we apply the model to filled time intervals that are equal in duration but extensions to unequal intervals are straightforward and given in the Appendix. Thus, for any pairwise judgment,  $i$ , we assume that two kinds of random variables exist that are distributed, respectively, as  $N(DE_i, \sigma_{Ei}^2)$  and  $N(DM_i, \sigma_{Mi}^2)$ , in which  $\sigma_{Ei}^2 = \sigma_{Mi}^2 = \sigma_i^2$ .

Formally, the Contrast Model differentially weights temporal contrast,  $DE_i$ , and mnemonic differences,  $DM_i$ , by the probability of future-oriented,  $W_i$ , and analytic attending,  $1 - W_i$  ( $0 \leq W_i \leq 1.00$ ). This logic leads to a binary mixture modeling of the two attending states (Townsend & Ashby, 1983). Let  $D_i$  be the expected value of a theoretical distribution of duration estimates that results from a binary weighting of the two states:

$$D_i = W_i DE_i + (1 - W_i) DM_i. \quad (3)$$

The predicted average of scaled duration judgments,  $JD_i$ , is a monotonic function,  $f$ , of  $D_i$ . For simplicity,  $f$  is assumed to be linear with scale parameters  $a$  and  $b$ . The intercept parameter,  $b$ , reflects response scale category; if  $b$  is the score for a null difference judgment, then average judged time magnitudes in a rating task are as follows:

$$JD_i = f[D_i], \quad (4)$$

and

$$JD_i = a[W_i DE_i + (1 - W_i) DM_i] + b. \quad (5)$$

Judgment bias is given by the point of subjective equality,  $PSE_i$ , and is gauged for the case in which two intervals have identical total durations. From Equation 5, the contrast model predicts  $PSE_i$  as follows:

$$PSE_i = JD_i - b = a[W_i DE_i + (1 - W_i) DM_i]. \quad (6)$$

The predicted variance of observed duration judgments is derived from assumptions regarding moments of the weighted distributions. This is  $SD_i^2$  for judgments of event(s)  $i$ :

$$SD_i^2 = a^2[W_i(1 - W_i)(DE_i - DM_i)^2 + \sigma_i^2]. \quad (7)$$

Equations 5–7 reflect predictions of the contrast model. They show that mean judgments, bias, and variance are a function of the weighting parameter,  $W_i$ , and the two attending mode parameters,  $DE_i$ , and  $DM_i$ .

### Parameters and Their Estimation

The  $W_i$  weight is sensitive to environmental structure, whereas the  $DE_i$  and  $DM_i$  attending parameters reflect the respective impacts of attending mode on time estimates. Parameter determinations rest on the attunement constraints described in Part 2. In particular, following the dominance region constraint, we assume that embedded periods play an important

role in guiding attending over pairs of events in comparative judgments. And with respect to ratio complexity, we assume that  $C_m$  determines the value of the coherence parameter,  $W_i$ , over both events. We now illustrate parameter determination using the hierarchical and nonhierarchical musical examples of Figure 3 to fill standard and comparison time intervals.

### The Weighting Parameter, $W_i$

The  $W_i$  parameter is identified with conditions that determine whether one uses analytic or future-oriented attending for time estimation. The most important of these is the temporal coherence of the events to be judged. Thus, assume that hierarchical (H) and nonhierarchical (NH) melodies fill standard and comparison intervals as shown in Figure 5 (H – NH or NH – H). Temporal coherence is greater when H melodies are involved and also when they function as standards. In these cases, dynamic complexity is low and hence  $W_i$  should be high. Thus,  $W_i$  is high when  $C_m$  for situation  $i$  is an invariant and low integer ratio. In other words, estimates of  $W_i$ ,  $\hat{W}_i$ , should be higher with (a) H events (than NH ones), and (b) pairwise comparisons of the form H – NH (than NH – H). Finally, instructions and task demands (goal/function) can also affect  $W_i$  by determining the likelihood of a given attending mode.<sup>14</sup>

### Attending Mode Parameters

In principle, although either attending mode is possible for a given event, they are not equally likely. Furthermore, because each mode is associated with its own brand of attunement failure, different time estimates can occur for the same situation. This is important because it reinforces the point that the way people attend influences their time estimates. We show this using two melodies filling identical time intervals.

*Estimation of  $DE_i$ .* If people attend in future-oriented ways to events of Figure 5, they will be biased by a pairwise contrast,  $DE_i$ . Estimates of  $DE_i$  ( $\hat{DE}_i$ ) are based on individual temporal contrasts in each event. Individual and pairwise contrasts at any level can be positive, negative, or zero. If contrasts involve the total time level of each event, then truly unequal time intervals obtain (see Appendix). But if they arise from time levels embedded within events of equal duration, as shown in Figure 5, then contrasts based on final embedded periods determine  $DE_i$ . Thus,  $DE_i$  is a difference of final contrasts,  $\Delta T_{ij} - \Delta T'_{ij}$ , from each of  $j$  patterns ( $[j = 1, 2]$  level subscripts suppressed). This parameter reflects a comparability of expectancy violations across both patterns:

$$DE_i = (\Delta T_{i2} - \Delta T'_{i2}) - (\Delta T_{i1} - \Delta T'_{i1}). \quad (8)$$

<sup>14</sup> Comparative judgment tasks have traditionally been distinguished in terms of prospective and retrospective designs. Such design differences can also influence  $W_i$ . In prospective tasks, people are told to compare the durations of two forthcoming intervals leading to a greater chance of future-oriented attending. In retrospective tasks, people engage in incidental activities during two time spans often based on analytic attending; later they learn they must compare durations. This variability contributes to divergence in the literature in that differential results between task type sometimes occur (e.g., Block, George, & Reed, 1980; S. W. Brown, 1985; Hicks, Miller, & Kinsbourne, 1976). We believe this mixture of results comes from interactions of instructions with event coherence.

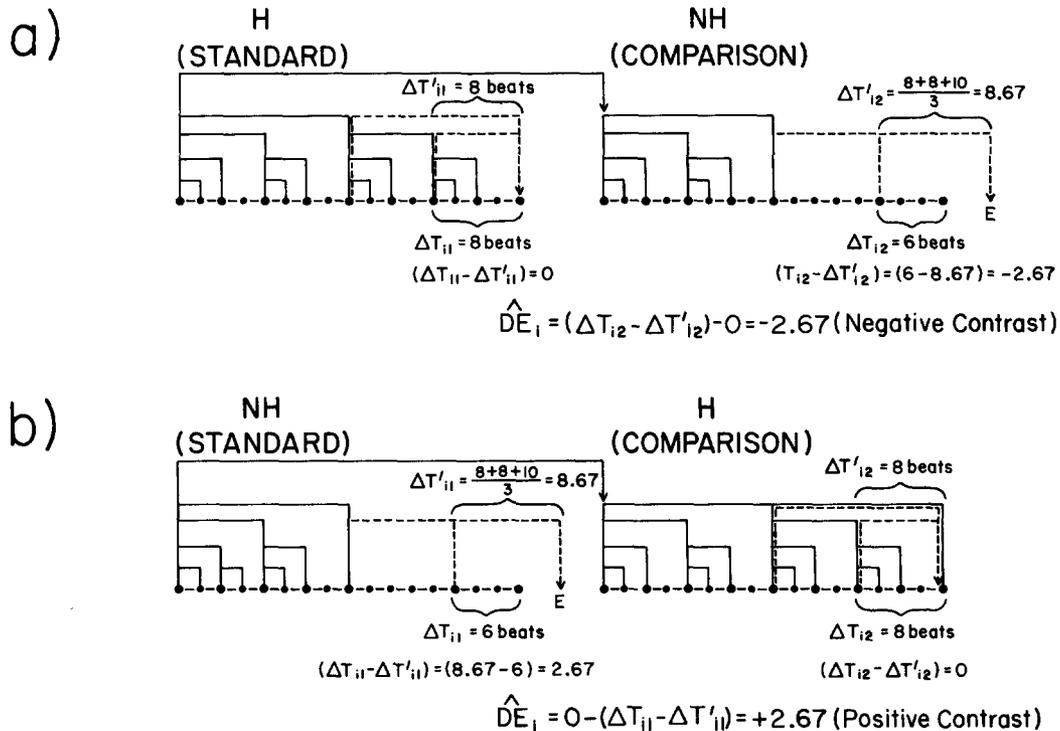


Figure 5. Abstracted time structures for two events, a hierarchical (H) and a nonhierarchical (NH) one, to be judged in a comparative time estimation task. (This event pair  $i$  produces two different kinds of temporal contrast,  $\Delta E_i$ , depending on event order. In (a) the H melody is the standard and the NH is the comparison yielding a negative contrast. In (b) the NH is the standard and the H is the comparison, yielding a positive contrast. These time structure represent the melodies of Figure 6. Smaller dots represent a common beat period and larger dots signify superordinate time periods of phrases marked by S accents. Expected ending time is denoted by E.)

Equation 8 simplifies whenever an expectancy about one pattern ending is confirmed because that contrast becomes zero. Thus, if the first pattern ( $j = 1$ ) of pair  $i$  confirms an expected ending time, then  $\Delta T'_{i1} = \Delta T_{i1}$  and

$$\Delta E_i = \Delta T_{i2} - \Delta T'_{i2}. \quad (9)$$

Different degrees and kinds of pairwise contrasts are given by Equations 8–9. For example, strong expectancies over the H – NH arrangement (Figure 5a) should appear because  $C_i = 2$  and a common referent period (quarter note) obtains for both events. Anticipations, reinforced in the standard, will occur in the comparison for focal periods of 8 beats or more (i.e.,  $\Delta T'_{i2}$  is the average duration of preceding phrases in the comparison:  $\Delta T'_{i2} = (8 + 8 + 10)/3 = 8.67$  beats). This expectancy is ultimately violated in the comparison in which a final tonic occurs too early (after 6 beats), yielding a negative pairwise contrast of  $\hat{\Delta E}_i = -2.67$  beats (Equation 9). Reversing the order of H and NH events (Figure 5b) leads to a positive pairwise contrast of the same magnitude ( $\hat{\Delta E}_i = 2.67$ ). More generally, Equations 8–9 indicate that the magnitude and direction of any pairwise contrast depend on discrepancies between observed and expected endings within both patterns.

*Estimation of  $DM_i$ .* If people attend analytically to these same events, they are biased by  $DM_i$ . This may apply if people resort to counting lower order features within each event. Estimates of  $DM_i$  ( $DM_i$ ) will not necessarily agree with contrast

estimates ( $\hat{\Delta E}_i$ ). If  $DM_i$  is based on a counting strategy, then let  $N1_i$  and  $N2_i$  denote, respectively, the number of items, chunks, or beats that are counted within the  $i$ th pair of standard and comparison events at the focal levels used. We assume that people often tacitly infer that associated analytic attending levels are the same for both events; this common focal level is denoted by  $\Delta T'_i$  (suppressing level subscripts). Thus,  $DM_i$  is estimated by

$$DM_i = \Delta T'_i(N2_i - N1_i). \quad (10)$$

For example, if attending is confined to the referent level, then  $DM_i$  is estimated from differences in the number of beats counted in each event. In the events of Figure 5a and 5b, the assumption of equivalent focal periods is correct and so correctly yields  $\hat{DM}_i = 0$ , because  $N1_i = N2_i$ . But, whenever beats (or other attended-to information) mark out different  $N1_i$  and  $N2_i$  values, this assumption is incorrect. Consequently, time distortions are predicted (filled duration effect). Supporting evidence for Equation 10 applied to beat counting is found in studies showing that time estimates are systematically biased when people judge melodies having identical durations but different total numbers of beats ( $N1 \neq N2$ ); Boltz, 1985).<sup>15</sup>

<sup>15</sup> Boltz (1985) found that when two melodies are identical in duration, as well as in number of tones and number of pause-defined chunks, they are nonetheless judged to differ in duration as a function of the number of different beats, with those containing more beats judged to be longer.

In sum, attending mode parameters can assume different values for the same pair of events. Furthermore, either parameter can yield over- or underestimations of a given interval. The degree to which one or another parameter estimate holds sway depends on its relative weight (Equations 5–7). Temporal contrast effects will predominate if future-oriented attending occurs (high  $W_i$ ), whereas mnemonic differences will predominate if analytic attending occurs (low  $W_i$ ). In general, this leads to predictions that variability and disagreements in time estimates are to be expected. Explanatory orderliness comes only with reference to event structure and dynamic attending.

### *Contrast Model Predictions*

The Contrast Model predicts that judgment accuracy (Equation 5), average bias (Equation 6), and variance (Equation 7) will vary systematically with model parameters. Among its important predictions are the following: (a) Time estimates will be biased in magnitude and direction by values of contrast ( $DE_i$ ). (b) Biasing effects of a given contrast will increase directly with temporal coherence, namely with increments in  $W_i$ . (c) Estimates of  $W_i$  will be a function of pairwise coherence, as indexed by  $C_m$ , such that  $\hat{W}_i$  will be high whenever a pair contains H patterns and especially when these serve as standards (ratio complexity constraints). (d) Judgment variability should decrease with increasing predominance of one attending mode (as  $W_i \rightarrow 1.0$  or  $0.0$ ).

Numerous other predictions can be deduced from Equations 5–10. Some concern effects of mnemonic strategies ( $DM_i$ ), which should be most pronounced when  $W_i$  is low. Others concern range effects arising from dominance region constraints. Finally, additional parameter influences on judgment variance can be anticipated from Equation 7.

These same equations also generate corresponding predictions about judgment accuracy and discriminations of unequal time intervals. These appear in the Appendix where related models such as Kristofferson's (1977) real time criterion model are considered. Systematic departures from Weber's function are predicted by the Contrast Model as a function of task, attending mode, skill, and event structure. However, our focus here and in Part 4 is on the case of equal time intervals and in particular on the first several predictions previously outlined.

### *Time Estimation in Production Tasks: Expectancies*

The Contrast Model only applies to comparative judgment situations. But time estimation has also been studied in temporal reproduction and production tasks (cf. Allan, 1979; Baker, 1962; Bartlett & Bartlett, 1959; Johnson, 1986; D. N. Lee, 1980). In reproduction, subjects recall a time span by overtly marking its beginning and end. In production, a specified span (e.g., 1 s) is constructed. A variant of the latter involves timed extrapolations in which part of an interval is presented and the appropriate remaining time must be produced. Analyses of reproduction and production tasks are beyond the scope of this article. However, temporal extrapolations are relevant because they directly involve expectancies.

Temporal extrapolations require an individual to use the pre-

ceding context to indicate (overtly) when an event should end in real time. Thus, they have potential for revealing expectancies because only the initial part of an event occurs and one must covertly continue it. Attending that controls performance occurs in the absence of the event itself; only "pure" expectancies remain.

Predictions based on expectancies rely on Equation 2. Assuming that a referent period,  $\Delta T_r$  and a rhythmic generator,  $C_r$ , are supplied by the preceding context, Equation 2 suggests the existence of sets of nested expectancies. Let  $\Delta T_n$  be the value of that focal period most relevant to a particular task. Produced and extrapolated times reflect such expectancies, and these can be predicted from average values of  $\Delta T_n$  within the relevant event. Thus, focal periods of 8 beat phrases within a musical event allow estimation of the extrapolated completion time, providing, of course, that the listener uses expectancies to time extrapolations. By contrast, analytic attending should not produce extrapolation times that are systematically related to an event's higher order time periods.

### *Summary of Part 3*

Time estimation in judgment and production tasks involves dynamic attending over various levels of structure within events that fill time intervals. According to the Contrast Model, means and variances of comparative duration judgments are determined by the likelihood of different attending modes (future-oriented, analytic) and their related potentials for attunement failure (temporal contrast, mnemonic activity). Over- and underestimates can occur with either future-oriented or analytic attending, with the nature and extent of biasing depending on event structure, its complexity, and attending mode itself. According to the production model, extrapolations involve only use of event structure. Thus, it predicts that produced time estimates will be systematically affected by expectancies but not by temporal contrast. Together, both models reinforce new definitions of complexity and, at the same time, suggest some reasons for seemingly diverse experimental findings regarding nontemporal information and time estimation.

## Part 4: Time Judgments and Productions

Two experiments address issues raised by the Contrast Model and other models presented in Part 1. Together, they also explore the relation between time judgment and production tasks. The first experiment tests the hypothesis that temporal contrast affects judgments of identical time intervals (Boltz, 1985). The second uses a production task to provide converging evidence that temporal contrast arises from temporal expectancies about event endings.<sup>16</sup> In both, events are melodies based on simple folk tunes.

<sup>16</sup> Experiment 1 is one of several from Marilyn Boltz's (1985) dissertation. Experiment 2 and the Contrast Model were conceived by Mari Riess Jones.

EXCERPT FROM "THERE'S MUSIC IN THE AIR" BY GEORGE ROOT

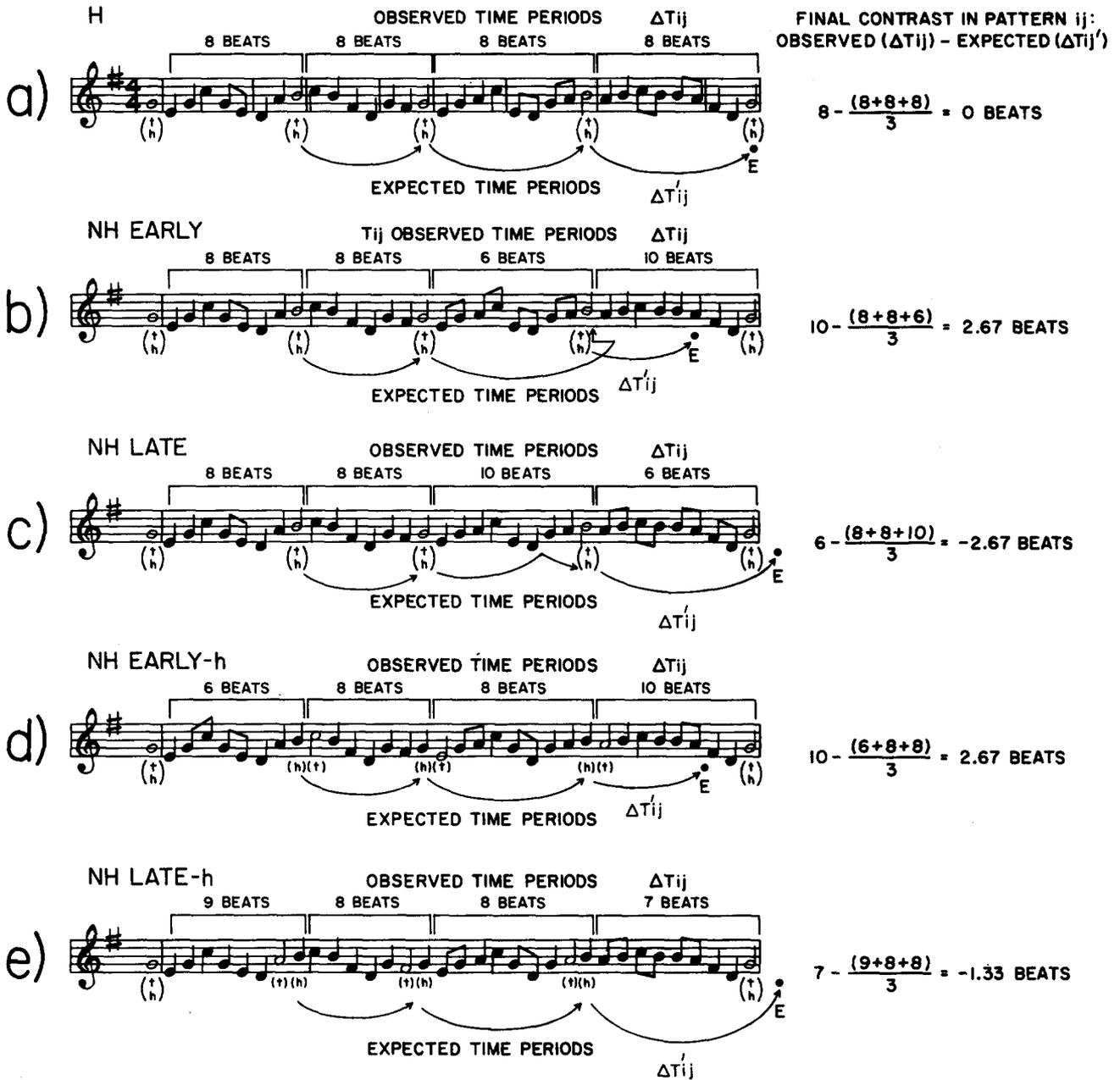


Figure 6. Excerpt from a folk tune in G major in which relative durations of phrases (meter constant) are varied. ((a) H is the original excerpt; (b) through (e) are temporal variations that adjust relative durations of structural phrases marked either by coupled t and h accents (b, c) or by h accents alone (d, e). Single pattern contrasts are based on observed minus expected times where expected times, in turn, are based on contextually governed expectancies about final phrase durations.)

Experiment 1: Judgments of Melodic Durations

The first study assesses predictions of various time estimation models, including the Contrast Model, in comparative judgment tasks. Moderately skilled listeners judged the relative duration of two melodies having identical total durations. Across

pairs of melodic patterns, their time structure was manipulated to vary (a) temporal contrast through confirmation and violation of expectancies and (b) degree of temporal coherence through structural accent timing (melodic phrasing).

Melodies within a pair always involved a common folk tune based on the same (duple) meter. However, one of the two tunes

could differ relative to the other in the timing of its melodic phrase structure. One tune and four temporal variations appear in Figure 6. The original tune (Figure 6a) has a hierarchical form where harmonic accents corresponding to tonic triad members coincide with temporal accents arising from prolonged durations. These joint accents functioned as structural markers of melodic phrase endings (Cone, 1968; Kramer, 1982; Todd, 1985). Within each H tune, structural accents always recurred after an invariant number of beats and were predicted to promote temporal expectancies about phrase and event endings that are reliably confirmed. All variations of an H tune introduced additive time changes through timing modifications of structural accents. In particular, all variants were NH tunes containing contrasts of specific magnitude and direction. Some ( $NH_{\text{early}}$  and  $NH_{\text{early-h}}$ ) were designed to promote temporal expectancies for early endings, and in these, melodic phrasing was initially regular so that focal time periods were, on the average, shorter than the final one. Others ( $NH_{\text{late}}$  and  $NH_{\text{late-h}}$ ) were similarly designed, but with the aim of promoting expectancies for late endings; thus, initial phrases had focal periods that were, on the average, longer than the final one.

Temporal coherence was varied over an event pair in several ways—all related to the overall degree of temporal regularity in musical phrasing (i.e., not meter; Part 2). Both the presence of a simple (H) time structure and its relative placement in the event pair (event order) were varied. A third potential source of coherence involved the clarity of focal periods. The prominence of structural phrasing was varied through the coupling or decoupling of harmonic and temporal accents. Temporal contrast effects should be enhanced in events with coupled accents ( $NH_{\text{early}}$  and  $NH_{\text{late}}$  melodies), relative to those with less prominent, decoupled accents ( $NH_{\text{early-h}}$  and  $NH_{\text{late-h}}$  melodies). Although the temporal accents of these melodies do recur with invariant time periods (hence they are hierarchical), temporal contrast arises from harmonic accents that appear either too early or too late. The extent to which such manipulations of time markers are, in fact, perceived by listeners was validated in an independent rating study in which a different group of skilled listeners were asked to judge accentuation patterns in all melodies (see Method section).

### Model Predictions

#### The Contrast Model

The Contrast Model assumes that attending mode and environmental structure jointly influence time judgments. Those predictions from Part 3 involve (a) attending mode and temporal contrast, (b) attending and mnemonic activity, and (c) environmental structure and temporal coherence.

*Future-oriented attending and temporal contrast.* The temporal contrast parameter,  $DE_i$ , is of primary interest. For a pair of equivalent intervals, time judgments should differ from equality as a function of the magnitude and sign of  $DE_i$ , assuming moderate to high values of its weight,  $W_i$  (Equation 6). In this study, pairwise contrasts range from minima in some H – H pairs to maxima in some NH – NH pairs (e.g.,  $NH_{\text{early}}$  –  $NH_{\text{late}}$ ).

*Analytic attending and mnemonic activity.* The  $DM_i$  values are estimated according to assumed mnemonic strategy. Al-

though numerous strategies are possible (see Predictions of Alternative Theories ahead), for purposes of evaluating the Contrast Model, we assume a beat counting strategy in which  $\overline{DM}_i = 0$ .

*Environmental structure and temporal coherence.* Event coherence involves the temporal regularity of marked time levels embedded within the event pair. Because the coherence weight,  $W_i$ , reflects ratio complexity constraints on  $C_m$ ,  $\hat{W}_i$  should be higher with pairs including H events and with those that begin with H tunes (i.e., H – NH). Theoretically, accent decoupling can also affect coherence by rendering either harmonic or temporal regularities less noticeable. The  $W_i$  parameter gives greater weight to prominently marked contrasts.

#### Predictions of Alternative Theories

Alternative theories are memory oriented in nature and base predictions about judgment bias on memory code complexity. With regard to melodies, specific predictions rely heavily on musical coding theory, which assumes that code complexity depends on the number of pitch rule changes and the number of encoded chunks (e.g., Deutsch & Feroe, 1981; see Jones, 1985, for a review). Complexity is given by the resulting code length for a melody. This approach to complexity is not dynamic and does not involve expectancies. Nonetheless, according to time estimation models in Part 1, encoding complexity can vary in these melodies. This leads to three testable models of time judgment.

*Null model.* A null model predicts no time differences among melodies, because in all melodic pairs (H and NH) the number of tones, chunks (designated by temporal accents), pitches, and pitch rule changes are equated. By these criteria, the nontemporal information that fills both time intervals is equally complex. Thus, one interpretation of the storage size hypothesis (Ornstein, 1969), attentional effort models (e.g., Underwood & Swain, 1973), and a context change model (e.g., Block, 1978) is that no memory code or strategy differences exist to affect duration judgments.

*Chunk model.* A chunk model can also be generated from the storage size hypothesis. It predicts that NH melodies should seem longer than H melodies because the former contain chunks (phrases) of different lengths and thus are more complex.

*Change model.* A change model in its strictest form (as reviewed in Block, in press) is somewhat difficult to adapt to musical patterns because it is not always clear which types of structural change promote changes in processing strategy. However, because accents reflect contextual change, the most plausible predictions center on their parsing potential. The NH melodies should be more complex than H melodies, and it is also plausible that NH melodies with decoupled accents should seem longest of all. The idea is that nonhierarchical melodies generally display more unpredictable changes than hierarchical ones, whereas NH – h melodies specify still more contextual change because chunks are defined by either h or t accents. If people encode using both harmonic and temporal accents, then in NH – h melodies, their strategies will undergo most changes, perhaps even ones associated with conflicting tonalities arising from decoupled accents.

In sum, the Contrast Model relies on a dynamic interpreta-

tion of complexity to predict systematic judgment biases due to contrast and temporal coherence. Alternative models rely upon different interpretations of complexity to predict judgment biases due only to memory codes.

### Method

#### Design

The design was a  $5 \times 5 \times 3 \times 2$  mixed factorial. Five time structures (H, NH<sub>early</sub>, NH<sub>late</sub>, NH<sub>early-h</sub>, NH<sub>late-h</sub>) existed for each pattern, yielding 25 pairs ( $5 \times 5$ ) of to-be-judged melodies ( $i = 1, 25$ ). Three folk-tune instances were used within each level. Counterbalance order (1, 2) was the single between-subjects' factor. A total of 16 subjects, each with at least 4 years of musical experience within the past 6 years, were randomly assigned to one of the two counterbalance orders.

#### Stimulus Materials

**Pattern construction.** All of the patterns consisted of computer-generated square waves of equivalent loudness levels. They were based on adapted excerpts (lengths ranging from 34 to 37 notes) from three folk tunes. All spanned four melodic phrases that covered intact thematic statements and that were completed by the tune's ending, which in all cases was the tonic (Figure 6a is an exemplar). Although H tunes could be in different keys, all were scored in a duple meter, where the SOA for  $\downarrow = 300$  ms. In addition, all H tunes contained (a) coupled structural accents of elongated notes ( $t$ ) and stable pitches ( $h$ ), and (b) an invariant number of beats (e.g., 8) between these accents. Other structural details common to all H and NH patterns were as follows: (a) All contained the same number of beats. (b) "Keyness" of all melodies was ensured by inclusion of certain key-specific musical intervals (cf. H. Brown & Butler, 1981) and by initiating each with a lengthened tonic. (c) Structural accents ( $h, t$ ) alone defined the relevant (structural) time hierarchy; no metric stresses reinforced the metric time hierarchy. (d)  $h$  accents always corresponded to the tonic, dominant or mediant of the established key (Krumhansl & Kessler, 1982); in two H instances (and their NH variations), this required changes in a few scored  $h$  accents for greater phrase emphasis. (e) In three of the five time structures,  $h$  and  $t$  accents were coupled (i.e., H, NH<sub>early</sub>, NH<sub>late</sub>). (f) In the two remaining conditions (NH<sub>early-h</sub>, NH<sub>late-h</sub>),  $h$  and  $t$  accents were decoupled, separated in time by  $\pm k = 300$  ms. In NH<sub>early-h</sub>,  $h$  accents occurred on notes immediately prior to  $t$  accents, thereby shortening the first phrase in these tunes; in NH<sub>late-h</sub>, the  $h$  accents followed  $t$  accents, thereby lengthening this phrase. Corresponding final time periods were 10 and 7 beats for the two tunes. In decoupled structures, only  $h$  accents marked nonhierarchical time levels;  $t$  accents remained strictly hierarchical. Thus, NH<sub>early-h</sub> and NH<sub>late-h</sub> corresponded to NH<sub>early</sub> and NH<sub>late</sub> in weakening temporal coherence via  $h$  markers, but they corresponded to H tunes with regard to  $t$  markers. They are termed *H-type melodies*.

**Pattern structure validation: Rating study.** To assess the psychological validity of accent manipulations in these excerpts, an independent group of 18 musically trained listeners were asked to rate all melodies (using a 10-point scale) on two dimensions. For a dimension of accent timing, subjects judged whether a melody's temporal accents seemed well timed—occurring when expected—or whether they seemed ill-timed—occurring earlier or later than expected. The second dimension was one of harmonic appropriateness. Here, subjects judged whether temporal accents seemed appropriate to the melody's key and whether they occurred on tonally appropriate pitches. As anticipated, accents within the hierarchical melodies were considered to be the most well timed ( $M = 3.10$ ) and tonally appropriate ( $M = 3.09$ ). Furthermore, all three H instances were judged to be highly appropriate, harmonically, with no significant tune differences ( $F \leq 1.0$ ), thus indicating that accent phrasing and  $h$ -plus- $t$  coupling in H versions of the adapted excerpts

Table 1  
Mean Judged Duration Ratings for Pattern Pairs  
in Experiment 1 as a Function of Melodic  
Time Structure (Boltz, 1985)

Standard melody ( $j = 1$ )	Comparison melody ( $j = 2$ )				
	1	2	3	4	5
1. H	4.04 <sub>(1)</sub>	4.60 <sub>(1)</sub>	3.51 <sub>(1)</sub>	4.78 <sub>(1)</sub>	3.22 <sub>(1)</sub>
2. NH <sub>early-h</sub>	3.50 <sub>(2)</sub>	3.99 <sub>(3)</sub>	3.61 <sub>(3)</sub>	4.51 <sub>(3)</sub>	3.59 <sub>(3)</sub>
3. NH <sub>late-h</sub>	4.64 <sub>(2)</sub>	4.55 <sub>(3)</sub>	4.02 <sub>(3)</sub>	4.67 <sub>(3)</sub>	3.66 <sub>(3)</sub>
4. NH <sub>early</sub>	3.48 <sub>(2)</sub>	4.05 <sub>(4)</sub>	3.83 <sub>(4)</sub>	4.07 <sub>(4)</sub>	3.52 <sub>(4)</sub>
5. NH <sub>late</sub>	4.51 <sub>(2)</sub>	4.22 <sub>(4)</sub>	4.12 <sub>(4)</sub>	4.66 <sub>(3)</sub>	4.03 <sub>(4)</sub>

*Note.* A value of 4 indicates equivalent duration; values less than 4 indicate that the comparison was judged shorter. Numbers in parentheses indicate assigned coherence group of Figure 7. H = hierarchical; NH = nonhierarchical;  $h$  = harmonic accents.

were tonally meaningful to listeners. With respect to variations on H tunes, those with coupled temporal and harmonic accents that deviated in temporal regularity (NH<sub>early</sub>, NH<sub>late</sub>) were rated the least well-timed ( $M = 6.22$ ,  $M = 6.11$ , respectively) but only slightly less appropriate to the melody's key ( $M = 4.09$ ;  $M = 4.02$ ). A set of Tukey post hoc comparisons confirmed that these ratings significantly differed from those of the H melody ( $p < .01$ ). The H-type melodies, namely those with regularly recurring temporal accents but decoupled harmonic ones (NH<sub>early-h</sub>, NH<sub>late-h</sub>), did indeed seem more well timed than the NH<sub>early</sub> and NH<sub>late</sub> tunes ( $M = 4.85$ ,  $M = 4.56$ ). As predicted, their harmonic accenting was deemed less appropriate ( $M = 6.02$ ;  $M = 6.38$ ). These ratings differed significantly from those of both H and other NH melodies ( $p < .01$ ).

Overall, these results suggest that listeners found H versions of folk-tune excerpts to have appropriate harmonic and temporal accentuations. These tunes confirm expectations that the correct pitch accents coupled with temporal ones occur in the right spots. But listeners also clearly discern  $t$  accent regularity of H-type tunes. Deviations from these regularities result in accents that are perceived as occurring relatively early or late.

#### Procedure

Subjects listened over AKG (Model K240) headphones to 150 randomized pairs of melodies (25 pairs for each of three songs with one repetition each), each pair signaled by a 1-s warning tone. Melodies within each pair always spanned the same total duration (10.3 s).

Subjects were asked to compare the total durations of two melodies within a pair (and to remove watches). They then heard all pairs of melodies and indicated on a 7-point scale whether the second melody seemed longer than, shorter than, or the same duration as the first—*very short* (1), *very long* (7), *same* (4), the null response score.

#### Results and Discussion

Mean duration judgments for the 25 pairwise conditions are shown in Table 1. Values less than 4.00 indicate that the second pattern of a pair seemed shorter; values greater than 4.00 indicate that it seemed longer.

#### Basic Findings

Duration judgments differed significantly as a function of information filling both periods, in spite of the fact that no durational differences really existed. Average absolute deviations

Table 2  
Predicted and Signed Bias Values (PSE<sub>i</sub>) of Four Models

Melodic pair	$\widehat{PSE}_i$ predictions									Observed data PSE <sub>i</sub>	
	Alternative models						Contrast model				
	Null DM <sub>1</sub> <sup>*</sup>	PSE <sub>i</sub>	Chunk DM <sub>1</sub> <sup>*</sup>	PSE <sub>i</sub>	Change DM <sub>1</sub> <sup>*</sup>	PSE <sub>i</sub>	$\widehat{DM}_i$	$\widehat{DE}_i$	$\widehat{PSE}_i$		
Group 1: High coherence											
	$\hat{W}_1 = 0$		$\hat{W}_1 = 0$		$\hat{W}_1 = 0$		$\hat{W}_1 = .89$				
H - H	0	0	0	0	0	0	0	0	0	.04	
N - NH <sub>late</sub>	0	0	+	+	+	+	0	-2.67	-.72	-.78	
H - NH <sub>early</sub>	0	0	+	+	+	+	0	2.67	.72	.78	
H - NH <sub>late-h</sub>	0	0	+	+	++	++	0	-1.33	-.35	-.49	
H - NH <sub>early-h</sub>	0	0	+	+	++	++	0	2.67	.72	.60	
SD <sub>1</sub>										.204	.273
Group 2: Medium coherence											
	$\hat{W}_2 = 0$		$\hat{W}_2 = 0$		$\hat{W}_2 = 0$		$\hat{W}_2 = .70$				
NH <sub>late</sub> - H	0	0	-	-	-	-	0	2.67	.56	.51	
NH <sub>early</sub> - H	0	0	-	-	-	-	0	-2.67	-.56	-.52	
NH <sub>late-h</sub> - H	0	0	-	-	-	-	0	1.33	.28	.64	
NH <sub>early-h</sub> - H	0	0	-	-	-	-	0	-2.67	-.56	-.50	
SD <sub>2</sub>										.334	.375
Group 3: Medium coherence											
	$\hat{W}_3 = 0$		$\hat{W}_3 = 0$		$\hat{W}_3 = 0$		$\hat{W}_3 = .40$				
NH <sub>late-h</sub> - NH <sub>early</sub>	0	0	0	0	-	-	0	4.00	.48	.67	
NH <sub>late-h</sub> - NH <sub>early-h</sub>	0	0	0	0	0	0	0	4.00	.48	.55	
NH <sub>late-h</sub> - NH <sub>late</sub>	0	0	0	0	-	-	0	-1.33	-.16	-.34	
NH <sub>late-h</sub> - NH <sub>late-h</sub>	0	0	0	0	0	0	0	0	0	.02	
NH <sub>early-h</sub> - NH <sub>early</sub>	0	0	0	0	-	-	0	0	0	.51	
NH <sub>early-h</sub> - NH <sub>early-h</sub>	0	0	0	0	0	0	0	0	0	-.01	
NH <sub>early-h</sub> - NH <sub>late</sub>	0	0	0	0	-	-	0	-5.34	-.64	-.41	
NH <sub>late</sub> - NH <sub>early</sub>	0	0	0	0	0	0	0	5.34	.64	.66	
SD <sub>3</sub>										.489	.334
Group 4: Low coherence											
	$\hat{W}_4 = 0$		$\hat{W}_4 = 0$		$\hat{W}_4 = 0$		$\hat{W}_4 = .22$				
NH <sub>early</sub> - NH <sub>late-h</sub>	0	0	0	0	+	+	0	-4.00	-.26	-.17	
NH <sub>early-h</sub> - NH <sub>late-h</sub>	0	0	0	0	0	0	0	-4.00	-.26	-.39	
NH <sub>late</sub> - NH <sub>late-h</sub>	0	0	0	0	+	+	0	1.33	.08	.12	
NH <sub>late</sub> - NH <sub>late</sub>	0	0	0	0	0	0	0	0	0	.03	
NH <sub>early</sub> - NH <sub>early-h</sub>	0	0	0	0	+	+	0	0	0	.05	
NH <sub>early</sub> - NH <sub>early</sub>	0	0	0	0	0	0	0	0	0	.07	
NH <sub>late</sub> - NH <sub>early-h</sub>	0	0	0	0	+	+	0	5.34	.35	.22	
NH <sub>early</sub> - NH <sub>late</sub>	0	0	0	0	0	0	0	-5.34	-.35	-.48	
SD <sub>4</sub>										.422	.373

Note. H = hierarchical; NH = nonhierarchical; h = harmonic accents.

from a null difference (4.0) took on a maximum value of 0.78, which represents a 22% unidirectional shift on the rating scale. Overall, the interaction of time structure associated with the first and second patterns was significant beyond the .0001 level,  $F(16, 224) = 13.67, MS_e = 0.088$ .

The nature of this interaction supports the Contrast Model predictions. People judged patterns that seemed to end “too soon” to be short and those that seemed to end “late” to be long. Post hoc analyses (Tukey) verified these findings: (a) Type of contrast produced significant differences ( $p < .01$ ) between

pairs with early versus late endings (e.g.,  $NH_{early}$ ,  $NH_{late}$ ). (b) Magnitude of pairwise contrast produced predicted differences in that negligible contrast effects appeared in pairs with equivalent single pattern contrasts (e.g.,  $NH_{early}$  with  $NH_{early}$ ) and in pairs with no contrasts (e.g., H – H). Pairs with nonzero contrasts, however, yielded systematic differences. (c) Predicted differences as a function of temporal coherence were significant, with temporal contrast having a greater impact in pairs containing an H melody ( $p < .01$ ). (d) Effects of contrast were significantly less for pairs containing a melody with decoupled accents ( $p < .01$ ). (e) Temporal order effects were significant in that a given contrast had more impact whenever an H ( $p < .01$ ) or H-type (i.e.,  $NH_h$ ) melody ( $p < .05$ ) functioned as a standard. (f) Overall, these findings were robust in generalizing across all three melodic instances ( $F \approx 1.0$  for melodic instance) and to 12 of the 16 subjects involved (4 subjects tended to show strong contrast effects only in H and not in H-type melodies).

### Model Predictions: Contrast Model Versus Alternatives

Several simplifying assumptions are introduced to facilitate predictions about judgment means (from Equations 5–6) and variance (Equation 7) as a function of contrast and temporal coherence.

1. To evaluate the Contrast Model, the linear function  $f$  with slope  $a$  and intercept  $b = 4$  (the null response) relates the response scale (1–7) to bias predictions of the model. From Equation 6:

$$PSE_i = JD_i - 4 = a[W_i DE_i + (1 - W_i) DM_i]. \quad (11)$$

2. To evaluate the Contrast Model, attending mode parameters in Equation 11 are estimated using Equations 8 and 10. The  $DE_i$  estimates (Equation 8) are based on the assumption that skilled attenders can use higher order focal periods of musical phrases, an assumption supported by the rating task data reported in the Method section. The  $DM_i$  estimates (Equation 10) are based on the assumption that listeners may alternatively rely on the referent level and count beat periods. This assumption is supported not only by research described earlier (Boltz, 1985; also Footnote 15), but by the present data. That is, under this strategy,  $DM_i = 0$ , and we show that alternative models based on  $DM_i \neq 0$  are not appropriate for the present task (Table 2). Here,  $DM_i$  is weighted by  $1 - W_i$ . This yields the most conservative test of the contrast model by reinforcing the true (null) time difference and forcing contrast alone to bias judgments.

3. Each of the 25 melodic pairs was assigned to one of four ( $k = 1, 2, 3, 4$ ) experimentally designated groups predicted to differ in temporal coherence, as shown in Figure 7. Pairs were categorized according to whether they contained an H or H-type melody and whether this melody was the standard. Thus, event pairs in Groups 1 and 2 have the same absolute amount of contrast but incorporate an H melody either as a standard (Group 1) or as a comparison (Group 2). Similarly, pairs in Groups 3 and 4 involve only NH melodies, in which an H-type pattern is the standard (Group 3) or comparison (Group 4). Interpretations of this parameter lead to the expectation that  $\hat{W}_1 > \hat{W}_2 > \hat{W}_3 > \hat{W}_4$  for  $\hat{W}_k$ . Note that this partitioning has been simplified at the expense of ignoring coherence effects due to marker prominence. For example, Group 1 includes both H –  $NH_{early}$  and H –  $NH_{early-h}$ . If marker prominence is, in fact,

very important, then this particular partitioning may yield poor fits of the Contrast Model.

Predictions of the Contrast Model are summarized in Tables 2 and 3. A priori estimates of  $DM_i$  and  $DE_i$  are in columns 8 and 9 of Table 2, respectively; the four  $W_k$  (above these columns) are estimated from the data. Preliminary model fits indicated that the common variance of difference scores,  $\sigma^2$ , was extremely small, and so it was fixed at zero. Thus, five parameters were estimated from the data: four  $W_k$  and a scale factor,  $a$ . Model fits, using Equation 11, were based on a nonlinear least squares curve-fitting technique (NONLIN), which minimized errors between predicted and observed scores.<sup>17</sup> Best-fitting  $\hat{W}_k$  are in Table 2;  $\hat{a}$  was .303. This version of the contrast model accounts for 88% of the variance ( $R^2$ ).

Also in Table 2 are predicted and observed values of  $PSE_i$  (columns 10 and 11, respectively) and  $\hat{SD}_k$  (pooled for the  $k$ th group) for the Contrast Model, along with predictions of other models.<sup>18</sup> Predictions of all models derive from Equation 6. However, those of the Contrast Model are based on nonzero estimates of  $W_k$  (and  $1 - W_k$ ), whereas those of all other models assume  $W_k = 0$ . In terms of Equation 6, when  $W_k = 0$ , effects of contrast are eliminated and memory effects are maximal [ $(1 - W_k) = 1.0$ ]. The memory parameter for alternative models is denoted  $DM^*$ , because it differs conceptually from  $DM_i$  of the Contrast Model and also because most models yield only ordinal assessments of memory differences. These are in columns 2, 4, and 6. Corresponding directional predictions of bias (greater or less than) for all melodic pairs are shown in columns 3, 5, and 7 for the Null, Chunk, and Change Models, respectively.

A summary of directional (+, –) prediction accuracy over the four models (of Table 2) appears in Table 3. These indicate that the Contrast Model fares best, predicting 24 of the 25 possible directional differences in time judgments. The Change Model is next with 11 of the  $PSE_i$ s in the correct direction. In other respects, the Contrast Model also proves superior. It correctly predicts (a) highest temporal coherence for Group 1 ( $\hat{W}_1 = .84$ ) and lowest coherence for Group 4 ( $\hat{W}_4 = .22$ ), (b) maximal temporal contrast effects with least temporal uncertainty, namely in Group 1, (c) greater contrast effects for a hierarchical or H-type standard (Groups 1 and 3) than with an NH standard (Groups 2 and 4) because  $\hat{W}_1 > \hat{W}_2$  and  $\hat{W}_3 > \hat{W}_4$ , and (d) least variability in Group 1 in which  $\hat{W}_1$  is highest.

The Contrast Model yields a profile that is more accurate and informative than alternatives. However, it does not accommodate all of the data within this simplified grouping scheme equally well, and its failures are instructive. It describes performance better in melodic pairs that have clearly marked focal time periods (i.e., in pairs with coupled accents). In fact, when significant model deviations do occur, it is in melodies with decoupled accents. These patterns present less obvious violations

<sup>17</sup> NONLIN (Ohio State Science Package) operates on a least squares algorithm developed by Marquardt (1963), which is basically an interpolation between Newton-Raphson and gradient methods.

<sup>18</sup> The  $PSE_i$  predictions of the Contrast Model (column 10) are based on Equation 11 and variance predictions on Equation 7 (with  $\hat{a} = .303$ ). The  $\hat{DM}_i$  and  $\hat{DE}_i$  calculations (Table 2, columns 9, 8) are based on Equations 10 and 8, respectively. See text for  $DM^*$  calculations of Null, Chunk, and Change Models.

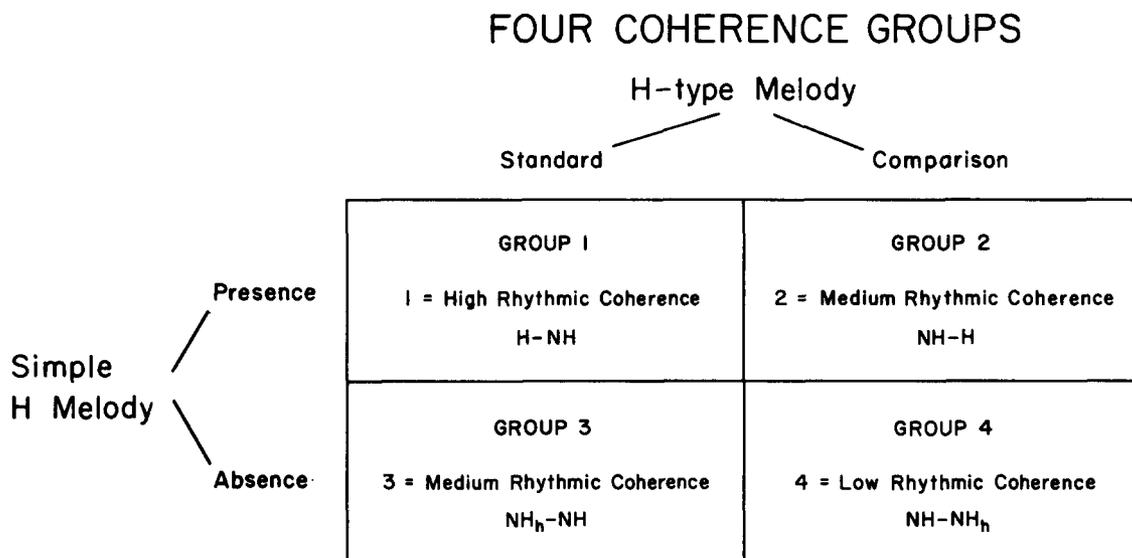


Figure 7. Four coherence groups of the 25 pairs of melodies used in Experiment 1: Groups 1, 2, 3, and 4.

of expectancies, and so this is not surprising. Although better model fits can be obtained with additional  $W_k$  estimates, effects of accentuation are nevertheless clear. We consider this application of the Contrast Model to be largely illustrative and take these discrepancies to underscore the broader theoretical point that the way in which nontemporal information marks time levels is important.

*Experiment 2: Productions of Melodic Durations*

A second study was designed to confirm that people actually expect patterns to end at particular times. Converging evidence for the psychological reality of expectancies is provided by a production task. A new group of listeners heard the same melodies, without the final three tones, and produced probable ending times with a button press. Thus, pure expectancies were made overt via extrapolated productions of the final focal period.

Some songs, such as NH<sub>early</sub>, should encourage anticipations for early endings by virtue of their shorter focal time periods. According to Equation 2, these melodies should lead to shorter final production times. Other songs, such as NH<sub>late</sub>, should generate expectancies about late endings and thus yield longer response times to completion. These predictions derive not from the Contrast Model, but from Equation 2 for timed productions.

Of interest is the reciprocity between produced time and

judged time. This involves the prediction that the pattern of produced times will be opposite to the pattern of judged times predicted (and observed) by the Contrast Model. In Experiment 1, for example, NH<sub>late</sub> patterns were predicted to be judged shorter than others, and they were. Here, the claim is if those duration judgments were based on temporal contrasts involving expectancies, then people will expect these same patterns to last relatively longer and this expectation will be evident in production. The converse, of course, should be true for NH<sub>early</sub> melodies.

Finally, the three alternative models do not predict a reciprocal pattern of results because none incorporates expectancies. Indeed, memory models taken literally are not particularly conducive to extrapolated times because time estimates are based on previously stored information. However, assuming that productions for a melody are governed by its code complexity, alternative models all imply that the pattern of produced times will be similar to judged times for the various H and NH tunes. Thus, either no differences in produced times should arise (Null Model) or NH melodies should yield longer produced times (Chunk and Change Models).

*Method*

*Design*

A 5 × 3 × 2 mixed factorial involved crossing the five types of accent/timing patterns (H, NH<sub>early</sub>, NH<sub>late</sub>, NH<sub>early-h</sub>, NH<sub>late-h</sub>) with three different songs (used in Experiment 1). Two counterbalance orders formed the single between-subjects factor; 9 subjects, each with at least 4 years of recent musical experience, were assigned to each counterbalance order.

*Stimulus Materials*

Stimulus materials were identical to those of Experiment 1, except that the final three tones were omitted from each melody. Three ran-

Table 3  
Summary of PSE<sub>i</sub> Predictions for Four Duration Models

No. of predictions	Alternative models			Contrast Model
	Null	Chunk	Change	
Total correct	6	10	11	24
Total possible	25	25	25	25

Table 4  
Results of Extrapolation Study (Experiment 2)

Measure	Conditions of melodic time structure				
	H	NH <sub>early</sub>	NH <sub>early-h</sub>	NH <sub>late</sub>	NH <sub>late-h</sub>
<i>Mdn</i> extrapolation time (in ms)	1154	899	1173	1209	1305
<i>Mdn</i> accuracy score (in ms)	-46	-301	-27	+309	+105
<i>M</i> no. of judged missing beats	2.67	1.59	2.56	3.11	3.24

*Note.* Entries in each row reflect, respectively, one of three dependent measures used in Experiment 2 to indicate expectancies about musical endings for five melodic conditions. Median extrapolation time for missing notes (row 1) and its accuracy (row 2; negative sign indicates anticipation) are given along with the mean number of beats judged missing (row 3). H = hierarchical; NH = nonhierarchical; *h* = harmonic accents.

domized blocks, each with 15 melodies, occurred within two counterbalance orders.

### Procedure

Subjects performed two tasks. The first involved extrapolations; the second required judgments of the number of missing notes.

*Production.* Recorded instructions told subjects to (a) listen to incomplete melodies whose final (1 to 5) notes, including the closing tonic, were missing, and (b) indicate when a melody should end if missing notes were played in real time. Four practice trials clarified the procedure.

On each trial, a 1-s warning tone (5000 Hz) preceded a melody by 2 s. When the last note of a melody was signaled by a panel light, subjects covertly continued the tune and indicated its true ending time by a button press. This extinguished the signal light, and the next trial began. This phase lasted 40 min.

*Judgments of missing tones.* Recorded instructions told subjects to judge, in writing, how many notes were missing from each melody. Subjects then heard the same 15 experimental melodies, now randomized according to one of two schedules. On each of 15 trials, a 5000-Hz warning tone preceded the melody by 2 s; after the melody, subjects had 5 s to respond.

### Results and Discussion

Table 4 presents the results in which averages are based on data collapsed over melodic instance and counterbalance order, both of which again produced negligible effects ( $F < 1.00$ ). Overall, the findings are remarkably congruent with those involving judged durations.

Average median extrapolation times for each melodic type are in the first row of Table 4. Mean differences were significant,  $F(4, 64) = 7.48$ ,  $MS_e = 54,724.97$ ,  $p < .0001$ , and reflect expectancies about early and late endings. Longest times were those predicted to generate expectancies for late endings (NH<sub>late</sub>, NH<sub>late-h</sub>), and the shortest times occurred for NH<sub>early</sub> tunes. Although NH<sub>late</sub> was significantly longer than NH<sub>early</sub> ( $p < .01$ , Scheffé), this was not true of NH<sub>late-h</sub> versus NH<sub>early-h</sub> ( $.10 < p < .05$ ), although differences here were in the predicted direction. Accuracy scores (in ms) appear in the second row of Table 4 and reflect temporal discrepancies between produced and observed endings. They should correspond to estimates of temporal contrast if people had actually heard the melody continue. Note

that pairs in which phrases were prominently marked by coupled structural accents, NH<sub>early</sub> versus NH<sub>late</sub>, yield greater differences than ones with less prominent (i.e., decoupled accents) markings, NH<sub>early-h</sub> versus NH<sub>late-h</sub>. People are relatively accurate in predicting endings of hierarchical melodies, and they are also accurate with NH<sub>early-h</sub> melodies. One reason for this may be that in those NH melodies with decoupled accents, temporal accentuation remains hierarchical. If people ignore harmonically marked phrasings and rely solely on *t* accents, then these NH melodies can indeed function as H-type time structures. At least in some cases, people did this.

Finally, the average number of missing beats reported by listeners (row 3, Table 4) converges with time measures. In all important respects, statistical analysis here confirms that similar profile differences exist and these generalize across the majority of subjects used in this study.

Overall, these data favor an interpretation that moderately skilled listeners detect and use temporal relationships to generate expectancies about musical endings. More recent work relying on different melodies and task requirements confirms these results (Van Vreden, Boltz, & Jones, 1989). Together, such findings reinforce the assumption that temporal contrast arises from expectancy violations. The pattern of over- and underestimated final periods, which is opposite to the pattern of duration judgments (Experiment 1), supports the idea that temporal contrast is based on listeners' extrapolations (temporal expectancies).

### General Discussion

Together, Experiments 1 and 2 suggest that duration judgments and overt productions of musical events reflect a common activity. Related conclusions have been drawn by others using different stimulus materials (Keele, Pokorny, Corcos, & Ivry, 1985; Kristofferson, 1984). Our explanation centers on the role of expectancies in extrapolations and in selective attentions in duration judgments. Equation 2, which describes extrapolations, and Equations 5–6, which describe judgments, together predict the observed pattern of reciprocal outcomes over the two tasks.

A primary aim of Experiments 1 and 2 was to approximate a natural context in which future-oriented attending would ap-

pear. Folk tunes, particularly coherent ones, fit the stimulus requirements. Reliance on moderately skilled musical listeners further ensured a possibility of future-oriented attending and simplified estimation of  $DE_i$ . Although both constraints limit experimental generality, they succeeded in demonstrating the existence of future-oriented attending. These findings suggest that when skilled attenders interact with relatively coherent time patterns, temporal expectancies will arise and duration judgments can be biased by temporal contrast. Related influences of temporal contrast have been reported in other contexts as well (Kidd, in press; Kristofferson, 1977; Lehiste, 1975, 1977).

The present results are difficult to explain in terms of contemporary models that emphasize memory for accumulated amounts of nontemporal information. For example, in Experiment 1 the Null Model predicts that all melodies should be judged equal in duration because the number of tones and pitch changes are identical. And yet some were judged longer (i.e.,  $NH_{early}$ ,  $NH_{early-h}$ ) and others shorter (i.e.,  $NH_{late}$ ,  $NH_{late-h}$ ). Differences in number of chunks or in changing accent patterns (Chunk, Change Models) also cannot account for these effects. Although melodies of a pair can differ in relative timing of certain accents, even models addressed to rhythmic chunking have difficulties because they fail to acknowledge that accent patterns incite dynamic and preparatory attending.

The Contrast Model weights the influence of contrast so that over- or underestimates of durations due to contrast are most evident in coherent events in which future-oriented attending occurs more often. It turns out that in event pairs with low temporal coherence (e.g., some NH-NH pairs), estimates of temporal contrast could be quite high. Yet, according to this model, these  $DE_i$  values should not have proportional impact on judgments, and they do not. According to the Contrast Model, expectancies are less likely and analytic attending with beat-counting strategies are more likely in judging these tunes. As a result, tendencies to correctly judge two events as identical in total duration are reinforced. Of course, in other cases, if events differ in number of beat periods, then tendencies to over- or underestimate duration will emerge (see Boltz, 1985). Thus, both attending modes may have been used in the present studies, a possibility supported by the fact that greater variability also appears in judgments involving pairs with moderate to low coherence.

Finally, although the overall emphasis here has been on  $DE_i$ , in principle the  $DM_i$  parameter lends itself to several interpretations. As with  $DE_i$ , this parameter can reflect the way someone uses time at any of several different levels (e.g., chunks or groups of beats). That is, people will not always rely on beat counting; they may analytically attend over other levels, levels that correspond to things other than beats. Part 1 indicates that amount of information (defined in various ways) does seem to bias time judgments in certain contexts. The present study and the Contrast Model suggest some reasons for these seemingly conflicting findings on the filled duration effect: The way information fills an interval to mark different time levels and the way a person attends to this information can affect time estimates.

In short, memory-oriented approaches to duration judgment must be supplemented, and the Contrast Model illustrates one way to do this. It incorporates some version of a memory-based model through the concept of analytic attending. It does so by

expressing the idea that effortful memory activities and expectancies are stimulated by different contexts, ones at opposite ends of the structural coherence continuum.

#### *Summary of Part 4*

A reciprocal pattern of findings emerged from two experiments involving comparative duration judgments (Experiment 1) and timed productions (Experiment 2) in which attending mode and temporal coherence were varied. Experiment 1 findings support the Contrast Model's proposition that duration judgments are a function of weighting attending mode parameters (temporal contrast, mnemonic differences) by degree of coherence. Experiment 2 findings suggest that contrast itself depends on temporal expectancies about musical endings.

#### Part 5: Conclusions

In this article we derived illustrative models of time judgment and production from a general theoretical framework in which both attending mode and event coherence figure. These models have some potential for explaining ways in which rhythmical attenders who use relative time nonetheless estimate intervals of time in different tasks. They also address divergent findings surrounding the filled interval effect. In Part 5, we conclude with general observations about the theoretical framework that relate to issues of generality, biological bases of attending, and philosophical views of time.

#### *Theoretical Generality: Is a Rhythmic Approach Warranted?*

The assumption of inherent rhythmicity in both the environment and the attender is a basic one. Because its claim to generality can be questioned, let us be clear about what it does and does not mean. It means that temporal patterning exists in environmental energy patterns and in the structure of attending over time. It does not mean that these paterings are immediately obvious, necessarily simple, or rigidly fixed. Indeed, the main point of our earlier discussion of environmental structure was to illustrate that formalization of time structure is a challenging goal, one demanding careful identification of functional time markers, marker couplings/strength, and various time transformations. The very success of another, quite different approach to rhythmicity, fourier analysis, should discourage ideas that inherent periodicities of events and attending activities can be superficially discerned.

At the level of the Contrast Model, the relative impact of simple and complex rhythmicities is expressed through attentional weights that assume a continuum of dynamic environmental complexity. It posits conditions in which evidence of simple rhythmicities will appear, and clearly these are not ubiquitous. Temporally incoherent events do exist. Our claim is simply that ready attunements to such complex events will fail in particular and sometimes meaningful (e.g., contrast) ways, given the nature of rhythmic attending. In this sense, the approach has greater generality than those time estimation models cited in Part 1 that focus largely on events with low temporal coherence. Interestingly, it achieves this by reconceiving the elusive con-

struct of complexity in rhythmic terms and suggesting a range of dynamic complexity based on temporal coherence.

There is no doubt that people are often forced to judge the duration of complex events with arbitrary time spans. The way this might occur has been illustrated in Parts 3 and 4 using musical events. Perhaps the Contrast Model will not generalize to comparable, nonmusical situations involving skilled attenders. Although this restriction remains a possibility, we think some formalization of constructs such as expectancy, contrast, and analytic attending will be useful for future models. Ultimately, the value of a fundamental assumption is found in its potential for stimulating explanations of diverse psychological phenomena.

### *Biological Clocks Versus a Rhythmic Attending Perspective*

Clearly, the basic assumptions about inherent rhythmicity, when considered with respect to the attender, imply a biological foundation for attentements. The bias toward a biological understanding is shared by others who provide additional support for the inherent rhythmicity of animate beings (e.g., Kelso et al., 1985).

With regard to theories of time estimation, however, biologically based theories have become unfashionable. Ornstein (1969) was most explicit in pointing to limitations of biological chronometers: "Although psychologists . . . have attempted to study time experience there has been a maze of postulated possible mechanisms proposed for time. One major reason . . . [is that] . . . time has been treated as . . . sensory process" (p. 34).

Ornstein (1969) correctly criticized earlier approaches for their disagreement on a specific time base. However, his alternative storage size hypothesis rests on equally confusing guidelines concerning complexity, ones that do not reliably predict time judgments. Nevertheless, because our approach envisions a biological basis, it is useful to distinguish the way time is conceived here from the way it is viewed in models of a strict biological clock.

Strictly speaking, a clock is a timekeeper that abstracts an absolute interval from the time dimension to serve as a fixed interval time base. Its appeal in time estimation often involves arousal effects that are handily accommodated by assuming that the interval time base shortens or lengthens with, for example, drugs or boredom. Nevertheless, as a biological metric the clock concept reinforces the idea that time is independent of the structure of particular events because the clock interval indiscriminately measures both arbitrary and nonarbitrary time periods. For example, a 5-min interval is arbitrary when defined only by clock numbers, but not when it coincides with the beginning and end of a bird's song, a child's game, or a conversation. Yet such distinctions are irrelevant for a clock model that exclusively counts elapsed time. Furthermore, the choice of a single time interval as a clock itself is a problem, as Ornstein (1969) noted. For a clock model to be sensitive to nested time periods, there must be postulates referring to clocks within clocks, as well as a master clock that somehow controls and monitors their interactions. Finally, each clock must be intelligent: it must "know" how to identify its salient markers, those that correspond to meaningful beginnings and ends of its particular intervals. Most psychological versions of clock models be-

gin as simple ones that do not directly reflect either relative time (e.g., time ratios, time contrasts) or structural markings. Consequently, they become complicated and unparsimonious when explanations of special goal-directed temporal sensitivities are required. For example, the roles of expectancy violations and temporal contrast as well as those of flexible and/or effortful attending remain difficult to incorporate into clock models. Instead, clock models encourage a search for a central timing mechanism (master clock) that monitors absolute intervals and for psychological time limits that are interval in nature. These are dubious undertakings in light of current data (e.g., Jones, 1976; Kolers & Brewster, 1985; van Noorden, 1971) and the arguments supplied here on the relational nature of timing and temporal limits (e.g., dominance region; see Footnote 12).

In our view, the biological basis for responses to event time takes the form of attunement rhythms that selectively entrain, that shift over nested levels, and eventually are shaped by the event itself. Attending is based directly on relative time, and this results in temporal flexibility in which neither the referent time base nor focal attending levels are rigidly fixed. Thus, various arousal effects can also be addressed within this theory via referent level shifts. But the major advantage of the current approach over strict clock models is found in its predictions concerning focal attentings relative to some referent level. For instance, expectancies and temporal contrast derive directly from analyses of dynamic interactions of an attender with patternings of environmental structure. In effect, the structure of a temporally coherent event can function as a natural timekeeper for the attender. Instead of a clock-timed world, the attender responds to a dynamic event-timed world in which time judgments depend on an event's characteristic timing and on how its structure confirms or disconfirms some expected course. Of course, this sort of attending is less likely with incoherent events. In this case people may resort to analytic attending or an even more practical strategy: They may check their watches. In fact, we predict that people will refer to external timekeepers more frequently in relatively unstructured situations.

In proposing that dynamic events afford different, biologically based attentional perspectives (e.g., future-oriented vs. analytic), we also raise other issues with regard to attending. Our experimental focus has been on time estimation, but the conceptual framework has broader implications, ones that distinguish it from other approaches to attention that are less concerned with attending over time. As a consequence, this approach has greater applicability to patterns of speech, music, dance, and so on, in which the concept of attunement levels may be fruitful (Jones, 1986). It introduces the possibility that people rely directly on event structure to achieve different temporal perspectives for different goals. These perspectives can result in different estimates of relative duration in time estimation. Dynamic complexity enters in to qualify how one uses event structure to achieve the requisite perspective. But in other situations and goals, attending perspectives and complexity may yield alternative models. Furthermore, by linking different attentional perspectives to distinctive time levels within each event, this approach avoids the homunculus problem that also troubles certain attentional theories (see Johnston & Dark, 1986). Finally then, this approach incorporates responsiveness to goals and event functions that is not evident in other models of attention and memory, including those adapted to time esti-

mation that were mentioned in Part 1. In sum, the construct of attending levels suggests a way in which environmental information can selectively control attending in time and also afford potential referent levels from which other aspects of the environment can be attended.

### Some Philosophical Implications

Finally, some of our distinctions inevitably invite comparison with Bergson's (1944, 1950) notion of "lived" time versus "measured" time. In reacting to Kant's (1900) claim that time is an internal intuition that allows separations of object representations on an infinite time line (i.e., "in relations of time"; Kant, 1900, pp. 28–31), Bergson (1950) argued that

Kant's great mistake was to view time as a homogeneous medium. He did not notice that real duration is made up of moments inside one another, and that when it seems to assume the form of a homogeneous whole, it is because it gets expressed in space. (p. 232)

For Bergson, nonhomogeneity characterizes experienced time because things experienced first always affect (interpenetrate) things experienced subsequently in time; this irreversibility describes time intervals as we live through them. It is not captured by the expression of a time line (Jones, 1987b).

Clearly, our concepts of temporal nesting, attunements, expectancy, and contrast capture some Bergsonian aspects of lived time. But Bergson and Kant present essentially phenomenological approaches in that both assume that only privately experienced time is real, because the time of external objects is constructed. And for both, this phenomenal time is essentially interval in character, and operates in an absolute fashion (see, e.g., Bergson's, 1965, reply to Einstein). By contrast, ours is not a phenomenological account. Nor do we conceive of time intervals in absolute terms. We assume that both the environment and its organisms exhibit temporal patternings that are reciprocally related and shared during attending activities. The internal time structure of attunements and expectancies is inextricably tied to the external time structure of transformationally related event changes. In order to interact, both must exist.

### References

- Abel, S. (1972a). Discrimination of temporal gaps. *Journal of the Acoustical Society of America*, 52, 519–524.
- Abel, S. (1972b). Duration discrimination of noise and tone bursts. *Journal of the Acoustical Society of America*, 51, 1219–1223.
- Adams, R. (1977). Intervening stimulus effects of category judgments of duration. *Perception and Psychophysics*, 21, 527–534.
- Allan, L. (1979). The perception of time. *Perception and Psychophysics*, 26, 340–354.
- Apel, W. (1972). *Harvard dictionary of music* (2nd ed). Cambridge, MA: Belknap Press of Harvard University Press.
- Avant, L., & Lyman, P. (1975). Stimulus familiarity modifies perceived duration in prerecognition visual processing. *Journal of Experimental Psychology: Human Perception and Performance*, 1, 205–213.
- Baker, C. H. (1962). On temporal extrapolation. *Canadian Journal of Psychology*, 16, 37–41.
- Barclay, C. D., Cutting, J. E., & Kozlowski, L. T. (1978). Temporal and spatial factors in gait perception that influence gender recognition. *Perception and Psychophysics*, 23, 145–152.
- Bartlett, N. R., & Bartlett, S. C. (1959). Synchronization of a motor response with an anticipated sensory event. *Psychological Review*, 66, 203–218.
- Benjamin, W. E. (1984). A theory of musical meter. *Music Perception*, 1, 355–413.
- Bergson, H. (1944). *Creative evolution* (A. Mitchell, Trans.). Toronto: Random House, Canada Limited.
- Bergson, H. (1950). *Time and freewill* (F. L. Pogson, Trans.). London: George Allen & Unwin.
- Bergson, H. (1965). *Duration and simultaneity* (L. Jacobson, Trans.). New York: Bobbs-Merrill.
- Berry, W. (1976). *Structural functions in music*. Englewood Cliffs, NJ: Prentice Hall.
- Block, R. (1974). Memory and the experience of duration in retrospect. *Memory and Cognition*, 2, 153–160.
- Block, R. (1978). Remembered duration: Effects of event sequence complexity. *Memory and Cognition*, 6, 320–326.
- Block, R. (1985). Contextual coding in memory: Studies of remembered duration. In J. Michon & J. Jackson-Roy (Eds.), *Time, mind, and behavior* (pp. 169–178). Heidelberg, FRG: Springer-Verlag.
- Block, R. (in press). Experiencing and remembering time: Affordances, context, and cognition. In I. Levin & D. Zakay (Eds.), *Time and human cognition: A life-span perspective*. Amsterdam: North Holland.
- Block, R., George, E. J., & Reed, M. A. (1980). A watched pot sometimes boils: A study of duration experience. *Acta Psychologica*, 46, 81–94.
- Block, R., & Reed, M. (1978). Remembered duration: Evidence for a contextual-change hypothesis. *Journal of Experimental Psychology: Human Learning and Memory*, 4, 656–665.
- Bolinger, D. L. (1958). A theory of pitch accent. *Word*, 14, 109–149.
- Boltz, M. (1985). *An expectancy model of judged duration: An ecological perspective*. Unpublished doctoral dissertation, Ohio State University Library.
- Boltz, M., & Jones, M. R. (1986). Does rule recursion make melodies easier to reproduce? If not, what does? *Cognitive Psychology*, 18, 389–431.
- Brown, H. (1988). The interplay of set content and temporal context in a functional theory of tonality perception. *Music Perception*, 5, 219–250.
- Brown, H., & Butler, D. (1981). Diatonic trichord as minimal tonal cue-cells. *In Theory Only*, 5, 39–55.
- Brown, S. W. (1985). Time perception and attention: The effects of prospective versus retrospective paradigms and task demands on perceived duration. *Perception and Psychophysics*, 38, 115–124.
- Burnside, W. (1971). Judgment of short time intervals while performing mathematical tasks. *Perception and Psychophysics*, 9, 404–406.
- Butler, D., & Brown, H. (1984). Tonal structure versus function: Studies of the recognition of harmonic motion. *Music Perception*, 2, 6–24.
- Carlsoo, S. (1972). *How man moves*. (W. P. Michael, Trans.). New York: Distributed by Cran, Russak, & Co.
- Clarke, E. F. (1984). *Structure and expression in the rhythm of piano performance*. Unpublished doctoral dissertation, University of Exeter, Exeter, Devon, England.
- Clarke, E. F. (1985a). Some aspects of rhythm and expressions in performances of Erik Satie's "Gnossienne No. 5." *Music Perception*, 2, 299–328.
- Clarke, E. F. (1985b). Structure and expression in rhythmic performance. In P. Howell, I. Cross, & R. West (Eds.), *Musical structure and cognition* (pp. 209–236). London: Academic Press.
- Clynes, M. (1978). *Sentics: The touch of emotions*. Garden City, NJ: Anchor Press.
- Coxeter, H. S. M., & Greitzer, S. L. (1967). *Geometry revisited*. New York: Random House.
- Condon, W. S., & Sander, L. W. (1974). Synchrony demonstrated between movements of the neonate and adult speech. *Child Development*, 45, 456–462.

- Cone, E. (1968). *Musical form and musical performance*. New York: Norton.
- Craik, R., & Lockhart, R. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 11, 671-684.
- Cutler, A., & Isard, S. D. (1980). The production of prosody. In B. Butterworth (Ed.), *Language production: Vol. 1. Speech and talk* (pp. 245-269). New York: Academic Press.
- Cutting, J. (1982). Two ecological perspectives: Gibson vs. Shaw and Turvey. *American Journal of Psychology*, 95, 199-222.
- Deutsch, D., & Feroe, J. (1981). The internal representation of pitch sequences in tonal music. *Psychological Review*, 18, 503-522.
- Devane, J. (1974). Word characteristics and judged duration for two response sequences. *Perception and Motor Skills*, 38, 525-526.
- Divenyi, P., & Danner, W. (1977). Discrimination of time intervals marked by brief acoustic pulses of various intensities and spectra. *Perception and Psychophysics*, 21, 125-142.
- Divenyi, P., & Sachs, (1978). Discrimination of time intervals bounded by tone bursts. *Perception and Psychophysics*, 24, 429-436.
- Dowling, W. J., & Harwood, D. L. (1985). *Music cognition*. New York: Academic Press.
- Essens, P. J. (1986). Hierarchical organization of temporal patterns. *Perception and Psychophysics*, 40, 69-73.
- Essens, P. J., & Povel, D. J. (1985). Metrical and nonmetrical representations of temporal patterns. *Perception and Psychophysics*, 37, 1-7.
- Fowler, C. (1983). Converging sources of evidence on spoken and perceived rhythms of speech: Cyclic production of vowels in monosyllabic stress feet. *Journal of Experimental Psychology: General*, 112, 386-412.
- Fraisse, P. (1956). *Les structures rythmiques* [Rhythmic structures]. Louvain, France: Editions Universitaires.
- Fraisse, P. (1964). *The psychology of time*. London: Eyre & Spottiswoode.
- Fraisse, P. (1982). Rhythm and tempo. In D. Deutsch (Ed.), *The psychology of music* (pp. 149-180). New York: Academic Press.
- 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, Sweden: Almqvist & Wiksell.
- Freyd, J. J. (1987). Dynamic mental representations. *Psychological Review*, 94, 427-438.
- Fry, D. B. (1958). Experiments in the perception of stress. *Language and Speech*, 1, 126-152.
- Gabrielsson, A. (1974). Performance of rhythm patterns. *Scandinavian Journal of Psychology*, 15, 63-74.
- Gabrielsson, A. (1985). Interplay between analysis and synthesis in studies of music performance and music experience. *Music Perception*, 3, 59-68.
- Gabrielsson, A., Bengtsson, I., & Gabrielsson, B. (1983). Performance of musical rhythm in 3/4 and 6/8 meter. *Scandinavian Journal of Psychology*, 24, 193-213.
- Gentner, D. R. (1987). Timing of skilled motor performance: Tests of the proportional duration model. *Psychological Review*, 94, 255-276.
- Getty, D. (1975). Discrimination of short temporal intervals: A comparison of two models. *Perception and Psychophysics*, 18, 1-8.
- Getty, D. (1976). Counting processes in human timing. *Perception and Psychophysics*, 20, 191-197.
- Gibson, J. J. (1966). *The senses considered as perceptual systems*. Boston: Houghton Mifflin.
- Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.
- Gray, J. (1968). *Animal locomotion*. New York: Norton.
- Hahn, J., & Jones, M. R. (1981). Invariants in auditory frequency. *Scandinavian Journal of Psychology*, 92, 627-651.
- Halpern, A., & Darwin, L. (1982). Duration discrimination in a series of rhythmic events. *Perception and Psychophysics*, 31, 86-89.
- Hamilton, P., & Hockey, R. (1974). Active selection of items to be remembered. *Cognitive Psychology*, 6, 61-83.
- Handel, S. (1984). Using polyrhythms to study rhythm. *Music Perception*, 1, 465-484.
- Handel, S., & Oshinsky, J. S. (1981). The meter of syncopated auditory polyrhythms. *Perception and Psychophysics*, 30, 1-9.
- Hicks, R., Miller, G., & Kinsborne, M. (1976). Prospective and retrospective judgments of time as a function of amount of information processed. *American Journal of Psychology*, 89, 719-730.
- Hirsh, I. J., Bilger, R. C., & Deatherage, B. H. (1969). The effect of auditory and visual background on apparent duration. *American Journal of Psychology*, 69, 561-574.
- Hogan, H. (1975). Time perception and stimulus preference as a function of stimulus complexity. *Journal of Personality and Social Psychology*, 31, 32-35.
- Inman, V. T. (1966). Human locomotion. *Canadian Medical Association Journal*, 94, 1047-1054.
- Jackson, J. L. (1985). Is the processing of temporal information automatic or controlled? In J. A. Michon & J. L. Jackson (Eds.), *Time, mind and behavior* (pp. 179-190). Berlin: Springer-Verlag.
- Jagacinski, R., Marshburn, E., Klapp, S. T., & Jones, M. R. (1988). Tests of parallel versus integrated structure in polyrhythmic tapping. *Journal of Motor Behavior*, 20, 416-442.
- Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception and Psychophysics*, 14, 201-211.
- Johansson, G. (1975). Visual motor perception. *Scientific American*, 232, 76-89.
- Johnson, W. W. (1986). *Studies in motion extrapolation*. Unpublished doctoral dissertation, Ohio State University, Columbus.
- Johnston, W. A., & Dark, V. J. (1986). Selective attention. In M. R. Rosenzweig & L. W. Porter (Eds.), *Annual review of psychology* (Vol. 37, pp. 43-75). Palo Alto, CA: Annual Reviews.
- 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 and Psychophysics*, 30, 492-504.
- Jones, M. R. (1981c). Music as a stimulus for psychological motion: Part 1. Some determinants of expectancies. *Psychomusicology*, 1, 34-51.
- Jones, M. R. (1982). Music as a stimulus for psychological motion: Part 2. An expectancy model. *Psychomusicology*, 2, 1-13.
- Jones, M. R. (1985). Structural organization of events in time: A review. In J. A. Michon & J. L. Jackson (Eds.), *Time, mind, and behavior* (pp. 192-214). Heidelberg, FRG: Springer-Verlag.
- Jones, M. R. (1986). Attentional rhythmicity in human perception. In J. R. Evans & M. Clynes (Eds.), *Rhythm in psychological, linguistic, and music processes* (pp. 13-40). Springfield, IL: Charles C Thomas.
- Jones, M. R. (1987a). Perspectives on musical time. In A. Gabrielsson (Ed.), *Action and perception in rhythm and music* (pp. 153-175). Stockholm, Sweden: Royal Swedish Academy of Music.
- Jones, M. R. (1987b). Some thoughts on the relevance of Bergson to contemporary psychology. In A. C. Papanicolaou & P. A. Y. Gunter (Eds.), *Bergson and modern thought: Towards a unified science* (pp. 250-270). New York: Harwood Academic Publishers.
- Jones, M. R., & Boltz, M., & Kidd, G. (1982). Controlled attending as a function of melodic and temporal context. *Perception and Psychophysics*, 32, 211-218.
- Jones, M. R., Kidd, G., & Wetzel, R. (1981). Evidence for rhythmic attention. *Journal of Experimental Psychology: Human Perception and Performance*, 7, 1059-1073.
- Kant, I. (1900). *Critique of pure reason*. (J. M. D. Meiklejohn, Trans.). New York: Colonial Press.
- Keele, S. W., & Hawkins, H. (1982). Explorations of individual differ-

- ences relevant to high level skill. *Journal of Motor Behavior*, 14, 3–23.
- Keele, S. W., Pokorny, R. A., Corcos, D. M., & Ivry, R. (1985). Do perception and motor perception share common timing mechanisms: A correlation analysis. *Acta Psychologica*, 60, 173–191.
- Kelso, J. A. S., Holt, K. G., Rubin, P., & Kugler, P. (1981). Patterns of human interlimb coordination emerge from the properties of nonlinear, limit cycle oscillatory processes: Theory and data. *Journal of Motor Behavior*, 13, 226–261.
- Kidd, G. (1984). *The perception of tonal relationships in music*. Unpublished doctoral dissertation, The Ohio State University Library.
- Kidd, G. (in press). Articulatory-rate context effects in phoneme identification. *Journal of Experimental Psychology: Human Perception and Performance*.
- Klapp, S. T. (1979). Doing two things at once: The role of temporal compatibility. *Memory and Cognition*, 7, 375–381.
- Klapp, S. T. (1981). Temporal compatibility in dual motor tasks: II. Simultaneous articulation and hand movement. *Memory and Cognition*, 9, 398–401.
- Klapp, S. T., Hill, M., Tyler, J., Martin, Z., Jagacinski, R., & Jones, M. R. (1985). On marching to two different drummers: Perceptual aspects of the difficulties. *Journal of Experimental Psychology: Human Perception and Performance*, 11, 814–827.
- Kolata, G. (1985). Finding biological echos in fetuses. *Science*, 230, 929–930.
- Kolers, P. A., & Brewster, J. M. (1985). Rhythms and responses. *Journal of Experimental Psychology: Human Perception and Performance*, 11, 150–167.
- Kozhevnikov, V., & Chistovich, L. (1965). *Speech: Articulation and perception* (Rep. No. JPRS 30543). Washington, DC: Joint Publications Research Service.
- Kramer, J. (1982). Beginnings and endings in Western art music. *Canadian University Music Review*, 3, 1–14.
- Kristofferson, A. B. (1976). Low-variance stimulus–response latencies: Deterministic interval delays? *Perception and Psychophysics*, 20, 89–100.
- Kristofferson, A. B. (1977). A real-time criterion theory of duration discrimination. *Perception and Psychophysics*, 21, 105–117.
- Kristofferson, A. B. (1980). A quantal step function in duration discrimination. *Perception and Psychophysics*, 27, 300–306.
- Kristofferson, A. B. (1984). Quantal and deterministic timing in human duration discrimination. *Annals of the New York Academy of Sciences*, 423, 3–15.
- Kronman, U., & Sundberg, J. (1987). Is the musical ritard an allusion to physical motion? In A. Gabriellson (Ed.), *Action and perception in rhythm and music* (pp. 57–68). Stockholm, Sweden: Royal Swedish Academy of Music.
- Krumhansl, C. L. (1979). The psychological representation of musical pitch in a tonal content. *Cognitive Psychology*, 11, 346–374.
- Krumhansl, C. L., & Kessler, E. J. (1982). Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychological Review*, 89, 334–368.
- Ladd, D. R. (1986). Intonational phrasing: The case for recursive prosodic structure. *Phonology Yearbook*, 3, 311–340.
- Laws, K. (1985). Physics and dance. *American Scientist*, 73, 426–432.
- Lee, C. S. (1985). The rhythmic interpretation of simple musical sequences: Towards a perceptual model. In P. Howell, I. Cross, & R. West (Eds.), *Musical structure and cognition* (pp. 53–69). London, Academic Press.
- Lee, D. N. (1980). The optic flow field: The foundation of vision. *Philosophical Transactions of the Royal Society of London*, 290, 169–179.
- Leeuwenberg, E. (1969). Quantitative specification of information in sequential patterns. *Psychological Review*, 76, 216–220.
- Lehiste, I. (1975, August). *The perception of duration within sequences of four intervals*. Paper presented at the 8th International Congress of Phonetic Sciences, Leeds, England.
- Lehiste, I. (1976). Suprasegmental features of speech. In N. J. Lass (Ed.), *Contemporary issues in experimental phonetics* (pp. 225–239). New York: Academic Press.
- Lehiste, I. (1977). Isochrony revisited. *Journal of Phonetics*, 5, 253–263.
- Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music*. Cambridge, MA: MIT Press.
- Lieberman, M., & Prince, A. (1977). On stress and linguistic rhythm. *Linguistic Inquiry*, 8, 249–336.
- Longuet-Higgins, H. C., & Lee, C. S. (1982). The perception of musical rhythms. *Perception*, 11, 115–128.
- Longuet-Higgins, H. C., & Lee, C. S. (1984). The rhythmic interpretation of monophonic music. *Music Perception*, 1, 424–441.
- Marquardt, D. W. (1963). An algorithm for least-squares estimation of nonlinear parameters. *SIAM Journal of Applied Mathematics*, 11, 431–441.
- Martin, J. (1972). Rhythmic (hierarchical) versus serial structure in speech and other behavior. *Psychological Review*, 79, 487–509.
- Maser, D. J. (1983). *Perceiving and counting sounds: The roles of pitch distance, rhythm and grouping*. Unpublished doctoral dissertation, Ohio State University.
- Michon, J. A., & Jackson, J. L. (1985). The psychology of time. In J. A. Michon, & J. L. Jackson (Eds.), *Time, mind, and behavior* (pp. 2–17). Heidelberg, FRG: Springer-Verlag.
- Monahan, C. B., & Carterette, E. C. (1985). Pitch and duration as determinants of musical space. *Music Perception*, 3, 1–32.
- Navon, D., & Gopher, D. (1979). On the economy of the human processing system. *Psychological Review*, 86, 214–255.
- Neisser, U., & Becklen, R. (1975). Selective looking: Attending to visually specified events. *Cognitive Psychology*, 7, 480–494.
- Newson, D., Hairfield, J., Bloomingdale, J., & Cutino, S. (1987). The structure of action and interaction. *Social Cognition*, 5, 191–237.
- Ornstein, R. (1969). *On the experience of time*. Baltimore, MD: Penguin Books.
- Paccia-Cooper, J., & Cooper, W. E. (1981). The processing of phrase structures in speech production. In P. D. Eimas & J. L. Miller (Eds.), *Perspectives on the study of speech* (pp. 311–336). Hillsdale, NJ: Erlbaum.
- Penner, M. (1976). The effect of marker variability on the discrimination of temporal intervals. *Perception and Psychophysics*, 19, 466–469.
- Pickett, J. M. (1980). *The sounds of speech communication*. Baltimore, MD: University Park Press.
- Pierson, K. (1976). Control of walking. *Scientific American*, 235, 72–86.
- Pike, K. (1945). *The intonation of American English*. Ann Arbor: University of Michigan Press.
- Piston, W. (1978). *Harmony* (4th ed., revised and expanded by Mark De Voto). New York: Norton.
- Povel, D. J. (1977). Temporal structure of performed music: Some preliminary observations. *Acta Psychologica*, 41, 309–320.
- Povel, D. J. (1981). Internal representations of simple temporal patterns. *Journal of Experimental Psychology: Human Perception and Performance*, 7, 3–18.
- Povel, D. J., & Essens, P. (1985). Perception of temporal patterns. *Music Perception*, 2, 411–440.
- Poynter, W. (1983). Duration judgment and the segmentation of experience. *Memory and Cognition*, 11, 77–82.
- Poynter, W., & Homa, D. (1983). Duration judgment and the experience of change. *Perception and Psychophysics*, 33, 548–560.
- Reber, A. S., Kassin, S. M., Lewis, S., & Cantor, G. (1980). On the relationship between implicit and explicit modes in the learning of a complex rule structure. *Journal of Experimental Psychology: Human Learning and Memory*, 5, 492–502.
- Scherer, K. (1979). Acoustic concomitants of emotional dimensions: Judging affect from synthesized tone sequences. In S. Weitz (Ed.),

- Nonverbal communication* (pp. 249–253). New York: Oxford University Press.
- Schiffman, H., & Bobko, D. (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. (1982). Rhythm and timing in skill. *Psychological Review*, *89*, 109–122.
- Shaffer, L. H., Clarke, E., & Todd, N. (1985). Meter and rhythm in piano playing. *Memory and Cognition*, *20*, 61–77.
- Shaw, L., McIntyre, M., & Mace, W. (1974). The role of symmetry in event perception. In R. MacLeod & H. Pick (Eds.), *Perception: Essays in honor of J. J. Gibson* (pp. 276–310). Ithaca, NY: Cornell University Press.
- Shepard, R. (1981). Psychophysical complementarity. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization* (pp. 279–341). Hillsdale, NJ: Erlbaum.
- Shepard, R. (1984). Ecological constraints on internal representation: Resonant kinematics of perceiving, imagining, thinking, and dreaming. *Psychological Review*, *91*, 417–447.
- 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 and Performance*, *102*, 250–255.
- Siegmán, A. (1978). The telltale voice: Nonverbal messages of verbal communication. In A. Siegmán & S. Feldstein (Eds.), *Nonverbal behavior and communication* (pp. 183–243). Hillsdale, NJ: Erlbaum.
- Simon, H. (1972). Complexity and the representation of patterned sequences of symbols. *Psychological Review*, *79*, 369–382.
- Sloboda, J. A. (1983). The communication of musical meter in piano performance. *Quarterly Journal of Experimental Psychology*, *35*, 377–390.
- Sloboda, J. A. (1985). *The musical mind*. Fair Lawn, NJ: Oxford University Press.
- Steedman, M. J. (1977). The perception of musical rhythm and meter. *Perception*, *6*, 555–569.
- Sternberg, S., Knoll, R., & Zukofsky, P. (1978). Timing by skilled musicians. In D. Deutsch (Ed.), *The psychology of music* (pp. 182–239). New York: Academic Press.
- Thomas, E., & Cantor, N. (1978). Interdependence between the processing of temporal and non-temporal information. In J. Requin (Ed.), *Attention and performance VII* (pp. 43–62). Hillsdale, NJ: Erlbaum.
- Thomas, E., & Weaver, W. (1975). Cognitive processing and time perception. *Perception and Psychophysics*, *17*, 363–367.
- Todd, N. (1985). A model of expressive timing in tonal music. *Music Perception*, *3*, 33–58.
- Townsend, J. T., & Ashby, F. G. (1983). *Stochastic modeling of elementary psychological processes*. New York: Cambridge University Press.
- Tuller, B., Kelso, J. A. S., & Harris, K. S. (1983). Converging evidence for the role of relative timing in speech. *Journal of Experimental Psychology: Human Perception and Performance*, *9*, 829–833.
- Turvey, M. T., & Carello, C. (1981). Cognition: The view from ecological realism. *Cognition*, *10*, 313–321.
- Underwood, G., & Swain, R. (1973). Selectivity of attention and the perception of duration. *Perception*, *2*, 101–105.
- van Noorden, L. (1971). Rhythmic fission as a function of tone rate. *IPO Annual Progress Report*, *6*, 9–12.
- Van Vreden, W., Boltz, M., & Jones, M. R. (1989). *Expected ending times and duration judgments*. Manuscript submitted for publication.
- Vos, P., & Handel, S. (1987). Playing triplets: Facts and preferences. In A. Gabriellson (Ed.), *Action and perception in rhythm and music* (pp. 35–47). Stockholm, Sweden: Royal Swedish Academy of Music.
- Vos, P. G., Leeuwenberg, E. L., & Collard, R. F. (1978). *What melody tells about meter in music* (Report No. 78 FU 03). Nijmegen, The Netherlands: University of Nijmegen.
- Wiemann, J., & Knapp, M. (1975). Turn-taking in conversation. *Journal of Communication*, *25*, 75–92.
- Winter, D. A. (1983). Biomechanical motor patterns in normal walking. *Journal of Motor Behavior*, *15*, 302–330.
- Woods, D., Sorkin, R., & Boggs, G. (1979). Stimulus context and duration discrimination. *Perception and Psychophysics*, *26*, 126–132.
- Yeston, M. (1976). *The stratification of musical rhythm*. New Haven, CT: Yale University Press.

## Appendix

### Judging Objectively Unequal Time Intervals

Contrast Model predictions apply to mean accuracy and variability (discrimination) of duration judgments about events that do differ in duration. Let this difference for pair  $i$  be denoted by  $\Delta t_i$ . Following convention in this field, let  $\Delta T$  refer to the total duration of the standard and  $\Delta T + \Delta t_i$  that of the comparison. In most settings,  $\Delta t_i$  varies in magnitude and direction over the set of (standard plus comparison) pairs based on a constant standard  $\Delta T$  value.

Contrast Model applications are direct: (a)  $DE_i$ , based on the assumption that expectancies associated with the standard ( $\Delta T$ ), are carried over to the comparison ( $\Delta T + \Delta t_i$ );  $DE_i$  provides one estimate of each true time difference,  $\Delta t_i$ . (b)  $DM_i$ , based on analytic attending and focal level differences between standard and comparison, provides another estimate of  $\Delta t_i$  for the same  $i$ . (c) The two estimates,  $DE_i$  and  $DM_i$ , are weighted, respectively, following Equations 5–7, by  $W_i$  to determine the likelihood of future-oriented and analytic attending.

Resulting predictions indicate that nontemporal information filling two different time intervals can enhance or degrade duration judgment accuracy and discrimination depending on event structure and attending mode. These predictions are outlined later in the Appendix along with special adaptations for determining model parameters when empty time intervals are involved.

#### Mean Judgment Predictions

Equation 5 applies unchanged, with model parameter estimates given by Equations 8–10 for two filled time intervals. The focal levels of relevance here involve total duration (of standard and comparison); these determine  $DE_i$  and  $DM_i$ . For  $DE_i$ , expected time is determined by the total duration of the standard, whereas for  $DM_i$  the assumption of a common focal level supporting, for example, counting, remains unchanged (and, in this case, may be a correct assumption).

Equation 5 yields many predictions about judgment accuracy: (a) The magnitude and direction of a true time difference will systematically affect average (rated) judgments about this difference. (b) If either  $DE_i$  or  $DM_i$  yields a good estimate of  $\Delta t_i$ , then performance is improved when this estimate gets more weight ( $W_i$ ) and degraded when it does not. If both  $DE_i$  and  $DM_i$  are good estimates of  $\Delta t_i$ , then accuracy will be high and unaffected by variations in  $W_i$ . Other predictions relate to effects of attunement constraints and practice and can be deduced from Equation 5.

#### Empty Time Intervals

Silences are always judged within a larger temporal context. Thus, adaptations of Equation 5 to judgments of empty intervals require that

(a) the temporal coherence parameter reflects the context provided by the succession of interval markers, for example, in a surrounding phrase (prose, musical) or in an experimental session (psychophysical setting), (b)  $DE_i$  remains tied to expectancies about total event durations, whereas  $DM_i$  reflects differences due to counting, tapping, and other temporally based activities in which a person engages and uses to subdivide some empty time period, and (c) the referent level is determined by the most frequent, prominently marked time contextual period. Given this, predictions about empty interval judgments follow from Equations 5–6. As before, these suggest that attending mode and temporal context should affect judgment accuracy and bias.

The role of temporal contrast in the latter predictions is reminiscent of Kristofferson's (1977, 1984) real time criterion in his modeling of duration judgments of empty intervals. However, there are significant differences. One is that the real-time criterion model addresses only speeded responding in two choice judgment tasks (long and short). That task encourages and the model depicts responding only to the relative temporal order of expected versus observed ending times. The Contrast Model applies to comparative time interval estimates of expected versus observed ending times. A second is that neither attending modes and temporal coherence are found in the real-time criterion model. A dichotomous attending model has also been presented by Thomas and Cantor (1978) for empty and filled intervals. However, it applies only to durations less than 100 ms, and although the model does consider effects due to different amounts of information, it cannot address contrast effects arising from temporal expectancies. Conceptually, its parameters differ.

#### Variance and Time Discrimination Predictions

When Equation 7 (variance prediction) involves  $DE_i$  and  $DM_i$  parameters, both of which estimate the difference in total time of the two intervals (i.e.,  $\Delta t_i$ ), it can be used to make predictions about time-discrimination performance in a classic psychophysical context. That is, if  $\Delta t_i$  values about some  $\Delta T$  are small, temporal threshold issues arise. Conventionally, time discrimination thresholds (JNDs) are estimated from the standard deviation of observed judgments over the event pairs ( $i$ ) associated with a given standard interval,  $\Delta T$ , namely  $SD(\Delta T)$ . A critical design constraint is that the  $\Delta t_i$  values form a balanced set for each  $\Delta T$ , so that their sum (over standard/comparison pairs) is zero. Thus, to adapt Equation 7 to express  $SD(\Delta T)$ , parameters  $DE_i$ ,  $DM_i$ , and  $W_i$  must be averaged over appropriate event pairs  $i$  having a common  $\Delta T$ . Taking the square root:

$$JND = SD(\Delta T) = [a^2[W(1 - W)(DE - DM)^2 + \sigma^2]]^{1/2}. \quad (A1)$$

In this form, discriminability is a function of temporal coherence ( $W$ ), attentional estimates ( $DE$ ,  $DM$ ) of the average  $\Delta t_i$  value (i.e.,  $\Delta t = 0$ ), and an average base error variance,  $\sigma^2$  (for either filled and empty intervals given previous assumptions). Thus, the discriminability function relating JND to  $\Delta T$  can take different forms, with some at variance with Weber's Law for time discrimination. Equation A1 also predicts that nontemporal information filling two different intervals can affect their discriminability. Both predictions must be justified.

Weber's Law holds that a temporal JND is a constant proportion of  $\Delta T$ . It is supported if  $SD(\Delta T)$  is a constant proportion of  $\Delta T$  for all  $\Delta T$ , implying a linear relation between JND and  $\Delta T$ . Weber's Law is not compatible with Equation A1 because it does not predict discriminability to vary as a function of attending mode, coherence, context, practice, and so on. Nonetheless, Equation A1 is worth exploring, given that little empirical support exists for a simple version of Weber's Law (e.g., Allan, 1979; Divenyi & Danner, 1977; Getty, 1976; Halpern & Darwin, 1982; Kristofferson, 1976; Povel, 1977). Although temporal discriminability is good over a range of time intervals (50–2,000 ms), JNDs are not linear functions of  $\Delta T$  as predicted by Weber's Law. In fact, the discriminability function varies with different factors. Two illustrative cases deal respectively with skilled and unskilled attending.

*Skilled attending and JND invariance with  $\Delta T$ .* Threshold invariance is predicted by Equation A1 if  $(DE - DM)^2 = 0$ . This occurs if people reliably estimate the true time differences,  $\Delta t_i$ , regardless of attending mode. That is, because  $DE_i$  and  $DM_i$  each estimate  $\Delta t_i$  values, their sums (over  $i$ ), like those of  $\Delta t_i$  values themselves, will be zero if all are good estimates. Because skilled attending is flexible attending, either of the two attending modes may be used to arrive at the correct  $\Delta t_i$  estimates. If  $DE = DM = 0$ , Equation A1 becomes

$$SD(\Delta T) = (a^2\sigma^2)^{1/2}. \quad (A2)$$

The  $\sigma^2$  is the (average) variance of difference distributions and with skilled observers its constancy is reasonable. Therefore,  $a^2\sigma^2 = V_r$ , a variance component that is independent of stimulus durations (Getty, 1975).

Some predictions are as follows: (a) Asymptotically, thresholds are invariant with  $\Delta T$ . (b) Given a fixed (modal) referent period, dominance region constraints imply range limits to predicted threshold invariance (i.e., invariance obtains only within limiting  $\Delta T$  values, with JNDs increasing monotonically with larger  $\Delta T$  values, ones outside the dominance region range). (c) Variations in temporal coherence do not affect asymptotic performance (i.e.,  $W[1 - W]$  vanishes!). Although some support exists for these first two predictions (Kristofferson, 1980, 1984), no evaluations exist for the third prediction.

*Unskilled attending and monotonically increasing JNDs.* Following Equation A1, a monotonic function occurs if the term  $(DE - DM)^2 \neq 0$  and increases with  $\Delta T$ . This is predicted with unskilled judges and/or a wide range of  $\Delta T$  values (see attunement constraints, Part 2). In such cases, regardless of attending mode, unreliable estimates of  $\Delta t_i$  occur, leading to nonzero average estimates of both  $DE$  and  $DM$  (see mean judgments). In general,  $(DE - DM)^2$  should increase with  $\Delta T$  because (a)  $|\Delta t_i|$  increases proportionately with  $\Delta T$ , and (b) dominance region constraints imply that disparities between expected and observed times will increase as the latter exceed the dominance region.

Finally, Equation A1 leads to Getty's (1975) generalized version of Weber's Law, if  $W$  and  $\sigma^2$  are invariant with changes in  $\Delta T$ . Thus, if  $\Delta T$  are constrained in range and without systematic sequential patterning within a session, then we assume  $k^2 = W(1 - W)$  and  $V_r = a^2\sigma^2$ . Whenever  $a^2(DE - DM)^2$  approximates  $\Delta T^2$ , then Equation A1 yields

$$SD(\Delta T) = [k^2(\Delta T^2 + V_r)]^{1/2}. \quad (A3)$$

Predictions regarding skill level and  $\Delta T$  range have some support (Kristofferson, 1980, 1984). With respect to temporal coherence ( $W$ ), at least with discriminations of empty intervals, the relative timing of markers of differential prominence within an experimental session (context) significantly affects performance (e.g., Abel, 1972a, 1972b; Divenyi & Sachs, 1978; Penner, 1976). In sum, the predicted discriminability function takes different forms depending on the way Contrast Model parameters reflect attending skill, task constraints, instructions, and context.

#### Nontemporal Information

Within two different intervals, nontemporal information can affect discriminability according to Equation A1. Although this prediction creates apparent problems for the Contrast Model because effects of nontemporal information on judgments about equal time intervals have not been reliably reported for time discriminations (Allan, 1979; Woods, Sorkin, & Boggs, 1979), it is resolved in Equations A1–A2. These equations predict that such effects do not always occur and moreover, that when nontemporal information does affect threshold performance, it will either enhance or degrade performance depending upon attending mode.

Received May 31, 1988

Revision received October 11, 1988

Accepted October 25, 1988 ■