

Dynamic pattern structure in music: Recent theory and research

MARI RIESS JONES

*The Netherlands Institute for Advanced Study, Wassenaar, The Netherlands,
and The Ohio State University, Columbus, Ohio*

Recent theory and research addressed to dynamic pattern structure in music is reported. Dynamic pattern structure refers to the ways in which rhythm and melody combine; it is summarized here in terms of the construct of *joint accent structure*. Properties of joint accent structures involving accent couplings and time symmetries are used to address standard psychological issues of pattern similarity and pattern simplicity as they are realized in musical tasks. In particular, experimental work on melody recognition reveals the importance of formalizations concerning dynamic pattern similarities, and other research on reproductions of music-like patterns reveals the potential of a time-based approach to dynamic pattern simplicity.

Musical sound patterns are created and performed by some humans in order to communicate with others. Along with speech, music is preeminent among communicative sound patterns that our species generates. Like speech, music is complex and diversified in style, context, and specific function. Commercial jingles, spontaneous shower humming, and artful tonal music all receive the label "music." Yet even here there are clearly major differences in the structure and function of musical sound patterns: They serve different communicative goals and their structure is accordingly different.

I suggest that, in spite of this diversity, there are two aspects of musical sound patternings that cut across several different styles of Western music, and these have to do with fundamental ways we attend to and remember structure in time (Jones, 1985). One involves melodic theme recognition and the way in which a composer or performer evokes different degrees of familiarity in the reactions of listeners to recurrent themes. The other involves the listener's ongoing attending and the way in which a musical artist can effect predictability or, conversely, surprise through the artist's control of the listener's reactions to the dynamically unfolding musical structure.

With regard to theme recognition, many musical compositions "work" because their creators assume listeners will come to recognize an earlier melodic theme at certain points. Interestingly, timing variations that often accompany a theme play an important part in enabling the artist to introduce various subtle shadings, variations, and degrees of melodic familiarity in the listener. The idea of theme and variation, or more formally, invariance and transformation, has been around for a while and has been a fruitful concept in the psychological literature. It seems

especially well suited to the study of musical pattern structure (Jones, 1981a, 1981b, 1982; Welker, 1982). Here I focus on some recent research addressed to melodic theme recognition that suggests that functional pattern similarities that control theme recognition may be usefully conceived of in terms of dynamic pattern structure.

The second function of music relates to a musical artist's control over the listener's ongoing attentional response to a sound pattern. Clearly, it is critical that the performer (and composer) have a sense of how to keep a listener's attention relatively high, tracking the right melodic line, or at least that he/she know when it will be low. There is already some evidence that the surrounding regularities in pitch and time relations can guide attending selectively in time to particular notes (e.g., Jones, 1984; Jones, Boltz, & Kidd, 1982; Jones, Kidd, & Wetzel, 1981). Thus it is reasonable that an understanding of this dynamic response will come from analysis of structural constraints associated with melodic and temporal information in these sequences. That is, if, by exploitation of these constraints, an artist can ensure that a listener will indeed pay attention at certain times, then it is wise to insert more important thematic information at those times. Perhaps there are relatively simple dynamic constraints that form the basis for this control. I assume there are. Here I also describe some research addressed to dynamic pattern simplicity in music.

In sum, I report on recent theoretical developments and research that speak to two issues in music perception. The first is *dynamic pattern similarity* and the second is *dynamic pattern simplicity*. The term *dynamic* in both is intentional in that I mean to emphasize the integration of temporal relationships with other aspects of musical pattern structure. Here I focus specifically on relationships between melody and rhythmic structure.

My goals in this article are twofold. First, I consider theory and some recent research addressed to dynamic pattern similarity and theme recognition. Second, I discuss a new approach to the issue of dynamic pattern sim-

This research was sponsored in part by National Science Foundation Grant BNS-8204811 to the author and in part by The Netherlands Institute for Advanced Study, where the author was a fellow during the 1986-1987 academic year. Requests for reprints should be addressed to Mari Riess Jones, 404 C West Seventeenth Street, The Ohio State University, Columbus, OH 43210.

plicity, and some supporting research. My emphasis is on structural analysis, and I attempt to tie both issues together by suggesting some common aspects of listeners' responses to common dynamic constraints.

SOME STRUCTURAL DESCRIPTIONS OF PITCH (MELODY) AND TIME (METER AND RHYTHM) RELATIONSHIPS IN WESTERN MUSIC

In principle, the pitch relationships in a musical pattern contribute to its melody (and harmony), and the time relationships contribute to meter and rhythm (and tempo). I begin by introducing some elementary ways of describing these respective structures in monophonic music, and then attempt a more integrated structural analysis, one that ultimately combines pitch and time relations. The rationale of the integrated approach to pitch and time relationships is simple. It involves (1) identifying those pitch and time relationships that function, respectively, as melodic and temporal accents, and (2) specifying a higher order dynamic structure in terms of temporal properties that emerge from the joint analysis of melodic and temporal accents. Dynamic pattern similarity and dynamic pattern simplicity are influenced by properties of this joint accent structure.

Pitch Relationships, Melodies, and Melodic Accents

Pitch is determined by tone frequency and by frequency relationships. In music, pitch transformations that carry melodic themes are logarithmically related to the frequency changes among the tones in the melody (Backus, 1969; Hahn & Jones, 1981; Jones & Hahn, 1986). Musical scales, along with musical grammars, reflect some of the pitch constraints that are found in melodic lines. Some frequencies and frequency changes are unlikely to occur if certain scales and grammars are at work in a particular musical context, and others are more likely to occur.

Pitch relationships and melodies. The equally tempered chromatic scale and its related diatonic subsets embody the logarithmic frequency invariant, in that both rely on invariant pitch "distances," or musical intervals, such as the semitone or perfect fifth. Notes from the chromatic scale, marked off in semitone (ST) pitch units, are shown on the ordinate of Figure 1. Other pitch intervals are made up of n semitones. Thus, a perfect fifth spans 7 ST ($n=7$), a major third has 4 ST ($n=4$) intervals, and so on (Hahn & Jones, 1981).

Monophonic melodies are created by a single line or sequence of tones. They display only successive (i.e., melodic) pitch intervals and lack the covarying simultaneous accompaniment of other pitch lines, which create harmonies. Nonetheless, the serial pitch constraints apparent even in this type of tune are multiple and covarying, and present challenging problems for psychologists concerned with analysis of structure (Jones, 1978, 1981a, 1981b, 1981c, 1982; Jones & Boltz, 1987).

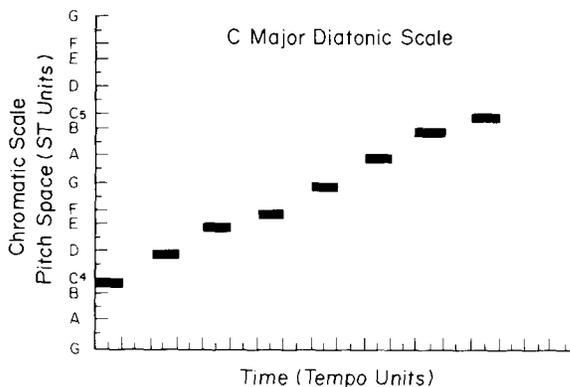


Figure 1. A pitch space-and-time coordinate system for presenting the dynamic structure of tunes. Steps on the ordinate reflect the chromatic scale unit, the semitone (ST); notes are chromatic scale notes. Steps on the abscissa reflect the tempo established by the value of the beat period. The space-time pattern is a diatonic C major pattern which itself establishes tonal relationships.

Consider the melodic sequence, drawn from the chromatic scale, presented in Figure 1. Perhaps the simplest constraint evident here is that each successive note must be higher (+) in pitch than its predecessor, creating overall an ascending pitch sequence. Serial rules that relate to these ordinal pitch changes will constrain the melody at this level of pitch contour (Dowling, 1978). But the sequence also displays a series of pitch intervals. Thus, this series of pitch relationships between adjacent tones, measured in semitone units, is +2 ST, +2 ST, +1 ST, +2 ST, +2 ST, +2 ST, +1 ST. Serial rules that relate to pitch intervals, that is, to combinations of pitch distances in time, will constrain the melody at this level of interval pitch relationships (Dowling & Bartlett, 1981).

Many melodies in Western music are tonal. This means that they reflect rules that constrain serial pitch relationships in special ways, ways that imply the existence of a single referent tone or pitch that is associated with a given musical scale. Typically, implied scales are diatonic. One example of a diatonic scale is the C major scale, which has as its referent pitch the pitch labeled C. It is shown in Figure 1. Tonal relationships are partly, but not entirely, summarized by the specified diatonic scale set. That is, tonal music follows grammatical rules that arrange these and related pitch transformations in ways that achieve a sense of "keyness" associated with tonal music. Various hierarchical systems have been proposed for depicting tonality structure and sense of key. Some are dominance hierarchies in which the key note (e.g., C) and those related to it by a perfect fifth interval (e.g., G) or a major third interval (e.g., E) are assumed to be strongest in conveying a sense of key (e.g., Krumhansl & Kessler, 1982). Others are relative frequency hierarchies in which commoner intervals in a scale, such as perfect fifths or major thirds, are seen as less informative in specifying the key note in a real musical context than are rarer intervals, such as the 6-ST pitch interval (the tritone) (Brown, in press; Brown & Butler, 1981; Browne, 1974;

Butler & Brown, 1984). Beyond issues of keyness, the tonal impact of any pitch relationship in a melody is ultimately constrained by its serial context and the tonal grammar that dictates specific temporal orderings.

In other words, there are at least three ways in which frequency relationships can determine pitch constraints within a melodic line. At an elementary level, we can identify a succession of “ups” (+, +, ...) and “downs” (-, -, ...) in pitch. These ordinal relations determine a melody’s *pitch contour*. Thematically, they constrain because a transformation of pitch contour, from, say, an ascending sequence of 5 notes (+, +, +, +) to a descending one (-, -, -, -), makes some notes and relationships impossible at certain serial locations in the melody. At another level, pitch intervals, measured in semitone units, determine a melody’s *interval* structure. Again, melodic themes based on recurrent transformations of pitch intervals, as when +2 ST, +2 ST, +3 ST, +3 ST becomes -2 ST, -2 ST, -3 ST, -3 ST, serially constrain the tune. Finally, with the overlay of tonal meaning in contexts where one pitch comes to function as a referent from which others are heard, changing key regions along with recurrent arrangements of notes having the same tonal functions will express tonal transformations and invariances. The sequence C4 E4 G4 C5, an upward arpeggio in C major, can recur in another key, say F major, as a downward sequence: F5 C5 A4 F4 (numbers indicate octave: C4 is middle C).

In short, serial constraints of various kinds can express relationships that carry melodic themes in music. Two additional facts are important. One is that the simplicity of these themes has to do not only with the various pitch relationships and their arrangements but also with the way they occur in real time. Second, these different kinds of pitch relationships clearly *covary* in music and reinforce one another. Both of these facts complicate experimental analyses of dynamic structure in interesting ways. The issue of covariation I consider later in the context of accent strength. At this point let us focus on the issue of timing of certain pitch changes.

Melodic accents. One way pitch relations are traditionally implicated in timing is through identifying *melodic (m) accents*. An accent is anything that is relatively attention-getting in a time pattern. Melodic accents may be of various sorts, but all tend to define beginnings or endings of more or less crucial time spans in a musical sequence. Things that mark beginnings and endings of coherent pitch groups will also mark durations of these groups. Consequently, any relatively prominent pitch relationship, given some melodic context, can function as a melodic accent and so tap into a higher time structure within the sound pattern. Ultimately, I will argue that rhythm, as a pattern of tone durations, can similarly offer temporal accents that join with melodic ones in this higher time structure. But first consider pitch changes and three kinds of melodic accents.

The first type of *m* accent involves pitch contour. Pitch contour yields *contour-based melodic accents*, as shown in Panel a of Figure 2. Successive pitch trajectories in a

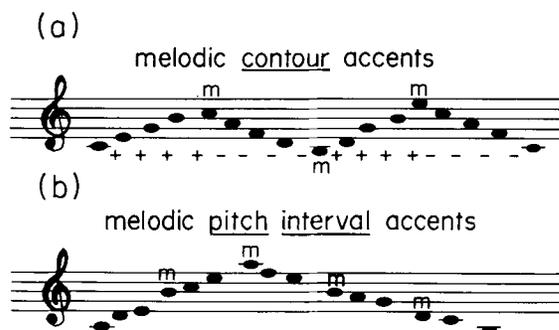


Figure 2. Two kinds of melodic accent (*m*). Panel a: Contour-based accents, which fall on locally higher and lower pitches. Panel b: Pitch-interval accents, which fall on the pitch following a locally larger pitch interval.

monophonic melody are created by transformations of contour relations (i.e., from + to - or vice versa), and these segment a melody meaningfully. Beginnings or endings of rising or falling trajectories can function as melodic accents. Locally higher and lower pitches, associated with contour changes, are potentially important melodic markers for durational spans (Boltz & Jones, 1986; Thomasen, 1982).

Melodic interval structure also groups and differentiates a pitch series via temporal ordering of pitch intervals of different sizes. Relatively small pitch intervals (1 or 2 ST) are common in Western music, and so when a theme introduces a relatively large pitch change (e.g., 4 or 5 ST), it is noticeable (Boltz & Jones, 1986). Therefore, relatively large pitch intervals, defined by certain pairs of tones, may produce melodic accents. Panel b of Figure 2 illustrates these. *Pitch-interval melodic accents* here are assumed to fall on the second tone of a pair, that is, on the one that completes the unusual pitch change.

Neither contour change nor pitch-interval change depends on an established musical key for its effect. Tonal accenting does. In contrast to contour and interval accents, which depend on local surprise or “differentness” for contextual salience, tonal melodic accents come from confirmation of a contextually governed expectation about tonal relationships (Bharucha, 1984). A resolving key note at a phrase ending exemplifies a tonal end accent that is often accompanied by a sense of confirmation or finality. These notes function as tonal end accents and their effectiveness depends partially upon culturally acquired attunements to tonal grammars.

In summary, contour, interval, and tonal pitch relationships are ordered in time to create melodic themes in monophonic Western music. They follow various rules that lead to more or less complex serial patterns. But a signal property of all is that each can function as a melodic accent. This renders their placement in real time critical.

Temporal Structure, Meter, Rhythm, and Temporal Accents

Temporal relationships refer to the various differences and proportionalities in time spans that obtain between musical elements with finite durations, things such as tones

and silences. The fact that any duration cannot be defined without some kind of energy change to indicate or mark its beginning and ending is significant. However, let us postpone questions associated with the interrelationship between markers and marked things and consider merely certain durational properties of musical time patterns. To do this, I will rely upon the generic term *temporal structure* to include three aspects of time relations in music: tempo, meter, and rhythm.

Tempo. As with pitch, in which the semitone functions as a unit relation, a functional musical time unit, namely the *beat period*, can be identified for temporal relations. The beat period communicates the tempo or pace of a piece. Theoretically, the beat period is a temporal invariant in that it is roughly constant over large musical segments. However, in practice, the beat period is not precisely invariant, for several reasons (see Jones, 1987; Shaffer, 1982).

Tempo can provide an important anchor in the time domain for a listener, in that it offers a temporal frame or perspective for perception of other time relations (Cone, 1968; Handel & Oshinsky, 1981; Jones, 1987). If we describe the pitch time structure of a melody in a space-time coordinate system, as in Figure 1, then tempo units are given by prominent markings on the abscissa or time axis. Other time periods, larger and smaller, become transformations of this unit time period. In Western metrical music, these time transformations, notationally, are ratio in nature, whereas in other kinds of music additive time transformations obtain (Jones, 1987).

Meter. Meter refers to a temporal invariance property that is related to the basic time transformations that are exploited by a musical style or genre. It involves two nested time levels, one of which is the beat period and the other of which is a larger period, the *measure* (Yeston, 1976). Metrical Western music typically embeds a constant number of unit time periods (beat periods) within the recurrent measure time span (e.g., 3 quarter notes per measure is $\frac{3}{4}$ meter). Thus a hierarchical nesting of time intervals is implied. This is a metric hierarchy.

The meter of a piece is typically known by a performer, but it must somehow be communicated through the sound pattern to the listener (Sloboda, 1985). Thus, both the composer, in delineating musical structure, and the performer, in producing it, must mark appropriate beginnings and endings of the time periods associated with beat periods and measures, so that a listener can detect that the time ratios of these periods are, for example, 1 to 2 (a duple meter) or 1 to 3 (a triple meter). Some have suggested that a performer's dynamic energy changes (stresses) associated with downbeats invariantly mark divisions of a meter and so communicate its hierarchical properties (e.g., Lerdahl & Jackendoff, 1983). However, for the listener, detection of invariance in metric relationships in the produced sounds also depends on a number of structural factors, including the tempo, the ratio invariant itself, and the degree of contextual regularity (Clarke, 1985a, 1985b; Essens, 1986; Handel, 1984;

Jones & Boltz, 1987; Povel & Essens, 1985; Povel & Okkerman, 1981).

Rhythm. The term *rhythm* refers to serial durational patterns. In theory, the durations are variations of a specified beat period and they preserve certain temporal proportions implied by a given meter. Thus, if the beat period is a quarter note, \downarrow , and the meter dictates 4 quarter notes to a measure, then the rhythmic pattern of $\downarrow \downarrow \downarrow$ (where $\downarrow = \downarrow + \downarrow$) forms a *rhythmic* figure of two shorter notes followed by a proportionately longer one. Rhythmic figures can serve as invariant themes or time patterns in their own rights (Jones, 1976, 1987), but they also offer properties that reinforce either ongoing metrical relationships or a recurrent melodic theme in some way. It is in the latter two functions that temporal accents figure prominently.

Rhythm and temporal accents. Critical aspects of dynamic pattern structure that come from rhythm are associated with the way in which a durational pattern offers *temporal (t) accents*. Like melodic accents, temporal accents are based on deviations from a norm that is contextually established by serial constraints. In this case, context and deviation are temporal: any relatively long or short tonal duration defines a temporal accent. Also, silences can determine *t* accents, which often follow or precede a pause depending upon the metric grouping, pause length, and tempo (Fraisse, 1978; Handel, 1984; Woodrow, 1951). Thus a sequence of short (S) notes, such as SSS pause SSS pause . . . , often has a *t* accent on the note following the pause. Assume each S note requires one beat (\cdot) and the pause takes two beats; then the accent sequence is $t \cdot \cdot \cdot t \cdot \cdot \cdot$ and so forth (I assume that initial pattern tones always receive *t* accents). However, if the rhythm is SSL pause SSL pause . . . , then the primary *t* accent is likely to fall on the lengthened note prior to the pause. In musical notation, if the short note is scored as a quarter note (\downarrow) (requiring one beat) and the longer one is twice its duration ($\downarrow\downarrow$), then this accent sequence becomes $t \cdot t \cdot \cdot \cdot t \cdot \cdot \cdot$ and so forth (I assume that the *t* accent falls on the first of two beats of the longer tone). In short, proportional transformations of the beat period create longer or shorter durations, some of which function as temporal accents.

Notationally, Western metrical music offers many time relationships that can be expressed as ratio time transformations within a temporal frame established by the tempo. Meter and rhythm reflect proportional variations of the referent time unit according to a fixed constraint (meter) or according to a serial patterning rule that yields temporal accents (rhythm). But a note of caution must be added. This analysis suggests that ratio time relationships, expressed notationally by a composer, are faithfully reproduced as such in musical sound patterns produced by performers. We know this is not so (cf. Jones, 1987). However, it is also true that notated time proportionalities can be, and often are, precisely expressed in music (e.g., Clarke, 1985a, 1985b; Todd, 1985). Furthermore, when deviations from notated ratios do occur, they can

be understood as *lawful* deviations, often based on tempo variations, that “play against” ideal time symmetries (Jones, 1981a, 1982; Jones & Boltz, 1987).

**Dynamic Pattern Structure in Pitch and Time:
Joint Accent Structure**

In this section I present a descriptive framework that addresses ways in which melodic and rhythmic relationships may combine to affect a listener’s response to the musical structure of a pattern as a whole. The framework is not complete. It represents a plausible initial description of interdependencies, or co-constraints, of melody and rhythm that explain recent experimental findings. Thus, it has been useful in pointing to new ways to attack dynamic pattern similarity and dynamic pattern simplicity in psychological experiments.

The key to this analysis of melodic and rhythmic interplay involves the way in which melodic and temporal accents combine into a common, higher order dynamic structure called a *joint accent structure*. Thus, serial pitch relations that enter into melodic themes have been shown to afford melodic accents (contour, interval, or tonal accents). These melodic accents mark higher order time periods in a sequence. Similarly, time relations that enter into rhythmic themes have been shown to afford temporal accents (lengthened tones, pauses, etc.). These too mark higher order time periods. Both melody and rhythm offer markers for time periods within the same sound patterns. The resulting higher order time periods and their interrelationships will form the basis of a joint accent structure analysis.

Joint accent structure refers to a particular combination of accents, accent strengths, and time symmetries that result when melodic and rhythmic accent patterns combine. One property of joint accent structure involves *accent couplings*, namely coincidences of accents in time, which, I suggest, introduce variations in accent strength. This is consistent with other researchers’ thinking along these lines (e.g., Benjamin, 1984; Berry, 1976). Another important property of joint accent structure involves its *time symmetries*, which, loosely speaking, refer to hierarchical (nested) temporal regularities (Jones, 1976, 1981b). It appears that resultant time symmetries (and asymmetries) in the joint accent structure are closely related to the way melodic and temporal accents are *temporally phased* with respect to one another. Thus, one way to assess or manipulate certain dynamic aspects of a joint accent structure involves variation of phase relations among *m* and *t* accents.

Here I briefly preview the basic idea of joint accent structure. Subsequently it is illustrated in more detail in the context of applications to recent research. To keep matters simple, consider a 6-tone melodic sequence, shown in Figure 3 as CDEABC (presented within a C major context). For illustrative purposes, consider only those *m* accents that derive from deviant pitch intervals in this sequence. In addition to an opening *m* accent, the

(a) Example 1: Rhythm is sss pause sss.



Notes:	C	D	E		A	B	C'
Beats:	•	•	•	•	•	•	•
	1	2	3	4	5	6	7
Melodic Accents:	m	•	•	•	m	•	•
Temporal Accents:	t	•	•	•	t	•	•
Joint Accent Structure:	a'''	a'	a'	a	a'''	a'	a'

(b) Example 2: Rhythm is ssLssL...



Notes:	C	D	E		A	B	C'
Beats:	•	•	•	•	•	•	•
	1	2	3	4	5	6	7
Melodic Accents:	m	•	•	•	m	•	•
Temporal Accents:	t	•	t	•	t	•	t
Joint Accent Structure:	a'''	a'	a''	a	a'''	a'	a''

Figure 3. The same melody is shown in two different rhythms. Panel a: Example 1 shows an SSS pause SSS... rhythm, producing a *t* accent after each pause. Panel b: Example 2 shows an SSLSSL... rhythm, producing a *t* accent on L notes. Example 1 produces the more strongly outlined joint accent structure since *m* and *t* accents are coupled.

pronounced pitch interval from note E to note A in this melody also produces an *m* accent on note A. In theory, this melodic string can be presented to a listener in different rhythms. Two rhythms appear in Figure 3. In Example 1 (Panel a), the rhythm is ♩ ♩ ♩ · ♩ ♩ ♩, and so forth. Here a 1-beat pause creates two 3-tone groups, and *t* accents fall on the first and fourth tones. Thus, the second *t* accent falls on note A and it coincides with the *m* accent.¹ The separate melodic and temporal beat patterns can be combined to yield a joint accent structure, and this is also shown in Example 1. Coupled or temporally coincident accents (*m* plus *t*) yield stronger markers than do single accents (*m* or *t*). Accents of different strengths are indicated in Figure 3 by primes (a''', a'', etc.), and this is related to degree of accent coupling. Thus, in integrating melody with rhythm we ask how melodic and temporal accent structure combine to strengthen a common or joint accent structure. We are also interested in how resultant accent couplings of various strengths define or mark different higher order time spans within the whole

tune. Notice that in this tune the coupling of *m* and *t* accents yields not only relatively strong markers, but markers that are separated by constant time spans (of 4 beats).

The same melody is combined with the rhythm ♪ ♪ ♪ ♪ ♪ ♪ . . . in Example 2 (Figure 3, Panel b). This rhythm differs from that of Example 1 in terms of the location of its *t* accent, which now falls on the lengthened final note of each 3-tone melodic group. As a result, *m* and *t* accents do not coincide. Considering only pitch interval accenting, this could weaken accent strengths in the joint pattern, relative to Example 1 (if we assume *no* tonal accenting, a dubious assumption in this case because notes E and C could, arguably, function as tonal *m* accents). Note, however, that in spite of this change some lawful temporal relationships remain. This lawfulness is related to the fact that *t* accents fall halfway between *m* accents in the joint sequence. Ultimately, I will summarize this lawfulness in terms of ratio-based time symmetries arising from the joint accent structure. For the present it is sufficient to observe that there is temporal regularity and that this regularity appears to be related to the fact that *t* accents observe a particular temporal phase relationship with respect to *m* accents.

SOME APPLICATIONS OF DYNAMIC PATTERN STRUCTURES

The idea of joint accent structure can be applied to the study of dynamic pattern similarity in delayed melody-recognition tasks and to the study of the effects of dynamic pattern simplicity in immediate melody-reproduction tasks.

Dynamic Pattern Similarity in Melody Recognition

Earlier I indicated that composers undoubtedly need to control the degree to which listeners sense recurrence of a familiar theme. Theme recognition is determined by the extent to which recurrent patterns preserve psychologically important structural invariances of the melody. These functional similarities, in turn, will depend on what initially captures a listener's attention and on how these relations are reinstated later during some recognition test. I have suggested that both the act of attending and the act of remembering have dynamic qualities and, indeed, that they are related psychologically (Boltz & Jones, 1986; Jones, 1976, 1981b). Given this perspective, it made sense to formulate a general hypothesis that the functional *dynamic pattern similarity* of tunes is determined by the way in which different patterns may provoke common modes of attentional following in a listener. Specifically, if people target their attention over distinctive higher order time periods, associated with combined accents, then commonalities in joint accent structures may provide a framework for understanding dynamic pattern similarity.

Recent research, addressed to a specific version of this hypothesis involving the way average listeners attend to music, has focused on the function of contour accents and

their relation to temporal accents. Specifically, Jones, Summerell, and Marshburn (1987) proposed that the joint relation of contour *m* accents and rhythmic *t* accents specifies a tune's dynamic shape, and that this global property is important in melody recognition. *Dynamic shape* refers to that part of the joint accent structure which involves the relative timing of the pitch contour *m* accents. It focuses on timing of a melody's ups and downs in pitch. Thus, dynamic shape is not simply pitch contour. It is well known that listeners confuse different melodies that have the same pitch contours (Dowling, 1978; Dowling & Fujitani, 1971; Dyson & Watkins, 1984; Edworthy, 1985; Kallman & Massaro, 1979; Watkins, 1985). But often these contour manipulations take place within a single rhythmic context, usually an isochronous one. Since real music relies only infrequently on isochronous rhythms, and indeed often covaries melody and rhythm, Jones et al. (1987) examined the effects stemming from the more typical covariant situation. The new structural construct, namely dynamic shape, therefore extends the study of pitch contour as a basis for melodic theme recognition, to include the *relative timing* of pitch trajectories that make up contour. Essentially the idea is that average people recognize musical themes in a rhythmical way involving the dynamics of pitch contour accents, and Jones et al. found support for this idea. Furthermore, confusions among themes in melody recognition were consistent with the hypothesis that dynamic pattern similarity is related to dynamic shape. Some features of this study are interesting in terms of psychological issues of tune recognition.

The paradigm was a delayed-recognition one in which listeners with modest musical training heard (in initial familiarization) three target tunes that differed from one another in melodic interval structure and contour, but not in key (tonality). Rhythm covaried with melodic structure; three different rhythms were used. Various degrees of initial familiarization (6, 21, and 33 initial trials) with these target tunes were given to different groups of listeners. After a 15-min delay, listeners were tested (unexpectedly) for recognition of key-transposed versions of target melodies. In the recognition phase, the melody of a given target could appear either in its original rhythm or in a rhythm equally familiar but which had been initially associated with a different target melody (new rhythm). Decoy melodies also occurred in the recognition phase. These consisted of novel melodies with notes of a given target rearranged either to preserve the pitch contour of a given target (same-contour decoy) or to violate its contour constraints (different-contour decoy). In these respects, the design paralleled that used by Dowling (1978), but it differed in reliance upon delayed recognition and in the addition of a rhythmic variable (original, new) in recognition. Thus, all decoys also appeared either in the original rhythm of their target or in a new one. Hence, some of the decoys precisely mimicked the dynamic shape of a given target (i.e., original-rhythm-plus-same-contour decoy). Other decoys differed from the target's dynamic shape in various degrees (e.g., original-

rhythm-plus-different-contour decoy vs. new-rhythm-plus-same-contour decoy). The design thus allowed for gradations in similarity of dynamic shape between a novel melody and a more or less familiar target during the recognition phase.

Two specific questions were relevant to the original hypotheses: How well do these listeners recognize the key-transposed target melody when its dynamic shape has been shifted by its placement in a new rhythm? How often would people be fooled by the various decoy melodies—in particular, by those that mimicked the target's dynamic shape?

The results of one experiment are shown in Figure 4. The probability of identifying a given tune as "old" (i.e., a familiar target) is plotted for all three kinds of melodies (target, same-contour decoy, different-contour decoy) as a function of rhythm (original, new) and degree of initial familiarization with the target tune. As I indicated above, people were, in fact, misled by pairing of a target with a new rhythm. Target recognition accuracy dropped significantly when the target was presented in the new rhythm. Thus, the answer to the first question is that people do not do very well at recognizing a target in a new rhythm. Second, the most confusing decoy melodies were those with the same dynamic shapes as targets (original rhythm plus same contour). False-alarm rates

were significantly higher to these than to decoys with different contours in either rhythm. A decoy employing a more subtle introduction of the recurrent contour theme with a different rhythm (new rhythm plus same contour) also provoked some, but significantly fewer, confusions. Thus, the answer to the second question is that people do find decoys with dynamic shapes similar to those of the targets most confusing. In these data (which have been converted to accuracy scores) a significant interaction involving effects of melody and rhythm was observed: decoys with the same dynamic shapes as targets were disproportionately more confusing than other decoys, a finding that speaks to the interdependence of melody and rhythm on recognition accuracy.

An interesting auxiliary finding had to do with the disproportionate influence of dynamic shape confusions in recognition when listeners were relatively unfamiliar with the targets themselves. With low familiarity, proportionately more errors were due to subjects' reporting a decoy melody with the same dynamic shape as its target to be the target itself. The ratio of false alarms to same-contour decoys in the original rhythm (i.e., same dynamic shape) to hit rates for targets in this rhythm was .62 with the 6 familiarization trials, and dropped to .367 with 33 familiarization trials. Taken together with other data, the findings reinforced the intuitive idea that less experience

EXPERIMENT 2

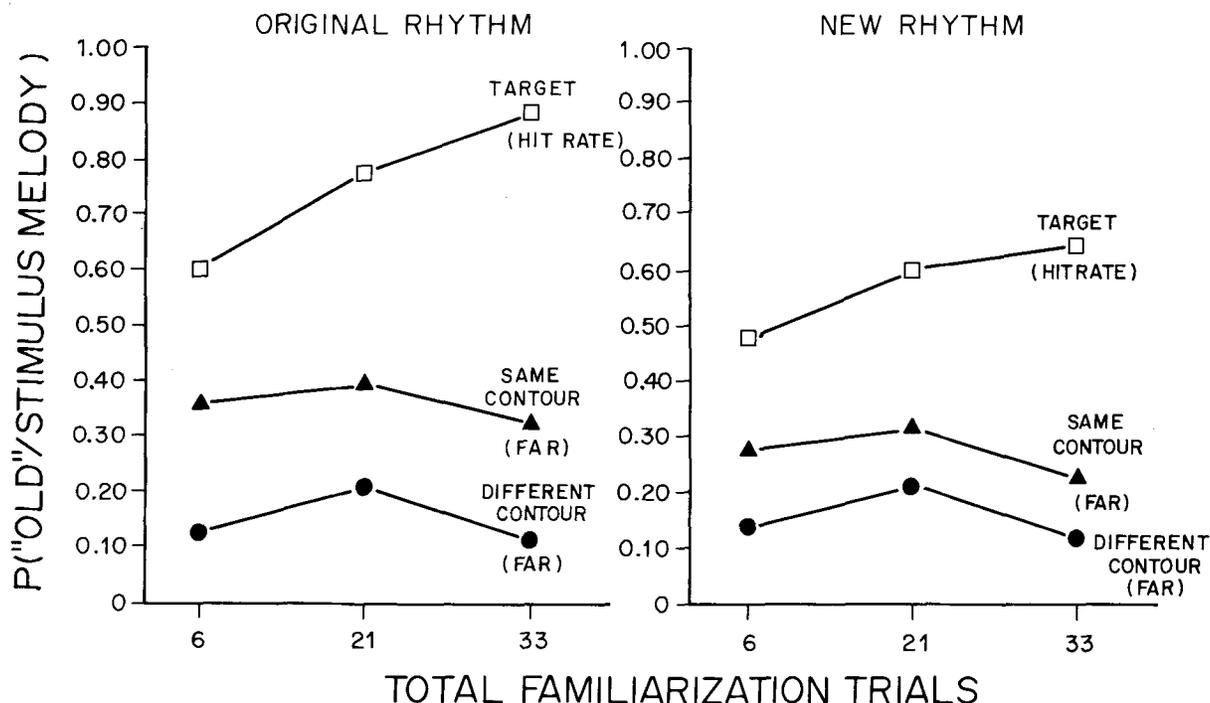


Figure 4. Proportion of "old" recognition responses (after 6, 21, and 33 target familiarization trials) to target melodies and to same-contour decoys and different-contour decoys as a function of rhythm. Note that a response of "old" is correct for targets but erroneous for decoys. (a) Original rhythm; (b) new rhythm (based on data from Jones, Summerell, & Marshburn, 1987).

with a tune would increase the chance that a listener would remember it globally, in terms of a rhythmical contour, rather than precisely, in terms of fixed pitch intervals.

In follow-up experiments that we have conducted, more specific manipulations of dynamic shape in terms of joint accent structure have been studied. For example, to boost dynamic shape prominence in initial targets, new target tunes were created that had many strong accents based on couplings of contour accents and temporal accents. Effectively, coincidences of *m* and *t* accents were maximized in the target tunes. Furthermore, new rhythms were also used (e.g., SSL pause S...LSS pause L... and SLS pause S...). Judgments of several musical experts confirmed that *t* accents always fell on the lengthened tonal duration (L). Additionally, in the recognition phase, decoy melodies with decoy contour accents were always realized by the same notes as were key-transposed targets.

These latter studies followed the same delayed-recognition paradigms as earlier ones. Furthermore, in these studies it was emphasized both in initial familiarization and in recognition that listeners should specifically *ignore* rhythm and concentrate on the pitch sequence or melody. However, the same patterns of results appeared. Confusions of decoys with the same dynamic shapes as targets were somewhat more pronounced than in previous studies, and this is undoubtedly because similarities in accent structure properties were enhanced.

Insofar as dynamic pattern similarity is concerned, all of these data suggest that rhythm and contour interact to affect a listener's sense of familiarity with a tune. One way to understand and explore this is in terms of joint accent structure. This concept suggests ways to enhance dynamic pattern similarities through strengthening of certain kinds of accents and through increasing commonalities in the relative timing of these accents. In these studies, target tunes and decoys with the same contours and rhythms as targets (i.e., with the same dynamic shape) possessed identical joint accent structures with respect to contour-based melodic accents and temporal accents.

Some additional points are relevant. The first is that even listeners instructed to ignore rhythm and focus on melody could not do so. Psychologically, rhythm and melody are integrated, and these data argue against an interpretation that the two are independent. They also argue against a simple segmentation interpretation of rhythmic effects on memory that relies on pause-based melodic codes (e.g., Deutsch, 1980). In these studies, rhythms were carefully constructed to present uniform pause structures: all rhythms parsed melodies via pauses into 3-tone groups. If such a coding account were applicable, it would mean that listeners should not be distracted by these rhythmic switches, because their segmental impact on melodic encoding is the same in all cases.

Finally, the fact that proportionately more errors are due to confusions based on dynamic shape when listeners are less familiar with the original tunes supports the idea that listeners initially respond to global relational features

that involve pitch contour. Dowling and Fujitani (1971) proposed that contour confusions might be more likely with unfamiliar tunes, and this is consistent with that hypothesis. But it goes beyond it as well. The data suggest that such contour confusions are, in part, specific to a particular rhythmic context. Furthermore, because the recognition paradigm introduced retention periods of 15 min, it is clear that these rhythmically specific contour effects on memory are relatively long-lasting.

In summary, one means of describing the interplay of melody and rhythm involves the joint accent structure. If the accent and timing relationships that characterize joint accent structure are taken as a basis for manipulating dynamic pattern similarity, this explains why people tend to judge a familiar melody as unfamiliar when it appears in a new rhythm. Its *dynamic* similarity to the old tune has changed. Likewise, if we manipulate only contour and temporal accents in a joint accent structure, then the dynamic shape of a novel melody can be made to mimic that of a familiar tune. In recognition tests listeners are fooled by dynamic shape, especially when familiarity with the target is not high. This suggests that initially the average listener attends to and remembers the dynamic shape of a tune (within some key).

Dynamic Pattern Simplicity in Reproducing Melodies

Pattern similarity refers to ways in which things are alike, regardless of how simple or complex they are. A second issue related to ways in which musical structure serves a communication function concerns musical pattern simplicity and attentional control. Simplicity, too, must be dynamically conceived, for it also pertains to the way in which melody and rhythm combine and permit more or less efficient attending. One advantage of the idea of joint accent structure is that it permits us to address not only functional similarities among dynamic patterns, but also their simplicity properties. This section considers *dynamic pattern simplicity*.

Various tasks are used to assess pattern simplicity (e.g., Garner, 1974; Jones, 1974; Simon, 1972); many involve ratings and judgments about simplicity (or complexity) as a psychological dimension. However, considering music's communicative function, it is also possible to gauge simplicity in situations where people must use pattern relations to achieve some goal. Practically speaking, in such contexts the empirically simple sequences should be those that allow more efficient use. The goal of a common task relevant to musical patterns is to train musicians to listen carefully to the arrangement of pitch relations in a melody and then reproduce them. In recent research, this melody-reproduction task served as a basis for verifying hypotheses related to predicting pattern simplicity on the basis of dynamic structure (Boltz & Jones, 1986).

One hypothesis about dynamic pattern simplicity in this context is that simple patterns are more efficiently tracked attentionally as they unfold in time. Again, because attending is assumed to be related to later dynamic acts of

remembering, patterns more readily tracked in real time should also be more accurately reproduced subsequently. Dynamically simple patterns are ones in which the initial pattern relationships more reliably forecast the “when” and “where” of later pitch changes in the pattern. To formalize this idea, Marilyn Boltz and I (Boltz & Jones, 1986) proposed that some combinations of *m* and *t* accents produce greater symmetry in higher order pattern-time relations, and it is these temporally symmetric configurations that are more effective in guiding dynamic attending. Another way of saying this is that compositions wherein an artist combines melody and rhythm so as to yield greater overall temporal regularity, as indexed by ratio time symmetries within their joint accent structure, are structurally simple and should be more accurately reproduced. Basically, this means the simplicity criteria can be based on symmetry properties of the joint accent structure.

At this point some elaboration and refinement of these constructs is necessary. Accordingly, I consider accent strength and time symmetries (temporal phasing) in more detail.

Accent strength and time hierarchies in joint accent structures. *Temporal regularity* refers to special relations among different time levels that are marked by accents in the joint accent structure. Time levels, in turn, are intrinsically related to accent strengths (Jones, 1976; Martin, 1972). It is also important to note that these time levels correspond to phrasings established by melodic and rhythmic themes, and so they do not necessarily coincide with time levels based on musical measures and meter in general. To illustrate, imagine that a melody displays phrases with relatively large pitch intervals between every fourth and fifth beat, and these catch the attention. According to the prior rationale, the tone on every fifth beat can function as an *m* accent. Table 1 shows a 16-beat melody with a pitch sequence of this type (Example 2).

Adding in rhythmic themes and temporal accents, let the accompanying durational pattern be one where a prominent pause (silence) terminates after every 7th beat. As shown, *t* accents then fall on every 8th beat. Considering *m* and *t* accents together, the joint accent structure associated with melodic-rhythmic phrasings is shown in Table 1.

Now consider accent strengths. I assume that marker strength is related to the degree of accent couplings. Here with an *m* and a *t* accent, the strongest marker is associated with coincidences of these two events. These are denoted by *a'''* with a strength metric of 3 (i.e., the number of primes). Singly occurring accents, *m* or *t*, are less prominent markers of time periods and are denoted by *a''* with a strength metric of 2. Onsets of ordinary tones create still weaker accents, namely *a'*, with a strength of 1, and simple beats, which occur within silences or longer notes, are unaccented, with 0 strength for a.

Accent strength differences amplify hierarchical time properties of the joint accent structure, as Figure 3 and Table 1 reveal. Consider the time spans of a single level of accent strength. Temporal regularity within any given level appears to be connected to the invariance of time spans associated with that level. Each span between accent markers of a given strength covers the same number of beats. But overall it is the relationship *between* strength-marked time levels that taps *m/t* global time symmetries. Notice that smaller (nested) time spans are marked by correspondingly weaker accents, but in these examples these spans are also regular. The idea of an inverse relationship between accent strength and time periods links levels; it was first proposed by Martin (1972) and later modified by others. Two modifications are relevant to the issue of dynamic pattern simplicity. One involves covariant structure and accent strengths; the other involves formalizing relationships among time spans in terms of time ratios and time symmetry.

Table 1
Two Melody-Rhythm Combinations

	Beats															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Example 1: A Simple Dynamic Pattern*																
Melodic Accents	<i>m</i>	.	.	.	<i>m</i>	.	.	.	<i>m</i>	.	.	.	<i>m</i>	.	.	.
Temporal Accents	<i>t</i>	<i>t</i>
Joint Accent Structure	<i>a'''</i>	.	.	.	<i>a''</i>	.	.	.	<i>a'''</i>	.	.	.	<i>a''</i>	.	.	.
$TR_{t/m} = \frac{8}{4} = 2$	Hierarchical Time Spans:								<i>a'''-a'''</i> : 8 <i>a''-a''</i> : 4,4,4							
Example 2: A Complex Dynamic Pattern†																
Melodic Accents	<i>m</i>	.	.	.	<i>m</i>	.	.	.	<i>m</i>	.	.	.	<i>m</i>	.	.	.
Temporal Accents	<i>t</i>	<i>t</i>	<i>t</i>
Joint Accent Structure	<i>a'''</i>	.	.	.	<i>a''</i>	.	.	.	<i>a''</i>	.	.	.	<i>a''</i>	.	.	<i>a''</i>
$TR_{t/m} = \frac{7}{4} = 1.75$	Hierarchical Time Spans:								<i>a'''-a'''</i> : 8 <i>a''-a''</i> : 4,3,1,4,2							

Note—Hierarchical time spans are given in beats. Assume that stronger (e.g., *a'''*) accents can also initiate and close time spans of adjacent weaker (e.g., *a''*) accents. Thus *a'''-a''* is summarized as *a''-a''*. **m* accent period is 4 beats; *t* accent period is 8 beats. †*m* accent period is 4 beats; *t* accent period is 7 beats.

Insofar as the problem of covariant structure is concerned, many covarying structural changes in music can be addressed in part by the idea that accent strength derives from the degree of accent couplings (Jones & Boltz, 1987). Thus if m accents arise from contour, interval, and tonal markers in a melodic context, and if all of these accents coincide on some note, then that note becomes a strong accent. In theory, the idea embraces additional ways of marking a given point in time as the beginning or end of a time span: clearly, t accents can and do additionally couple with m accents. But dynamic accents, which come from relative changes in loudness, can also "add" in. With regard to dynamic accents, the suggestion of Lerdahl and Jackendoff (1983) that only loudness changes associated with particular stresses (e.g., downbeats) mark meter, and hence metric hierarchies, is a narrower definition of dynamic accents than I intend here. But it does raise important issues concerning different kinds of accent hierarchies, as distinct from different strengths of accents; these issues are discussed elsewhere (Jones, 1987; Jones & Boltz, 1987).

Time hierarchies associated with the meter of a rhythm, which may or may not be outlined strictly by stress accents, are different from the time hierarchies that are connected with joint accent structure. The joint accent structure is based on accents that come from melodic and rhythmic figure phrasings, and perhaps other (e.g., loudness) patternings as well. According to this analysis, whenever multiple accents, of whatever sort, coincide to mark those time levels that are associated with melodic phrasings, then the joint-accent time hierarchies will be more clearly outlined. This means that if a metric hierarchy happens to coincide precisely with that from a melody-rhythm analysis, it will strengthen the demarcation of levels in the joint accent structure.

Time symmetries and asymmetries (temporal phasing). Another aspect of accent-time hierarchies relates to their time symmetries and time ratios (Jones, 1976). Some have maintained (Boltz & Jones, 1987; Jones, 1976; Jones & Boltz, 1987) that musical listeners attend in a rhythmical way and "use" time symmetries. If so, then functionally speaking, dynamic pattern simplicity should be related to a formalization of those time properties within the joint accent structure. This formalization can be expressed in terms of time ratios.

As an illustration, note that in Table 1 (Example 1) the pattern's higher order time spans, of a''' to a'' , neatly divide the total duration in half. Spans of these two levels form a time ratio of 2.00. Furthermore, global temporal regularity here is evident from the fact that these periods, in turn, are evenly divided by lower order time spans, ones involving correspondingly weaker accents (a''' to a'' spans). In fact, a common time-span ratio (TR) of 2.00 formalizes the relationship between several adjacent hierarchical levels. For example,

$$TR = \frac{a''' - a''}{a'' - a'} = \frac{8 \text{ beats}}{4 \text{ beats}} .$$

This integer ratio is based directly on two adjacent time levels in the joint accent structure. It is invariant over the whole time pattern, but it has recursive properties. It relates several embedded levels, and so summarizes temporal symmetries.²

In this illustration it is clear that the time-symmetry property of the joint accent structure, which is captured by the TR value, is directly related to spans of t accents and m accents and their interrelationships. The t to t spans are always 8 beats in duration; the m to m spans always cover 4 beats. That is, the ratio invariant time property of a joint accent structure stems from the time relation between invariant periods of m and t accents. Thus, a second time ratio, based directly on the two accent periods themselves, is possible. Denote this by $TR_{t/m}$, where the larger period (here t) always forms the numerator subscript:

$$TR_{t/m} = \frac{t-t}{m-m} = \frac{8 \text{ beats}}{4 \text{ beats}} .$$

The $TR_{t/m}$ value is also 2.00, and hence in this case $TR = TR_{t/m}$. However, this equivalence is true primarily because t and m accents themselves have invariant time periods. It is also true because the pattern *begins* with a coincidence of the two accents (m with t). Thus, these m and t accent periods, as such, bear a simple harmonic relation to one another, and furthermore they are not phase-shifted with respect to one another. It can be shown that joint-accent-structure hierarchies that are derived from these kinds of accent relations have rather elementary recursive time properties that connect the different time levels (Jones & Boltz, 1987).

There are several ways in which the temporal relationships of m and t accents can be meaningfully manipulated experimentally, and these manipulations will influence time properties of joint accent structures. One obvious manipulation involves simply changing the time value of the constant period of the m and/or t accents. Note, for example, that in Table 1, Example 1, the 4-beat m accent period is shown paired with a t accent period of 8 beats. In Example 2, however, the same melody appears in a rhythm having a t accent period of 7 beats. In the first case, a $TR_{t/m}$ ratio of 2.00 occurs, reflecting a coherent joint-accent structure time hierarchy; but in the second example, the simple harmonic relationship between m and t periods is broken, because $TR_{t/m} = \frac{7}{4} = 1.75$. This means that there is less coherence in the joint accent structure in the second case. In this sense, this time ratio indexes time symmetries of the joint accent structure: smaller integer values suggest greater temporal regularity and predictability. This, then, may serve as a basis for gauging dynamic pattern simplicity.

One reason that a change in the t accent period produces asymmetries in the joint-accent-structure time pattern is that t accents that begin in phase with m accents get out of phase with them as the pattern unfolds. Out-of-phase relations introduce more temporal asymmetries. Temporal

symmetries are more pronounced when these time ratios are small integers, whereas asymmetries are indexed by more complex ratios. This raises the issue of temporal phasing more directly.

In another manipulation of *m* and *t* relations, *m* and *t* accent periods remain invariant and, in fact, are harmonically related to one another. The *m* accent period is based on 4 beats and the *t* accent is generally 8 beats, so $TR_{t/m} = 2.00$, as in Example 1 of Table 1. It is possible to devise a rhythmic change in which *t* accents are initially phase-shifted relative to *m* accents, but neither the *m* nor the *t* accent periods are changed. An example of this occurred with the rhythm changes described in the last section, where the same melody appeared in different rhythms with identical *t* accent periods. These phase shifts, which manipulated similarity of dynamic shape through decoupling of contour and temporal accents, were achieved by modifying only the length of the first *t* accent period.

Other examples of phase-shifting are shown in Tables 1 and 2. (Example 1 of Table 1 and the phase-shifted versions in Table 2 are all assumed to refer to the same melody.) In Table 1, Example 1, with a *t* accent period of 8 beats and no phase-shift, $TR = TR_{t/m} = 2.00$. If the rhythm is changed so that the 8-beat *t* accent sequence is initially phase-shifted to various degrees, the different temporal phase relations of *m* and *t* shown in Table 2 appear. Here the initial phase shifts of the 8-beat *t* accent period are 1, 2, 3, and 4 beats long.

The accent-time hierarchies associated with all phase-

shift manipulations are also shown in Tables 1 and 2. Notice two aspects of these joint-accent-structure hierarchies: (1) The hierarchical time spans are less clearly outlined when *t* accents are shifted by a time interval greater than zero (no phase-shift) but less than the 4-beat *m* accent period (i.e., by 1, 2, or 3 beats) than when there is no phase-shift or a phase-shift equal to the *m* period (i.e., 4 beats). Fewer strong, coupled accents (*a'''*) appear. (2) The higher order time periods that are outlined by weaker accents are more uniform in the hierarchy resulting from a 2-beat phase shift than in those resulting from 1- or 3-beat phase shifts. That is, the former hierarchy has only periods of 2 and 4, whereas in the latter, periods of 3, 1, and 4 appear.

One way of assessing dynamic pattern simplicity relies on time ratios to gauge the impact of a phase-shift on a time hierarchy. Call a TR ratio based on the *initial* part of the pattern a shift-TR. For instance, in this example, if the larger initial accent period (the *m* period) is alternately divided by *t* accent periods of 1, 2, 3, and 4 beats, then shift- $TR_{m/t}$ values of 4, 2, 1.33, and 1, respectively, result for these phase shifts. Whole-integer and small-integer values of the shift-TR appear to be correlated with simpler time hierarchies. This sort of temporal-phasing metric further clarifies dynamic simplicity properties. Overall, the symmetrical time hierarchies are those with time ratios (TR and shift-TR) based on small integer values.

The argument here is that simplicity properties of joint-accent-structure time hierarchies can be gauged and manipulated through variations of ratio time relations

Table 2
Four Different Phase-Shifts of a Rhythm

Shift Condition	Beats															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1 Beat																
Melodic Accents	<i>m</i>	.	.	.	<i>m</i>	.	.	.	<i>m</i>	.	.	.	<i>m</i>	.	.	.
Temporal Accents	<i>t</i>	<i>t</i>	<i>t</i>
Joint Accent Structure	<i>a'''</i>	<i>a''</i>	.	.	<i>a''</i>	.	.	.	<i>a''</i>	<i>a''</i>	.	.	<i>a''</i>	.	.	.
$TR_{t/m} = \frac{1}{4} = 2$; Shift- $TR_{m/t} = \frac{1}{4} = 4$ Hierarchical Time Spans: <i>a''-a''</i> : 1,3,4,1,3																
2 Beats																
Melodic Accents	<i>m</i>	.	.	.	<i>m</i>	.	.	.	<i>m</i>	.	.	.	<i>m</i>	.	.	.
Temporal Accents	<i>t</i>	.	<i>t</i>	<i>t</i>
Joint Accent Structure	<i>a'''</i>	.	<i>a''</i>	.	<i>a''</i>	.	.	.	<i>a''</i>	.	<i>a'</i>	.	<i>a''</i>	.	.	.
$TR_{t/m} = \frac{1}{4} = 2$; Shift- $TR_{m/t} = \frac{1}{2} = 2$ Hierarchical Time Spans: <i>a''-a''</i> : 2,2,4,2,2																
3 Beats																
Melodic Accents	<i>m</i>	.	.	.	<i>m</i>	.	.	.	<i>m</i>	.	.	.	<i>m</i>	.	.	.
Temporal Accents	<i>t</i>	.	.	<i>t</i>	<i>t</i>
Joint Accent Structure	<i>a'''</i>	.	.	<i>a''</i>	<i>a''</i>	.	.	.	<i>a''</i>	.	.	<i>a''</i>	<i>a''</i>	.	.	.
$TR_{t/m} = \frac{1}{4} = 2$; Shift- $TR_{m/t} = \frac{1}{3} = 1.33$ Hierarchical Time Spans: <i>a''-a''</i> : 3,1,4,3,1																
4 Beats																
Melodic Accents	<i>m</i>	.	.	.	<i>m</i>	.	.	.	<i>m</i>	.	.	.	<i>m</i>	.	.	.
Temporal Accents	<i>t</i>	.	.	.	<i>t</i>	<i>t</i>	.	.	.
Joint Accent Structure	<i>a'''</i>	.	.	.	<i>a'''</i>	.	.	.	<i>a''</i>	.	.	.	<i>a'''</i>	.	.	.
$TR_{t/m} = \frac{1}{4} = 2$; Shift- $TR_{m/t} = \frac{1}{4} = 1$ Hierarchical Time Spans: <i>a'''-a'''</i> : 4,8 <i>a''-a''</i> : 4,4																

Note—Hierarchical time spans are given in beats. Assume that stronger (e.g., *a'''*) accents can also initiate and close time spans of adjacent weaker (e.g., *a''*) accents. Thus, *a'''-a''* is summarized as *a''-a''*.

among various accents associated with melodic and rhythmic patterns. Tables 1 and 2 indicate, for example, that greater uncertainty of higher order accent ("a"-a") time spans appears in sequences with more complex (e.g., $TR_{t,m}$ and shift- $TR_{m,t}$) time ratios. Given the history of simplicity research, one is tempted to develop a simplicity metric. But it is probably premature to advance a sweeping simplicity metric at this stage. A more prudent course is to develop specific task-oriented metrics that reflect particular hypotheses about the time relationships that listeners with various skill levels use in different contexts. One example of such an adaptation is described below.

Application of time ratios to dynamic pattern simplicity. Recent evidence for the usefulness of time-ratio properties in indexing hierarchical time symmetries comes from experiments using reproduction tasks.

In one (Essens, 1986), people had to reproduce the time intervals themselves, using a rhythmic tapping paradigm. This experiment found support for time hierarchies wherein time levels relate as integers with prime factors 2 or 3. Simpler patterns according to these ratio criteria showed fewer distortions when reproduced.

In another (Boltz & Jones, 1986), musically sophisticated listeners had to reproduce notationally the melodic line of patterns with various combinations of m and t accent periodicities. The m accents in these monophonic melodies came from both contour change and pitch-interval differences, and the t accents were determined by pauses within simple rhythms. Of interest, from the point of view of temporal phasing, were several conditions wherein m accents either were or were not phase-shifted by one half of the fixed period of t accents. Other, more complicated, variations of accent timing relations were also introduced. A ratio metric, more elaborate than the ones outlined here, was developed to predict accuracy levels. This metric essentially enumerated the deviations from simple time ratios, based on m and t periods, that were associated with each melody-rhythm combination. The psychological rationale was that if ratio time symmetries within a joint accent structure controlled attending, then the number of deviations ("surprises") from highly symmetric schemes would be an index of less efficient attending and poorer recall.

Three aspects of Boltz and Jones's (1986) research are relevant here. First, significant interactive effects of melodic and rhythmic combinations on reproduction accuracy were again found, reinforcing the idea that rhythm and melody are not psychologically independent in tasks in which listeners must "use" serial relations. Second, the tunes easiest to reproduce were those based on simpler m and t ratios with no phase shifts, and the next easiest were tunes that incorporated the phase-shift ratio of 2.00. Third, overall, the metric developed for this task, which was sensitive to time-ratio symmetries and accent phasings, rather nicely predicted performance, with an overall correlation of $r = .85$. In short, these findings indicated that dynamic pattern simplicity in tasks in which listeners must "use" pattern structure is related to sym-

metries in the time structure associated with the joint accent pattern. This study led to further clarification of the concepts of joint accent structure, time symmetries, and asymmetries as they have been outlined here.

Other evidence for ratio complexities in music comes from work with polyrhythms, in which, typically, temporal patterns created by co-occurring pulse trains established by two different recurrent pitches are presented to listeners. Using a tapping task, Handel and Oshinsky (1981), for example, presented various polyrhythm patterns, based on different time ratios, to listeners. They found an effect of ratio complexity. More generally, Yes-ton (1976) argued that musical rhythms based on simple integer ratio relations between time levels are consonant, whereas those with noninteger values yield polyrhythms and are dissonant at the rhythmic structure level. The parallel with joint accent structure is that a melodic accent train can either create a simply phased consonant rhythm with a temporal accent train or form a polyrhythmic-like structure based on melodic and rhythmic interrelations (among other things). This is a new way of thinking about melodic and rhythmic relations. The resulting higher order polyrhythmic-like patterns, of course, must serve the communicative and aesthetic goals of the musical artist. However, from the perspective of a listener, the ease of attentional tracking and reproduction should be greatest for tunes with joint accent structures that have generally small whole-number time ratios. These are tunes that are dynamically simple.

Finally, these ideas return us to the problem of multiple time hierarchies mentioned earlier. Metrical music has time hierarchies associated with measure spans and meter (metric hierarchies), as well as ones based on melody-rhythm phrasing (joint-accent-structure hierarchies) (Lerdahl & Jackendoff, 1983). The ratio time properties of the two kinds of hierarchies may or may not agree. It is well known that musical phrasings do not always coincide with metrical divisions appointed by the meter. The current framework opens up new ways to assess the complexity of relationships between these two kinds of time hierarchy. Among other things, the ratio relationships between the two kinds of hierarchy can be meaningfully varied.

In sum, time hierarchies of joint accent structures suggest new ways to approach dynamic simplicity in musical patterns. This approach assumes that functional simplicity in musical patterns arises, in part, from interdependencies of melody and rhythm.

IMPLICATIONS AND CONCLUSIONS

Recent theory and research addressed to complex problems of integrating melodic with rhythmic structure have been reported in this article. Some new constructs lead to new ways of experimenting with multiple components of musical structure, such as melody, rhythm, meter, and even amplitudinal changes (dynamics)—ways that capture the covarying and structurally interleaved nature of musical relationships. The emphasis is on analysis of dynamic

interrelationships in structure. Accordingly, the theoretical bias is clear: It suggests that a meaningful way to study functional similarity and simplicity involves examination and manipulation of interrelationships of melodic and rhythmic structures (among others). The dynamic framework proposed offers concepts such as joint accent structure, accent strength, and time symmetries and asymmetries. The joint accent structure refers to a higher order time structure based on time intervals marked jointly by melodic and rhythmic accents. Accent strength and symmetry properties relate to this structure. Preliminary support for these concepts comes from research on melodic theme recognition and dynamic pattern similarity and on melodic reproductions and dynamic pattern simplicity.

One abiding implication of this presentation is that the active listener attends to and remembers temporal patterns in ways that "use" a pattern's higher order time relationships. The simplest way to express this idea is to say that attending is a time-based activity with special sensitivities to temporal lawfulness in world structure. Remembering, then, is related to attending in that it reflects a recapitulation, perhaps time-scaled, of the original acts of attending. This idea explains both recognition memory and reproduction performance. Structural control of attending and remembering comes by way of properties of the joint accent structure. Thus, beyond these broad assumptions, there are specific implications with regard to musical attending and ways in which it can be temporally directed. Attending exhibits temporal selectivity and reliance on time symmetries and temporal recursivity. Temporal selectivity refers to the idea that people learn, on the basis of initial temporal regularities, when to anticipate future events of importance. This is useful in music because it forms the basis of more efficient attending and detection of new relationships. It also offers a dynamic foundation for the artful surprise as a deviation from the anticipated "where" and "when" in musical space and time. Time symmetry and temporal recursivity refer to lawful relationships between time levels in higher order time patterns. Time hierarchies of the joint accent structure reveal recursivity when they have a high degree of ratio time symmetry. These are dynamically simple patterns. Psychologically, this means that listeners can more easily relate time levels and can shift attending over one or another. Some temporal phasings of accents create higher order time asymmetries in the joint accent structure, and these patterns are more difficult to track in real time. Many of these ideas about dynamic attending are developed more fully elsewhere, where additional psychological notions, also implied by the present analyses, are considered in more depth (Jones & Boltz, 1987).

REFERENCES

- BACKUS, J. (1969). *The acoustical foundations of music*. New York: Norton.
- BENJAMIN, W. E. (1984). A theory of musical meter. *Music Perception*, **1**, 355-413.
- BERRY, W. (1976). *Structural functions in music*. Englewood Cliffs, NJ: Prentice-Hall.
- BHARUCHA, J. J. (1984). Anchoring effects in music: The resolution of dissonance. *Cognitive Psychology*, **16**, 485-518.
- 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. (in press). The interplay of set content and temporal context in a functional theory of tonality perception. *Music Perception*.
- BROWN, H., & BUTLER, D. (1981). Diatonic trichord as minimal tonal cue-cells. *In Theory Only*, **5**, 39-55.
- BROWNE, R. (1974). Review. *Journal of Music Theory*, **18**, 390-409.
- BUTLER, D., & BROWN, H. (1984). Tonal structure versus function: Studies of the recognition of harmonic motion. *Music Perception*, **2**, 6-24.
- 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*. New York: Academic Press.
- CONE, E. (1968). *Musical form and musical performance*. New York: Norton.
- DEUTSCH, D. (1980). The processing of structured and unstructured tonal sequences. *Perception & Psychophysics*, **28**, 381-389.
- DOWLING, W. J. (1978). Scale and contour: Two components of a theory of memory for melodies. *Psychological Review*, **85**, 341-354.
- DOWLING, W. J., & BARTLETT, J. C. (1981). The important interval information in long-term memory for melodies. *Psychomusicology*, **1**, 30-49.
- DOWLING, W. J., & FUJITANI, D. S. (1971). Contour, interval and pitch recognition in memory for melodies. *Journal of the Acoustical Society of America*, **49**, 524-531.
- DYSON, M. C., & WATKINS, A. J. (1984). A figural approach to the role of melodic contour in melody recognition. *Perception & Psychophysics*, **35**, 477-488.
- EDWORTHY, J. (1985). Melodic contour and musical structure. In P. Howell, I. Cross, & R. West (Eds.), *Musical structure and cognition*. New York: Academic Press.
- ESSENS, P. (1986). Hierarchical organization of temporal patterns. *Perception & Psychophysics*, **40**, 69-73.
- FRAISSE, P. (1978). Time and rhythm perception. In E. C. Carterette & M. P. Friedman (Eds.), *Handbook of perception* (Vol. 8). New York: Academic Press.
- GARNER, W. R. (1974). *The processing of information and structure*. Potomac, MD: Erlbaum.
- HAHN, J., & JONES, M. R. (1981). Invariants in auditory frequency. *Scandinavian Journal of Psychology*, **92**, 627-651.
- 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 & Psychophysics*, **30**, 1-9.
- JONES, M. R. (1974). Cognitive representation of serial patterns. In B. Kantowitz (Ed.), *Human information processing: Tutorials in performance cognition*. Potomac, MD: Erlbaum.
- JONES, M. R. (1976). Time, our lost dimension: Toward a new theory of perception, attention, and memory. *Psychological Review*, **83**, 323-355.
- JONES, M. R. (1978). Auditory patterns: The perceiving organism. In E. C. Carterette & M. P. Friedman (Eds.), *Handbook of perception* (Vol. 8). New York: Academic Press.
- JONES, M. R. (1981a). Music as a stimulus for psychological motion: Part I. Some determinants of expectancies. *Psychomusicology*, **1**, 34-51.
- JONES, M. R. (1981b). Only time can tell: On the topology of mental space and time. *Critical Inquiry*, **7**, 557-576.
- JONES, M. R. (1981c). A tutorial on some issues and methods in serial pattern research. *Perception & Psychophysics*, **30**, 492-504.
- JONES, M. R. (1982). Music as a stimulus for psychological motion: Part II. An expectancy model. *Psychomusicology*, **2**, 1-13.
- JONES, M. R. (1984). The patterning of time and its effects on perceiving. *Annals of the New York Academy of Sciences*, **423**, 158-167.
- JONES, M. R. (1985). Structural organization of events in time: A review.

- In J. A. Michon & J. L. Jackson (Eds.), *Time, mind, and behavior*. Heidelberg: Springer-Verlag.
- JONES, M. R. (1987). Perspectives on musical time. In A. Gabrielsson (Ed.), *Action and perception in rhythm and music*. Stockholm: The Royal Swedish Academy of Music Publications.
- JONES, M. R., & BOLTZ, M. (1987). *Dynamic attending and responses to time*. Manuscript submitted for publication.
- JONES, M. R., BOLTZ, M., & KIDD, G. (1982). Controlled attending as a function of melodic and temporal context. *Perception & Psychophysics*, **32**, 211-218.
- JONES, M. R., & HAHN, J. (1986). Invariants in sound. In V. McCabe & G. Balzano (Eds.), *Event cognition: An ecological perspective*. Hillsdale, NJ: Erlbaum.
- JONES, M. R., KIDD, G., & WETZEL, R. (1981). Evidence for rhythmic attention. *Journal of Experimental Psychology: Human Perception & Performance*, **7**, 1059-1073.
- JONES, M. R., SUMMERELL, L., & MARSHBURN, E. (1987). Recognizing melodies: A dynamic interpretation. *Quarterly Journal of Experimental Psychology*, **39**, 89-121.
- KALLMAN, H. J., & MASSARO, D. W. (1979). Tone chroma is functional in melody recognition. *Perception & Psychophysics*, **26**, 32-36.
- 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.
- LERDAHL, F., & JACKENDOFF, R. (1983). *A generative theory of tonal music*. Cambridge, MA: MIT Press.
- MARTIN, J. (1972). Rhythmic (hierarchical) vs. serial structure in speech and other behavior. *Psychological Review*, **79**, 487-509.
- POVEL, D.-J., & ESSENS, P. (1985). Perception of temporal patterns. *Music Perception*, **2**, 411-440.
- POVEL, D.-J., & OKKERMAN, H. (1981). Accents in equitone sequences. *Perception & Psychophysics*, **30**, 565-572.
- SHAFFER, L. H. (1982). Rhythm and timing in skill. *Psychological Review*, **89**, 109-122.
- SIMON, H. (1972). Complexity and the representation of patterned sequences of symbols. *Psychological Review*, **79**, 369-382.
- SLOBODA, J. A. (1985). *The musical mind*. Fair Lawn, NJ: Oxford University Press.
- THOMASSEN, J. M. (1982). Melodic accent: Experiments and a tentative model. *Journal of the Acoustical Society of America*, **71**, 1596-1605.
- TODD, N. (1985). A model of expressive timing in tonal music. *Music Perception*, **3**, 33-58.
- WATKINS, A. J. (1985). Scale, key, and contour in the discrimination of tuned and mistuned approximations to melody. *Perception & Psychophysics*, **37**, 275-285.
- WELKER, R. L. (1982). Abstraction of themes from melodic variations. *Journal of Experimental Psychology: Human Perception & Performance*, **8**, 435-447.
- WOODROW, H. (1951). Time perception. In S. S. Stevens (Ed.), *Handbook of experimental psychology*. New York: Wiley.
- YESTON, M. (1976). *The stratification of musical rhythm*. New Haven: Yale University Press.

NOTES

1. The *t* accent rules will depend on tempo and serial grouping, among other things, as indicated earlier. The *m* accents can arise from contour, interval, or tonal relations; here a more realistic version of *m* accenting might include tonal accents on the final C and on note E.
2. Temporal recursiveness was originally formalized as a time ratio, C_1 , relating adjacent time levels in a time hierarchy (Jones, 1976). Here $C_1 = TR$. In cases where TR is the same across all levels, then $TR = C_1 =$ a constant, and the recursive formulas of Jones (1976) and Jones and Boltz (1987) apply to relate nonadjacent, embedded time levels.