Tracking Musical Patterns using Joint Accent Structure

MARI RIESS JONES AND PETER Q. PFORDRESHER
The Ohio State University

Abstract Joint Accent Structure (JAS) is a construct that uses temporal relationships between different accents in a melodic pattern as indices of its complexity. Concordant patterns are ones in which the periodic recurrence of melodic accents form simple ratios with the period of temporal accents (e.g., 1:1, 1:2), whereas Discordant patterns have periods of melodic and temporal accents with a more complex accent period ratio (e.g., 3:2). Participants were told to selectively attend to and synchronize finger taps with accents in two experiments that examined attentional tracking to musical patterns having a "concordant" or "discordant" JAS. Results indicated that tapping was more variable with discordant than with concordant JAS patterns, both with respect to produced inter-accent time periods and with respect to the phase of taps relative to accent onsets. These findings are interpreted in terms of real time attending and its control by event time structure.

In listening to music one often experiences a sense of temporal anticipation that is crucial to the impact of a piece. We contend that part of this experience derives from the establishment of regular accent relationships in that piece. An accent may be defined as any element in an auditory sequence (e.g., a tone) that stands out from others, usually because it disrupts the context established by surrounding elements (e.g., it is longer in time or higher in frequency). Accents arise from changes along different specifiable dimensions (such as time or frequency); we refer to an accent defined by a change along one dimension as a distinct accent "type." Instances of various accent types may mark out distinct time periods within a musical event, lending to it a coherent temporal structure. For example, a coherent melodic line may designate regularly occurring melodic-type accents by virtue of changes in contour, pitch distance or implied harmony, as this information arises from distinct movements of pitch in time. Similarly, a melody's rhythm may designate a series of temporal-type accents through lengthened tones and/or rests.

Much research has been directed toward understanding the effects of single accent types (e.g., melodic only) on the perception of musical events. In contrast, we know little about the temporal coherence of more intricate patterns that result from combinations of different accent types. Indeed, the relationships between different accents in time may contribute to the temporal coherence of a pattern. When different accent types combine in a musical event, they outline a time structure that has been termed that event's "Joint Accent Structure" (Jones, 1987, 1993). Different Joint Accent Structures (JAS's) result when the positioning of accents form different higher order time relationships. The present research examines the role of different JAS's in real time attending to simple musical events.

We suggest that listeners who are sensitive to accented relationships respond directly to time intervals between various accents within a JAS, and may use higher-order invariant aspects of this time structure to guide attending in a preparatory fashion. In this way a JAS enables listeners to monitor unfolding events in time. Consider, for instance, the first few bars of Beethoven's fifth symphony shown in Figure 1. The fourth note of the piece receives an accent from two different sources: an increase in duration (a temporal accent) and a change in pitch (a melodic accent). For this excerpt, the Joint Accent Structure created by the co-occurrence of a temporal and a melodic accent provides a clear time structure, one characterized by a strongly marked and recurrent time period. This invariant periodicity, in turn, creates a strong expectation for future accents, which Beethoven fulfills in the next several measures.

The preceding example outlines a relatively simple JAS. Here, two different accent types (melodic, temporal) coincide to outline a common recurrent time period. In more complex JASs, different accents specify different, possibly conflicting, time relationships. When such time relationships conflict, the accent types involved may neither coincide in time nor outline the same recurrent

Opening Motif for Beethoven's 5th Symphony

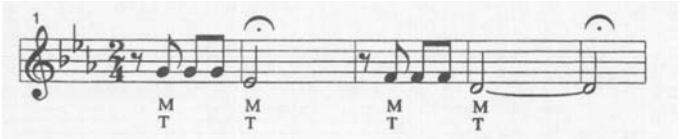

Figure 1. The opening theme of Beethoven's 5th symphony. Melodic and temporal accents are indicated by the letters "M" and "T", respectively.

inter-accent period. We hypothesize that a complex JAS will hinder a listener's ability to monitor, and hence anticipate, the trajectory of a melodic event. This might be evident in listeners' responses to accents as they occur. To test this hypothesis, we constructed two sets of music-like patterns that differed in JAS temporal coherence: concordant (simple) and discordant (complex). In both sets, melodic accents (**m**) were derived from contour inflections, and temporal accents (**t**) were derived from silences (rests) in the rhythmic pattern. In Experiment one, all experimental sequences comprised melodic accents that occurred every four beats; thus, **m** accents have a four beat "accent period." In this experiment, the temporal accents for concordant patterns assumed a four beat accent period whereas those for discordant patterns assumed a six beat period. The relations between periods of different accent types in concordant patterns therefore coincided (both four-beat periods) whereas the relations between the two different accent periods in discordant patterns created a polyrhythm. In Experiment two, melodic and temporal accent relationships were reversed. Our aim in these experiments is to examine the influence of these different JAS's on how and how well people attentionally monitor a sequence in real time.

BACKGROUND

In this section, we provide background for our experimental approach and predictions. We first consider evidence for various accentuation manipulations; next we review related research on JAS.

Accents and Accent Strength

Our experiments concentrate primarily on different combinations of melodic and temporal accents. They build on findings that indicate the perceptual salience of both types of accentuations. Traditionally, an accent has been defined as an element of a larger auditory event that is "marked for consciousness" (Cooper & Meyer, 1960; p. 8), i.e., it attracts attention. To this we add the stipulation that an accent derives from change along a dimension of sound that deviates from the preexisting serial context. Furthermore, successive accents help the listener organize a melody into coherent groups of elements.

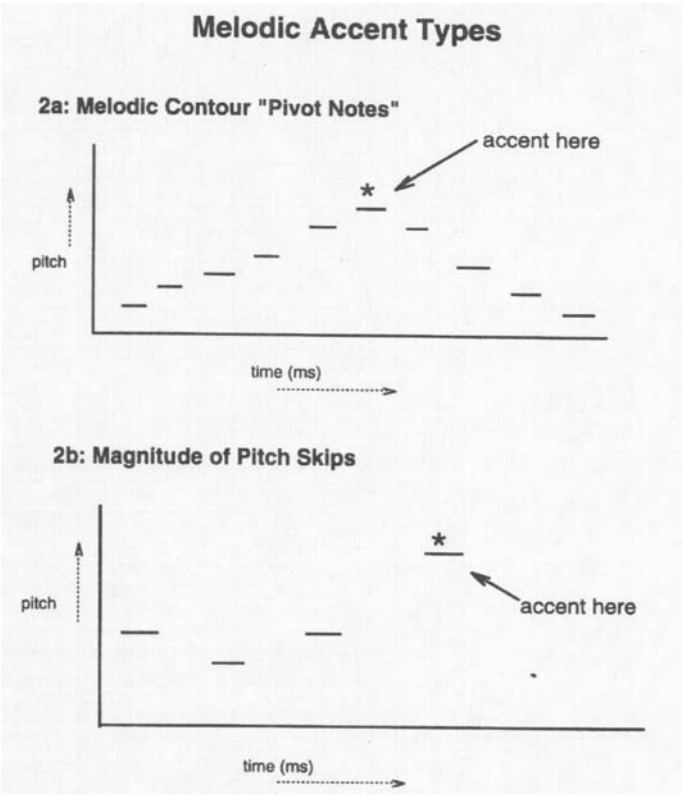

Figure 2. Two examples of melodic accentuation, using a pitch space vs. a time coordinate system.

Accentuation, therefore, involves grouping, and accentuation may operate with or against natural biases people have to group sequences in certain ways. For instance, even when no explicit tone lengthenings or pauses appear in a monotone sequence people tend to "hear" groups, often in sizes of two, defined by illusory lengthenings, pauses and/or heightened intensity. In addition, they typically place subjective accents on the first element of such groups (Woodrow, 1909). More recently, other evidence suggests that binary temporal groupings are preferred (e.g., Drake, 1993; Dawe et al., 1993, 1994; Parncutt, 1994). Accents that determine some other groupings (e.g., 3 or 6), thus, may conflict with this bias.

Melodic Accents. A variety of melodic accents are possible in musical events; among these are contour inflection, pitch skip, and tonal accentuation (Jones, 1987). Most arise from a change in pitch that signals a deviation from some pre-existing pitch motion.

First, the direction of pitch movement itself ("melodic contour") can produce accentuation. This accent, therefore, results when the deviation from pitch motion involves a change in direction. For example, as a melody rises in pitch some have maintained that listeners expect

future events to continue this trajectory (within limits) (Narmour, 1992a; Jones, 1976; 1990); in other cases, expectancies are more complexly linked to higher-order contour patterns as in the gap-fill configuration (Schmuckler, 1989)[1]. In both, it is likely that a change in trajectory direction stands out perceptually as suggested in Figure 2a; indeed, current research provides evidence for the salience of contour changes as accents (Bigand, 1997; Boltz, 1991; Boltz, Marshburn, Jones, & Johnson, 1985; Huron & Royal, 1996; Jones & Boltz, 1989; Thomassen, 1982). Most suggest that the accent falls on the "pivot" note - i.e., the note directly preceding a direction shift (see Figure 2 a). It is this melodic accent that our experiments use.

A second melodic accent type involves pitch skips (Jones, 1987; 1993). In these, the deviation in pitch movement involves a change in the relative magnitude of semitone (ST) intervals, usually signalled by an enlarged interval between a note "i" and "i + 1" as illustrated in Figure 2b (although unisons often provide accentuation as well). The event following the interval change will receive some accentuation regardless of pitch contour.

Finally, a third type of melodic accentuation has been termed "tonal accent." Tonally accented elements may draw attention by virtue of a change in the direction of some tonal implication. Thus, a note that significantly decreases or increases tension associated with moving toward or away from the tonic within an established tonal frame is often heard as accented (Jones, 1987; 1993). Little research on tonal accentuation exists, although there is evidence that shifts in harmony create melodic accents (Dawe, Platt, & Racine, 1993, 1994). Holleran, Jones, and Butler (1995) found that within an established key, individual tones that violate expectancies about implied tonal harmonies are more noticeable than tones bearing larger pitch interval changes that do not violate these expectancies. Although much research exists on the powerful expectancy effects that various tonal contexts exert on judgments about single tones and chords, this work does not explicitly address tonal accentuation (Bharucha, 1994; Bharucha & Krumhansl, 1983; Bharucha & Stoekig, 1987; Butler & Brown, 1984; Krumhansl & Kessler, 1983).

Our choice of contour accents as a basis for manipulating melodic accentuations is a pragmatic one. First, there exists more evidence to suggest how melodic contour accents arise than for, say, tonal accents. Second, contour changes appear to influence the performance of both musicians and nonmusicians in many tasks. To be sure, Deliege (1987), using a segmentation task, found only a modest influence of contour in brief melodies. But others, using memory tasks, have demonstrated strong effects of contour on memory for even brief melodies (Dowling & Fujitani, 1971; Dowling, 1978; Watkins, 1985). Although this research has not directly connected melodic contour manipulations to accent structures, other work does imply such a connection (Boltz & Jones, 1986; Jones & Ralston, 1991). Indeed, an empirical basis for choosing contour over other melodic accents stems from Huron and Royal (1996). In a sample including European classical and folk tunes, the authors found empirical support for only contour inflection points (see also Thomassen, 1983), using correlation with strength of metrical position as an index of accent strength. This implies that contour accents may override other melodic accents. For our purposes, it suggests that a reasonable starting point in testing the JAS hypotheses is with contour accents.

Temporal Accents. Temporal accents stem from a change in timing that signals a break in the pre-existing temporal flow. For instance, a long note that follows a string of shorter notes will receive accentuation. Similarly, pauses within an acoustical sequence break up the flow into discrete but accented groups, where temporal accents are heard on either the first or last group member.

Of the two kinds of temporal accents that exist, namely those created by tone lengthening and those created by pauses, we focus on the latter. Again this is a pragmatic choice. Our aim is to show that accents function as attentional targets within a JAS. If we use lengthened tones, it could be argued that their attentional prominence might not result from the structural properties of accents, but rather that the tones are more memorable because lengthening gives the listener more time to process them (e.g., Massaro, 1972).

In addition we aim to use beginning accents, wherein the temporal change confers an accent on the beginning of each accent-based group, rather than end accents where the accent falls on the final element. Lengthened tones lead to closure and tend to be heard as end accents (e.g., Narmour, 1996)[2]. For these reasons we rely on distinct silences to convey beginning **t** accents.

Our stimulus patterns incorporate two structural features designed to promote beginning temporal accentuations; these involve use of relatively long silences and of group sizes greater than two. No systematic research exists on positioning of temporal accents in

1 To be clear, neither Narmour nor Schmuckler consider these melodic contours in the manner suggested here; both, however, point to the importance of certain melodic shapes.

2 Narmour (1996) distinguishes cumulative temporal patterns such as (short-long) from counter-cumulative (long-short) indicating that the former tend to produce closure whereas the latter are nonclosural (p. 287). In this sense, when a group ends with a lengthened duration, this functions as an end-accent in that it provides closure.

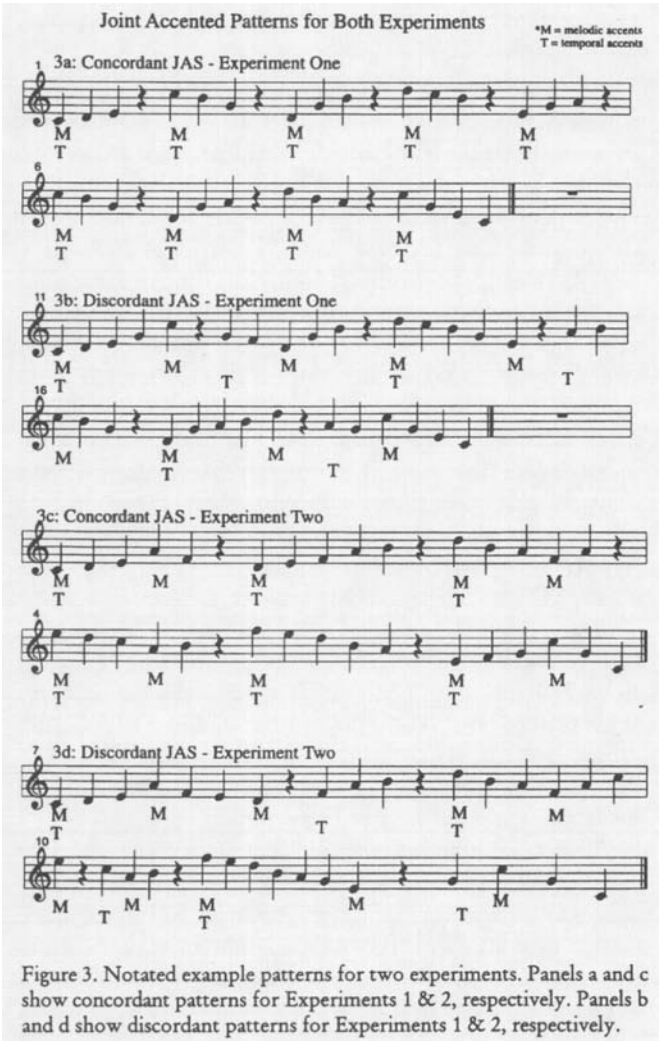

Figure 3. Notated example patterns for two experiments. Panels a and c show concordant patterns for Experiments 1 & 2, respectively. Panels b and d show discordant patterns for Experiments 1 & 2, respectively.

groups greater than two in melodic sequences. Research that does exist relies on brief monotone sequences; it suggests that tones that begin or end groups of tones are the most likely to be accented with the likelihood of one or the other or both functioning as temporal accents varying with circumstances (e.g., Povel & Okkerman, 1981; Povel & Essens, 1985). However, it is not uncommon to find that t accents based on rests determine beginning rather than end accents in groups of more than two tones and we follow this convention.

Finally, with respect to the strengths of temporal accents relative to melodic accents (including contour accents), there is relatively little empirical data that specifically assesses perceived accent strength. Some evidence suggests that temporal accents, or at least temporal aspects of an event, may override melodic accentuation. Narmour (1992b) outlined situations in which temporal accents can overwhelm (i.e. "envelop") other accents; these involve complex combinations of intervallic, registral and other properties. The data of Huron and Royal (1996) also imply that temporal aspects of musical structure may drive melodic accents. Even with simple tunes, reproductions indicate that rhythmic accent patterns exert a more powerful effect on performance than melodic ones (Drake, Dowling, & Palmer, 1991; Drake & Palmer, 1993). Given these considerations, our instructions and design attempted to ensure that listeners attended to the melodic contour as well as the rhythmic groupings of experimental patterns.

Joint Accent Structure

The JAS construct rests on the assumption that the melody and rhythm of an unfolding musical event are perceptually interdependent. However, there is evidence challenging this assumption. For instance, performance on certain rating tasks indicates that melodic and temporal factors have an additive and not an interactive influence when listeners rate a melody's tonal closure (Palmer & Krumhansl, 1987). Additionally, neuropsychological research has shown a double dissociation in the effect of brain injury on the perception of melody vs. rhythm (see Peretz & Morais, 1989, for a review). Both pieces of evidence support the idea that these two factors work independently. However, in other experimental situations normal listeners are unable to comply with selective attending instructions that require listeners specifically to "ignore" the rhythm of a melodic pattern, suggesting that decoupling of temporal from melodic structure is difficult (e.g., Jones & Ralston, 1991). We suggest that many listening contexts tacitly encourage listeners to integrate melodic and rhythmic aspects of a melody. Moreover, the relative timing of different accent types, as accounted for by JAS, may facilitate the integration of accents. In much musical composition, accent types arising from melody and rhythm work in tandem to realize a composer's aesthetic goals; and it is likely that listeners respond to these calculated dependencies.

Our view is that timing brings about integrated relationships between accent types. In a JAS two aspects of timing relationships are of concern: (1) ratios formed by different accent periods, and (2) serial phase ratios between locations of different accents. In a JAS, both aspects of timing work in relation to a beat period implied by the structure of an unfolding series of tone onsets. A third property of the JAS construct depends on the two aforementioned relative timing features; this is accent strength (3). We consider each of these three JAS properties in this section.

Accent period ratios. We have already briefly described the idea of accent period ratios as it figures in the JAS construct. Figure 3 illustrates the idea in more detail using example stimulus patterns from Experiments 1 and 2. A

time ratio juxtaposes the accent period defined by one accent type with the accent period of another type. In Figure 3a, for example, the four-beat melodic accent can be juxtaposed with the four-beat temporal accent yielding an accent period ratio of 4:4 or 1:1, a simple time ratio. A similar JAS relationship is shown in Figure 3c where the **m** and **t** accents confer periods of six beats (1:1 period ratio)[3]. In other cases, where the **t** accent retains a four beat period but the **m** accent is 2 or 8, the ratio involved continues to remain rather simple (1:2, 2:1). These are Concordant JAS patterns because their accent period ratios can be expressed as small integers (e.g., 1,2,3...).

Discordant JAS patterns are those which manifest accent period ratios that cannot be expressed as small whole numbers. A Discordant stimulus example appears in Figure 3b. In this case, the melodic accents maintain a basic period of four beats but the temporal accents occur with a period of six beats, leading to a more complex time ratio of 2:3. Still more complex JAS patterns would involve ratios of 4:5 or 4:7 where different **t** accent periods produce more complex time ratios.

Serial Phase Ratios. Although time ratios based on fixed periods between accents of different types reflect overall aspects of the time structure of a JAS, they do not necessarily capture the specific phase relations between **m** and **t** accents at different points within a serial pattern. When **m** and **t** accents do not coincide, the sequence of inter-accent time intervals within a pattern can become complex due to continually changing phase relationships among the two accents. Consider, for example, the Concordant pattern in Figure 3a where both **m** and **t** accents observe 4 beat periods. If both accent types retain four-beat periods but **m** accents are shifted to occur one beat later, the two accents would follow one another by a fixed phase lag, instead of the two accents co-occurring throughout the event. In this case, a phase ratio of .25 would obtain at each serial accent location because the melodic accent now falls one quarter of the way through the temporal accent's period. Inter-accent time intervals of one and three beats would therefore alternate throughout the pattern. Of course, for a JAS with an accent period ratio other than 1:1, phase differences between accent types cannot be constant. If the period ratio is harmonic, then one accent period will be some multiple (2,3,4...) of the other, but if it is nonharmonic (more complex) then the onset of one accent period will wind around that of the other, creating a systematic and often complicated series of phase ratios throughout the pattern.

Accent Strength. An inevitable consequence of the manipulation of period and phase ratios among accents is that the perceived prominence of a particular tone may be affected. This is illustrated in Figure 3a. Notice that in the concordant pattern, where inter-accent periods are identical (both 4 beats), and the phase ratios are zero, all **m** and **t** accents coincide. In this case, it is possible that the individual strengths of coincident two accents combine in some way, thereby changing the overall prominence of a given element relative to its perceived strength as a single accent type. By contrast, such combinations are impossible in sequences where accents do not coincide (e.g., Discordant patterns). There is evidence that coincidences of two or more accent types enhance the psychological impact of the tone on which these occur; Drake et al. (1991) found that accuracy of performing a given note increased with the number of different accent types falling on it (cf. Jones, 1987).

In sum, the three distinctive features of a JAS are accent period ratios, phase ratios and accent strength. These features are tightly interwoven such that changes in one often produce changes in another.

Joint Accent Structure and Real Time Attending

The present research is designed to examine the effect of manipulating accent period ratios in JAS patterns on real time attending and expectancies about accents. Our view of expectancy is that it refers to a future-oriented aspect of attending that is inherently time-based. Thus, listeners come to use temporal regularities within JAS patterns to guide attending toward "when" critical accents will occur. Elsewhere it has been proposed that the psychological mechanism underlying the control of attending by an event's time structure involves entrainments of multiple attentional rhythms (Jones, 1976). Recent models of real-time attentional tracking based on this approach have been formulated in terms of entrainable attentional oscillators (Large & Kolen, 1995; Large & Jones, 1997).

Although there is evidence that more complex JAS patterns hinder performance relative to simple ones, past research along these lines has focused largely on performance in various sorts of memory tasks (Boltz & Jones, 1986; Deutsch, 1980; Drake et al., 1991; Drake & Palmer, 1993; Jones & Ralston, 1991; Monahan, Kendall, & Carterette, 1987). In other cases, manipulation of a JAS was shown to affect listeners' estimation of a melody's total duration (Jones & Boltz, 1989; Boltz, 1991). Whereas these findings are consistent with the hypothesis that people respond to melodic and rhythmic accent patterns as interdependent, they do not establish that such integrated

3 Technically, the ratio of melodic to temporal accents in Experiment 2 is not precisely 1:1, as the melodic accent period alternates successively between accents periods of 3 and 6 beats. This melodic accent pattern creates alternating accent period ratios between melodic and temporal accents of 1:1 and 2:1. As we classify both ratios as "simple," we refer to AP ratios for concordant patterns in Experiment 2 as 1:1, for simplicity.

accent patterns guide attending in real time. In fact, Peretz and Morais (1989) cogently argue that independence of melody and rhythm may reside in early stages of processing which are not reflected in the aforementioned designs. These authors suggest that more research using on-line tracking procedures may assist in sorting out various influences on attending. Although it cannot be argued that overt tapping necessarily reveals early processing stages, it is superior to memory tasks in reflecting properties of real-time attending.

Recent models of dynamic attending offer a means of expressing distinctions related to the dependence/independence issue. If different internal (attending) oscillators entrain respectively to melodic and temporal accent periodicities, then these two oscillators may function in either an interactive or independent manner with respect to one another in guiding attending in a tracking task (e.g., Large & Kolen, 1995; Large & Jones, 1997). If they operate independently, then performance on concordant and discordant patterns should not differ. However, if attending oscillators locked into melodic accent periods interact with those tracking periods of t accents, then performance will be better with simple ratios of their periods (i.e., concordant JAS patterns) than with complex ones (discordant patterns).

Accordingly, in the present research we relied on a perceptual-motor task in which listeners with different amounts of musical training were asked to tap in synchrony with perceived accents within each tune. First they listened to a pattern and then, during a second presentation, they attempted to anticipate/respond to accent points *as these occurred in the unfolding pattern.* Concordant and discordant melodies (similar to those of Figure 3) along with a variety of control melodies were used. As described earlier, melodic contour accents occurred with a period of four beats in Experiment 1 whereas they occurred with periods of three or six beats in Experiment 2. Concordant and discordant JASs were created by crossing temporal accents (from rests) of periods of four or six beats, respectively, with these melodies.

We had two general predictions for these tracking experiments. First, we predicted that the taps produced for concordant patterns should be more accurately positioned on (or before) the onset of the accented tone. Second, the periods produced for discordant patterns should be more variable than those produced for concordant patterns. We tested each of these predictions using two different methods. The first method involved the timing of produced periods, which is simply the time elapsed between any two taps. For these analyses, accuracy can be indexed by the degree to which the average produced period for any given pattern approximates the inter-accent period of that pattern. The variance of this produced period relates to our second prediction regarding period variability. The second method of analysis indexes the phase relationship between each tap and the onset of the subsequent melodic accent. Relative phase synchrony between taps and accents therefore entails perfect accuracy on this measure, and the variance of the average relative phase for a pattern shows the consistency with which a participant adhered to a particular phase relationship. Note that the first analysis technique, which uses more traditional statistics, does not relate produced taps to the onset time of accents whereas the latter technique does.

Although it is reasonable to expect that musically trained participants will perform better in such a task than less skilled individuals, we anticipate that the specific effects of JAS may transcend skill level. Furthermore, because listeners must monitor a melody and selectively anticipate its future accents, the tapping requirement in this task differs from the synchronization/continuation paradigm commonly used in motor production research (eg. Jagacinski, Marshburn, Klapp, & Jones, 1988; Summers, Rosenbaum, Burns, & Ford, 1993; Wing & Kristofferson, 1973a & b). To assess the possibility that listeners are not responding to accent patterns as such, various control melodies lacking melodic and/or temporal accents are also enlisted in the present design.

Experiment 1

METHOD

Participants

Twenty participants were drawn from the Columbus, Ohio area; half were musically experienced (5 or more years of formal training on a musical instrument) and half were not (fewer than 4 years musical training). Experienced listeners were paid 10 dollars for their participation, whereas non-experienced listeners participated in exchange for course credit in an introductory psychology course at the Ohio State University (OSU). Participants' data were eliminated if the following appeared: (1) a tendency to respond on every note of non-baseline patterns; (2) an obvious discrepancy between stated experience level and performance on a post-experiment questionnaire; (3) perfect pitch. Of 11 musically experienced musicians tested, 3 were eliminated, 2 because of perfect pitch and the other because of poor performance on a post test questionnaire that measured musical knowledge. Of 12 non-experienced participants, 4 were eliminated due to a tendency to tap on every note of the sequence.

Stimulus Conditions

Melodic variables. Four melodies were generated from one base melody; all members of this set had transformally equivalent contour accent patterns. All melodies were based on the C Major diatonic set; all were isochronous

TABLE 1
Condition for Experiment I

		Melodic accent periodicity four beats				
		a1	a2	d1	d2	monotone
Temporal accent periodicity	four beat	Concordant patterns				Temporal controls
	6 beat	Discordant patterns				
	no pause	Melodic controls				Baseline

with 36 IOI's of 300 ms (200 ms tone durations). Contour accents (inflection points, or changes in pitch movement direction) of the base melody were constrained in the following way: All contour accents were separated throughout the base melody by four tones, yielding a fixed inter-accent period of 1200 ms; the first contour accent occurred on the 5th tone and thereafter on every 4th tone. Thus, assuming a beat period equal to the IOI, contour accent groups contained a 4 beat period. When possible, the four melodies adhered to the following additional constraints: (1) Two successive contour inflections must engage different pitches, (2) Pitches other than the tonic pitch (C) should occur on inflection points (exceptions were a pattern's initial and final inflection points), and (3) all pitch skips must be less than a perfect fifth.

The base melody was transformed in two ways. The first transformation reversed the base melody's pitch interval distribution; thus, a sequence of tones from the base melody such as E4-G4-B4-C5, containing pitch intervals (in semitones) of +3, +4, +1, would become E4-F4-A4-C5 in the second melody (+1, +4, +3). Because both base and reversal melodies open with ascending intervals, these patterns are denoted by "a1" and "a2," respectively. The second transformation inverted the pitch intervals of "a1" and "a2" melodies. Thus, this transformation produced two descending melodies, "d1" and "d2", respectively, having pitch intervals similar to "a1" and "a2." All "d" melodies began on C5, whereas "a" melodies began on C4, thus equating the range of pitches in ascending and descending patterns. These generative constraints were relaxed if the above rules yielded an awkward sounding musical pattern. All generated melodies preserved the 4 beat melodic contour structure of the original base melody.

Rhythmic variables. To manipulate joint accent structure, two kinds of experimental variations in rhythm were constructed by combining melodic, **m**, with temporal, **t**, accents. The **t** accents were always produced by substituting rests for selected tones in the isochronous melodies; silences were distributed so as to generate **t** accents with periods of either 4 or 6 beats, resulting in two different "temporal structures." Temporal structures based on a metrical grouping of 4 created a concordant JAS with coincident **m** and **t** accents. In these, a one-beat rest replaced every fourth tone, excluding the last tone of the pattern, yielding a mean inter-accent time interval for concordant patterns of 1,200 ms. Temporal structures with a metrical group of 6 created the discordant JAS patterns. In these, the one-beat rest was substituted for every sixth tone, excluding the last tone of the pattern. In discordant patterns, **m** and **t** accents were not always coincident but instead formed a 3:2 polyrhythm, with contour inflection points appearing at unexpected positions given the temporal structure of the pattern, and vice-versa. The mean inter-accent period of discordant patterns was 900 ms. (See Figures 3a and 3b for examples of concordant and discordant patterns used in Experiment 1).

Control sequences. We created three types of control patterns: melodic, temporal and baseline. Four melodic controls corresponded to the four original isochronous melodies; these controls contained only contour accents (a1, a2, d1, and d2). Two temporal controls were monotone (C_4) pitch sequences based on either four (t_4) or six (t_6) beats; these patterns contained no contour accents. Finally, a single baseline control sequence (t0) was an isochronous monotone string of C4 tones, lacking both contour and temporal accents.

Design

The complete design was a 5 (melodic instances) × 3 (temporal accent structure) × 2 (counterbalance order - order of stimuli within each block) mixed factorial, with counterbalance order being the only between subject variable. All within-subject variables are shown in Table 1. Melodic conditions included a1, a2, d1, d2 melodies plus a fifth instance, the monotone control sequence. Thus, the melodic instance variable included four melodies with four-beat melodic contour accents, with a fifth lacking melodic accents. Temporal structure included **t** accents with periods of four- and six-beat periods and a third instance with no temporal accent segmentation. These two within-subject variables (melodic instance and temporal structure) therefore comprise the 15 pattern types for the experiment shown in Table 1. This design assigns the monotone patterns t4, t6, and t0, to the three levels of the temporal accentuation variable where they function as controls for concordant, discordant and melodic controls, respectively. Similarly, the four melodic control patterns are assigned to each melodic instance of both concordant and discordant patterns, as both these joint accented patterns have the same melodic accent structure.

Apparatus and Stimulus Generation

All melodies were created using the 5.0 version of the MIDILAB software package (Todd, Boltz, & Jones, 1989) an IBM PC compatible computer interfaced by a Roland MPU-401 MIDI Processing unit that controlled a Yamaha TX81Z FM Tone Generator set to the "Pan Flute" voice (patch B-12). The MIDILAB package allows for the construction of sound patterns, their organization into blocks of trials, and the recording of participant responses. The sound signal was transmitted to a separate experimental room and amplified using a Rane HC-6 Headphone console. Each participant listened to patterns over AKG-K270 headphones, at a comfortable listening level. Instructions were recorded and played over a cassette recorder.

Tapping responses were produced on a box that included a central "home base" pad surrounded by four optical sensors arranged in a semicircle. Participants were permitted to tap that optical sensor at the most comfortable distance from the preferred finger of their dominant hand; they had to use the same sensor throughout, however. Between taps, participants were told to rest their finger on the home base.

Procedure

Each trial consisted of a single pattern presented twice; the second pattern occurred one second after the first and was raised an octave. A trial began with a high pitched warning tone; a C-major triad then signaled the onset of each pattern within a trial.

The experimental session was divided into five blocks. The first block, a practice block, consisted of 12 trials including both melodic control patterns and joint-accented patterns. For these trials, a new set of melodies was constructed in which the melodic contour points were exaggerated in relation to experimental patterns, with the aim of insuring that listeners were primed to attend to contour as well as temporal accents. This was achieved by increasing the interval skip by one scale step (avoiding the tonic) for the existing inflection points. Melodies a1 and d2 were used. Temporal structure was not similarly exaggerated. Data from these trials were not analysed.

Blocks 2 and 3 consisted of 20 joint accented patterns each: concordant and discordant patterns were randomly intermingled, with the constraint that no more than two trials of one accent structure or melody type followed in succession. Across these two blocks, each of the concordant and discordant versions of the four melodies (a1-d2) were presented. In block four, temporal and baseline controls were randomly intermingled within 15 trials, followed by 10 melodic control trials. Block five consisted of 10 melodic control trials followed by 15 temporal and baseline control trials. Both temporal controls and the monotone baseline melody were presented 10 times per subject, whereas each of the melodic control melodies was presented 5 times per participant.

During a trial, the participants were instructed to "hear out" accents during the first presentation of the sound pattern. An accent was defined as "any tone that stands out from the others around it." They were told to listen for between five and fifteen accents per pattern. When listening to the second presentation, participants were instructed to anticipate the occurrence of accents, and to tap selectively (on an optical sensor) only to onsets of tones they perceived as accented.

An experimental session lasted approximately one and a half hours (a total of 102 trials, including practice) with breaks between each block. Participants were tested either singly or in sets of two. Afterwards, participants completed a questionnaire concerning musical training and knowledge.

RESULTS

The primary data for assessing performance consist of produced inter-tap time intervals and the phase of successive taps with respect to accent onsets (expressed as phase lag within a given referent period). Accordingly, our findings are presented in two sections, devoted to summaries of produced periods and relative phase of taps.

In order to remove artifactual beginning and ending effects, we eliminated any tap occurring earlier than 1400 ms into the pattern (which corresponds to the offset of the second melodic accent) or later than 9500 ms (which corresponds to the offset of the tone preceding the 8th melodic accent). All subsequent analyses proceeded with an omnibus F-test, followed by specific planned comparisons. Although the omnibus test is useful, given our design it is not especially informative with respect to important comparisons; namely, those involving concordant and discordant patterns. In what follows, we therefore focus first on planned comparisons involving the JAS patterns and then consider other comparisons that involving various control patterns. We will relate these planned comparisons to the omnibus F test, which shows how these planned comparisons are revealed in the general design.

Produced periods: Mean and variance

We first considered time intervals that subjects produced in response to a certain pattern. Preliminary analyses focused on the mean and variance of produced time intervals between successive taps for each listener across the n repetitions of a given pattern instance ($n = 5$ or 10 for experimental and control patterns, respectively). Each listener received 15 scores (5 melodic $\times$ 3 temporal accent conditions). Table 2 shows the means and variances, averaged over listeners, in each of the 15 conditions. From

TABLE 2
Mean and variance of produced period by conditions for Experiment 1

		Melodic accent periodicity 4 beats				
		a1	a2	d1	d2	monotone
Temporal accent periodicity	four beat	M = 1360 V = 39, 971	M = 1351 V = 62,491	M = 1436 V = 66,721	M = 1421 V = 65,203	M = 1044 V = 17,764
	six beat	M = 1549 V = 141,096	M = 1546 V = 126,797	M = 1600 V = 74,194	M = 1596 V = 76,016	M = 1280 V = 59,614
	no pause	M = 1603 V = 83,034	M = 1501 V = 117,987	M = 1476 V = 82,093	M = 1569 V = 42,119	M = 961 V = 36,947

these data, we calculated a mean difference score between produced and observed inter-accent period (MD) and a coefficient of variation (CV) for each listener. We focus on the these measures in our primary analyses. With reference to our two general predictions, the MD measure relates to accuracy of produced periods and CV relates to the variability of produced periods.

Mean Difference (MD). Mean inter-tap intervals produced by listeners when tapping to accents are evaluated relative to the mean time interval between accents in the attended patterns. For the majority of the patterns (i.e. concordant, melodic controls and t4 controls), the average inter-accent period is 1200 ms (four beats); for the discordant patterns, however, this interval is 900 ms because it is derived from the polyrhythmic intervals resulting from combining the four beat melodic period (1200ms) with the six beat temporal one (1800ms). The MD analysis rests on the presumption that participants may try to mimic a pattern's inter-accent period. Thus, the MD, or "mean difference," measure subtracts the mean inter-accent period of each pattern from the mean produced period for that pattern for each listener.

Generally, participants overestimated average inter-accent periods in their mean produced periods. This can be seen in Figure 4 where overestimation is absent only with the two temporal control patterns (t_4, t_6). The most important outcome, however, is that overestimation for discordant patterns (mean MD = 673.48 ms) was significantly greater than that for concordant patterns (mean MD = 192.34 ms), as indicated by a planned comparison, $F(1, 18) = 70.67$, $MS_e = 125{,}454$, $p < .001$. Additionally, MDs for concordant patterns were significantly larger than for their t4 controls, $F(1, 18) = 13.48$, $MS_e = 117{,}947$, $p < .01$; in fact, with the t_4 controls, participants consistently produced periods smaller than the pattern's mean inter-accent period (mean MD = −155 ms). A similar result was found in mean MD scores for discordant patterns when compared to their t6 controls (mean MD = −519.09 ms), $F(1,18) = 157.19$, $MS_e = 135{,}000$, $p < .001$. Discordant MDs were significantly greater than those found for melodic controls, $F(1,18) = 16.42$, $MS_e = 239{,}359$, $p < .001$, although a similar difference was not found regarding concordant patterns. The omnibus ANOVA buttressed these findings, indicating a main effect of temporal accent, $F(2, 30) = 13.51$, $MS_e = 204{,}796$, $p < .001$, of melodic accent, $F(4, 60) = 28.17$, $MS_e = 65{,}780$, $p < .001$, and an interaction between the two, $F(8, 120) = 74.32$, $MS_e = 29{,}587$, $p < .001$.

Three unpredicted interactions arise from the omnibus ANOVA. One was a significant interaction between musical training and melodic instance, $F(4,48) = 3.45$, $MS_e = 76{,}460$, $p < .05$; the second was a significant interaction between counterbalance order and melodic instance $F(4,48) = 3.71$, $MS_e = 76{,}460$, $p < .05$. In each of these cases a similar pattern of results occurred: For both non musicians and musicians, all melodically accented melodies (a1, a2, d1, d2) showed consistently higher MDs than temporal controls. However, this effect was greater for nonmusicians than for musicians. Similarly, listeners showed higher MDs for all melodically accented melodies (relative to others) in one counterbalance order than the other. A third unpredicted interaction involved musical training and counterbalance order, $F(1, 15) = 5.38$, $MS_e = 1{,}555{,}596$, $p < .05$. This interaction revealed that musicians produced, on average, smaller overestimations in counterbalance order 2 in comparison with counterbalance order 1, whereas the opposite was true for non musicians. A post-non musicien hoc Tukey's HSD test revealed that musicians produced greater overestimations than non musicians in one counterbalance order ($p < .05$).

These data suggest that at least when tapping to discordant patterns, listeners do not precisely reproduce a pattern's average inter-accent period. However, we evaluate these data with caution because the average MD score for discordant patterns is based on a smaller subtra-

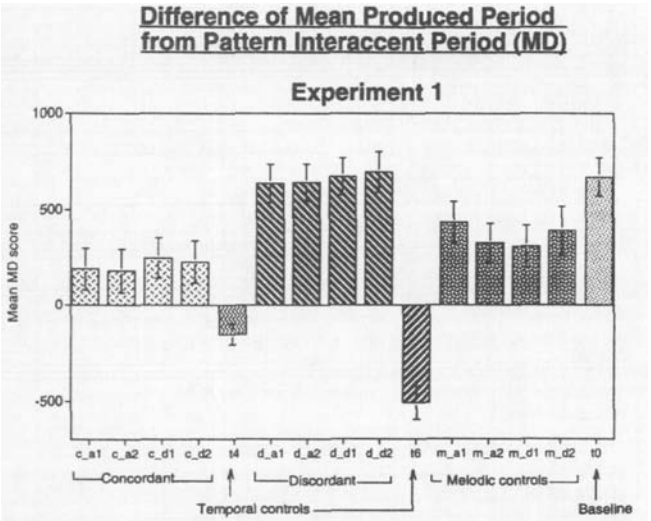

Figure 4. Mean MD scores for temporal accentuation and melodic instance conditions collapsed over counterbalance order and musical training in Experiment 1.

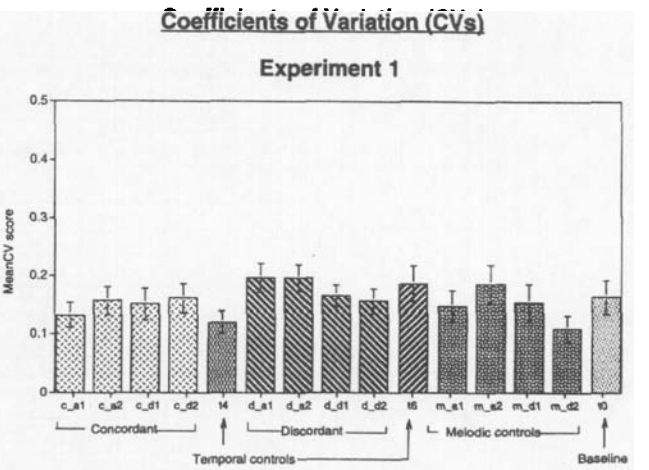

Figure 5. Mean CV scores for temporal accentuation and melodic instance conditions collapsed over counterbalance order and musical training in Experiment 1.

hend (i.e., 900 ms) than that for the concordant patterns (1200 ms). However, the difference in raw mean produced period between concordant and discordant conditions was also significant, $F(1,18) = 10.145$, $MS_e = 125{,}454$, $p < .01$, thus arguing against the point that the MD difference is an artifact of differences their subtrahends.

Coefficient of Variation (CV). The most conservative and informative measure of performance variability involves CV scores. A CV score combines the listener's standard deviation and mean of produced periods for a given pattern (CV = standard deviation / mean). It therefore indexes variability in produced time intervals as a proportion of the mean produced period, thereby eliminating the possibility that variance differences derive from differences in mean produced periods (i.e., in general, longer produced periods often result in greater variance estimates). Figure 5 presents mean CV scores as a function of pattern type (as defined by accent conditions), collapsed over musical skill level and counterbalance order.

On average, discordant patterns yielded higher CV scores than concordant patterns (0.174 vs. 0.148, respectively), particularly for ascending melodies. But overall, this difference was not significant, $F(1, 18) = 1.38$, $MS_e = .017$, $p > .10$. Although in the anticipated direction, this outcome is not consistent with the prediction that concordant patterns should be produced with less variability than discordant ones. Among our planned comparisons, the only significant contrast was between t4 and t6 control patterns, $F(1, 18) = 9.83$, $MS_e = .005$, $p < .01$. The omnibus ANOVA indicated a significant interaction between training and counterbalance order, $F(1, 15) = 4.74$, $MS_e = .067$, $p < .04$, such that musicians tended to produce less variable periods in counterbalance order one whereas non musicians tended to produce more variable periods in this counterbalance order. Additionally, the four way interaction between counterbalance order, training, temporal condition and melodic condition was significant, $F(8, 120) = 2.71$, $MS_e = .007$, $p < .01$. In short, the predicted pattern of results (concordant less variable than discordant) was found only for musicians in one counterbalance order and nonmusicians in the other counterbalance order.

Although CV scores do not differ reliably as a function of JAS, less conservative variability scores, such as raw variance, do show predicted effects (Table 2). Concordant patterns elicit significantly lower variance than discordant ones, $F(1, 18) = 5.51$, $MS_e = 128 \times 10^8$, $p < .05$, and both were found to yield significantly higher variability than their respective temporal control patterns, $F(1, 18) = 10.30$, $MS_e = 245 \times 10^7$, $p < .01$ for concordant vs. t4; and $F(1, 18) = 4.72$, $MS_e = 533 \times 10^7$ $p < .05$ for discordant vs. t6. However, these results are qualified by a number of effects, among these a difference between ascending and descending melodies $F(1, 18) = 7.66$, $MS_e = 495 \times 10^7$, $p < .05$. The predicted difference between concordant and discordant melodies achieves significance for only ascending melodies (a1, a2).

In sum, participants overestimated inter-accent periods marked out by different accent types in various experimental and control patterns, with longer periods, on average, appearing in discordant than in concordant melodies. In addition, although discordant patterns tend to produce more variable tapping intervals than concordant ones, analyses of CV scores indicate that these differences are not statistically significant.

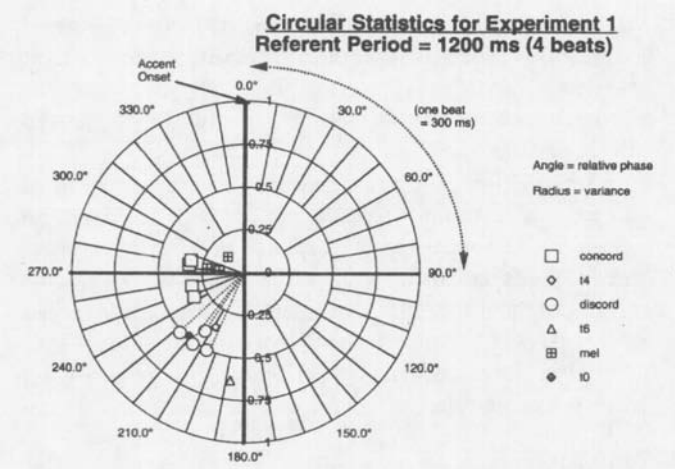

Figure 6. Circular statistics for Experiment 1. A phase angle reflects the mean phase relative to melodic accent onset time (0°) for each temporal accentuation and melodic instance condition. These are displayed within a reference period of 1200 milliseconds (4 beats = 360°). A beat, or one IOI (300ms) = 90 degrees; tone duration = 200ms (or 60°). Mean angular variance is shown along radial spokes for each of the 15 conditions (see legend).

Relative Phase Timing: Mean and variance

The actual timing of taps relative to the melodically accented elements in a pattern speaks to how well people coordinate their taps in order to synchronize with the four-beat melodic accent pattern that is present in all experimental patterns. Thus, a tap close to the onset of a 200 ms accented tone represents good synchronization, hence high accuracy. The time period specified by a recurring accent determines a polar coordinate system in which each angular degree represents temporal distance from the accent's onset. The time of any tap is expressed as relative phase — an angular deviation from accent onset. Using relative phase, perfect synchrony corresponds to a phase of zero within a referent period.

In the present experiment, if the melodic accent period of four beats (present in all but the monotone control sequences) is taken as a temporal reference frame of 1200 ms, then we can determine two angular statistics: the mean and variance of relative phase[4]. In discordant patterns, by definition two possible reference frames exist, one corresponding to the melodic accent periodicity (4 beats) and the other to the temporal accent periodicity (6 beats). If temporal accents provide the relevant time frame in discordant patterns and in control sequence (t6), then the appropriate analysis period is six beats (1800 ms). For purposes of comparison, we use the same referent period for all patterns in the experiment. Our primary analysis relies on the more prevalent four-beat period; however, where appropriate we report data based on a six-beat period. Consequently, our primary analysis uses angular descriptive statistics to assess how accurately (mean relative phase) and variably (variance of relative phase) people respond to the same melodic accents when they occur in different temporal accent contexts.

Mean Relative Phase. Figure 6 presents the mean relative phase angle within a polar coordinate graph based on a referent period of four beats (1 degree = 3.33 ms; 90 degrees = 300 ms or one beat). The accuracy with which listeners selectively time their taps is shown as mean relative phase (angular deviation from the accent onset) for each condition, averaged over the five repetitions of each pattern musical training level and counterbalance order. The planned comparison between concordant and discordant patterns using mean relative phase was not significant, $F(1, 19) = 3.18$, $MS_e = 9.35$, $p < .10$. In fact, the only planned comparison yielding a significant result was that discordant patterns had significantly larger values of relative phase than did melodic controls, $F(1, 19) = 4.68$, $MS_e = 12.52$, $p < .05$.

An ANOVA on listener's mean relative phase scores indicated that the only significant effect came from the manipulation of temporal accents, $F(2, 32) = 3.85$, $MS_e = 8.07$, $p < .05$. Post hoc comparisons indicated that this effect reflected the fact that listeners produced larger relative phase values for patterns with no temporal accents than for those with 6 beat temporal accent periods (Tukey's HSD $p < .05$).

It is important to note that large values of relative phase do not necessarily indicate a long delay (i.e., lag) in responding to an accented tone. To the contrary, these large values may result from anticipating the onset of an accent. To tease this out, we examined frequency distributions of listeners' taps within concordant and discordant patterns. In both these conditions, a disproportionate number of taps occurred roughly at negative phase angle (-15 degrees), relative to zero, indicating that listeners indeed anticipated accent onsets.

In sum, people displayed reasonable levels of accuracy in timing their taps relative to melodic accents. They tended to tap at similar points in a four beat period for all experimental patterns. Concordant and melodic control patterns featured slightly smaller relative phase scores than the three temporal controls and than all but one of the discordant patterns (d2).

4 Angular means and variances for each participant were calculated in the following way. Each tap was converted into an angle according to its lag in milliseconds from the onset of the most recent marker set by the referent period. Sums were then generated for the sine (S) and cosine (C) of each angle (i.e., tap) across each pattern of a given pattern type. Mean relative phase was then found using the arc tangent times (S/C) if values of both S and C are greater than 1. A more detailed account can be found in Fisher (1993, pp. 30-35).

Variance of Relative Phase. Figure 6 also illustrates the average variability with which listeners responded to accent onsets. In this graph, it is shown as the variance of relative phase, or angular variance (ranging from 0 to 1), which is expressed as a radial distance from the polar center for each condition. Points farther from the center reflect more variable conditions.

A planned comparison between concordant and discordant patterns indicated that concordant patterns produced reliably lower angular variance than discordant ones, $F(1,18) = 7.18$, $MS_e = .130$, $p < .05$. Comparisons of the experimental patterns with their respective monotone controls (t4, t6), indicated that concordant patterns did not produce a different level of angular variance than the t4 control sequence. Discordant patterns produced only marginally less variability than the t6 control pattern, $F(1,18) = 3.62$, $MS_e = .092$, $p < .10$, although this becomes significant when one (extreme) participant's data are removed from the comparison, $F(1, 17) = 6.55$, $MS_e = .076$, $p < .05$. For the majority of participants, therefore, average variance of the t6 control was higher than that of discordant patterns. Additionally, both concordant and discordant patterns yielded significantly higher variability than melodic control patterns, $F(1,18) = 8.74$, $MS_e = .103$, $p < .01$ and $F(1, 18) = 18.55$, $MS_e = .198$, $p < .001$, respectively.

Three sources of variance in the omnibus ANOVA proved significant. All confirmed our predictions regarding variability. Two main effects were found, one of temporal accent, $F(2, 30) = 14.17$, $MS_e = .137$, $p < .001$ and melodic instance, $F(4, 60) = 8.48$, $MS_e = .039$, $p < .001$. The temporal accent effect results from higher variability in patterns with a six-beat temporal period relative to those patterns with a four-beat temporal period and patterns with no temporal accents (Tukey's HSD $p < .05$ for both comparisons). The melodic instance effect derives from inclusion of the monotone controls (t6, t4, t0); the monotone patterns lead to more variability in targeting taps to accent onsets than did others, (Tukey's HSD $p < .05$ for each pair comparison) largely because of performance in the t6 and t0 conditions. A two way interaction was found between melodic instance and temporal accent condition $F(8, 120) = 3.11$, $MS_e = .034$, $p < .01$. This finding relates to the aforementioned planned comparisons.

Although angular variance is lower for concordant than for discordant patterns, people are not reliably less variable on the former than with the relevant monotone control (t4). This suggests that, at least with concordant patterns, lower variability in the phasing of taps within concordant patterns could merely be due to the presence of the four beat temporal accent. Arguing against this is the fact that concordant patterns show significantly *more* variability than their melodic controls. In rank order of increasing variability are melodic control, concordant patterns, discordant patterns. This suggests that the presence of two distinct accent types within a pattern affects performance.

Perhaps, then, the most reasonable interpretation of this pattern of results is that the concordant JAS, based on a four-beat temporal accent, does not necessarily *improve* the listener's consistency in targeting taps to melodic accent points in this four beat reference frame, but it does not hurt performance either. By contrast, tapping to accents in the discordant JAS reflects an *interference effect* from t accents relative to concordant and melodic controls.

Finally, the greater variability found with Discordant patterns and the t6 sequence cannot be attributed to the engagement of an inappropriate reference frame for angular statistics computations. If these listeners responded primarily to temporal accents, then their angular variance should be lower when performance is evaluated with the 1800 ms time frame. This is not the case. In fact, in the larger, six-beat, frame the angular variance increases for all sequences. Although it is not surprising that the variability increases in this (inappropriate) frame for concordant and melodic control patterns, it is surprising that angular variance for the discordant patterns also increases. This provides converging evidence that listeners are not dominated by the temporal accent period in discordant patterns. And it means that the average difference in angular variance between concordant patterns, evaluated on a four-beat reference frame, and discordant patterns, evaluated on a six-beat reference frame, is actually *greater* than when both are evaluated on a four beat reference frame (where they were reported above to be reliably different). In other words, the finding that angular variance is minimal with the 1200 ms frame for both concordant and discordant patterns validates our use of it for our primary analysis.

To summarize our findings concerning relative phase, it appears that listeners tapped at largely similar locations for both concordant and discordant JASs, often anticipating future accent onsets. However they were more variable in the latter condition than in concordant patterns. In contrast to the analyses of produced period, therefore, these analyses confirm our prediction regarding variability, but do not support our prediction regarding accuracy.

DISCUSSION

Overall, the pattern of results is fairly consistent with the hypothesis that people are more accurate in matching the produced period to accent periods in concordant than in discordant patterns. Support for this derives from MD

scores. In addition, as predicted people respond with less variability to those JAS patterns based on simpler time ratios. Support for this hypothesis comes from angular variability.

In terms of accuracy, the mean produced period more closely resembled the pattern's inter-accent period in concordant than in discordant patterns. In fact, produced periods for discordant patterns (M = 1572.8 ms) approached a simple average of a four-beat and six-beat period (i.e., 1500 ms) rather than the aforementioned inter-accent period (900 ms) for these patterns. One explanation of this finding is that in the discordant condition some listeners may have favoured a four-beat (melodic) period whereas others may have favoured a six-beat (temporal) period. Auxiliary analyses explored this idea by examining the distribution of produced periods as a function of the number of beats spanned by that period; these analyses partially confirm this explanation, insofar as produced periods for discordant patterns were somewhat more likely to span six beats than were produced periods for concordant patterns. Nevertheless, a mode of four beats obtained for all patterns. With respect to another index of accuracy, mean relative phase, there is little evidence that listeners are more accurate in anticipating accents in concordant than in discordant patterns. Assuming that relative phase angles close to zero index greater accuracy, people were somewhat more accurate with melodic control patterns than others (but not significantly so). Ultimately, it seems that listeners selectively tapped at similar locations for concordant and discordant patterns within a four-beat referent period, but the times spans they produced to reach these locations differed as a function of JAS.

Similarly, our prediction regarding variability is supported by one of the two variability analyses. Although the variability of produced periods, indexed by CV scores, did not differ as a function of JAS, angular variance did. People were less variable in aiming their tapping toward melodic accents with concordant than discordant patterns. In other words, when the melodic accent period is constant an interleaving of incompatible temporal accents in the discordant case interferes with synchronizing taps to the melodic accents. It is unlikely that this difference is due to melodic accents in the discordant condition seeming less prominent than in the concordant because they do not coincide with temporal accents. If this were so, then the angular variance in the concordant patterns would be less than in the melodic controls (which contain no temporal accents); but it is not. In short, less reliable targeting of taps to melodic accents in the discordant conditions appears to result primarily from the juxtaposition of temporal accents at unexpected places in these sequences.

Finally, it is clear that listeners do respond to both melodic and temporal accent patterns in this task. This is evident both with respect to mean produced periods that are sensitive to both kinds of accent periods, and with respect to mean phase lag which indicates that in most conditions listeners are tapping during an accented tone. In conditions where both accents are present (concordant, discordant), performance often reliably differs from that in conditions where fewer accents are present, namely only temporal accents (t4, t6), only melodic accents (melodic controls) or no accents at all (t0). However, these differences are complicated, depending both upon the type of accent and the dependent measure involved. We return to this issue in the general discussion.

Experiment 2

Perhaps differences in tracking performance due to the manipulations of temporal accentuations in Experiment 1 do not arise from the way melodic and temporal accent patterns combine in time. Perhaps, instead, they derive from the greater simplicity of a four-beat temporal accent which reinforces listeners' inherent biases for temporal groups based on some multiple of two (Jones, 1990). One could argue, given the design of Experiment 1, that the mere presence of a four-beat accent period contributes to superior performance found with concordant patterns. In reproduction tasks, there is some evidence that rhythmic accents dominate melodic ones (e.g., Drake & Palmer, 1993). Experiment 2 introduced manipulations designed to address this possibility by "playing against" any inherent biases listeners might have for base 2 time patterns. In Experiment 2 six-beat melodic accent periods were aligned with six beat temporal accents to form concordant JAS patterns. Thus, we employ a similar set of temporal accent patterns (temporal accents based on duple and triple beat periods plus the no accent control), but the melodic line changes.

The set of melodic instances constructed for the second experiment address other concerns as well. In designing these melodies we introduced greater variability into the melodic accent pattern by allowing for two different melodic accent periods, both consistent with a triple rhythm: melodic contour accents occurred with alternating three- and six-beat periods within any given melodic instance. This controls similarity between concordant and discordant patterns in several respects. First, as with discordant patterns, the melodic and temporal accents will not always coincide in these concordant patterns. Second, the inter-accent period of melodic accents is not constant for these melodies. As a consequence, greater inter-accent variability appears in concordant patterns than in Experiment 1. Third, this manipulation holds the disparity between the mean inter-accent period of concordant and

TABLE 3
Mean and variance of produced period by conditions for Experiment 2

		Melodic accent periodicity 6 beats				
		a1	a2	d1	d2	monotone
Temporal accent periodicity	6 beat	M = 1392 V = 159,269	M = 1276 V = 126,059	M = 1365 V = 180,194	M = 1337 V = 153,366	M = 1459 V = 166,524
	4 beat	M = 1461 V = 241,549	M = 1299 V = 223,067	M = 1480 V = 279,941	M = 1514 V = 396,169	M = 1091 V = 60,958
	no pause	M = 1404 V = 325,724	M = 1329 V = 190,391	M = 1306 V = 224,007	M = 1325 V = 265,655	M = 1230 V = 220,409

discordant patterns (i.e., 1,080 versus 830 ms, respectively) within the same range as that of Experiment one (i.e., 1,200 versus 900 ms) which is markedly less than that resulting from a six beat melodic accent period of 1,800 ms (for the concordant patterns). Finally, however, the concordant patterns continue to reflect simpler accent period time ratios than the discordant. These changes permit a stronger test of the same predictions evaluated in Experiment 1.

METHOD

All aspects of Experiment 2 were identical to those of Experiment one with the following exceptions.

Participants

Sixteen participants who had not served in Experiment 1 volunteered for Experiment two in return for credit in an introductory psychology course (OSU). Two groups of eight listeners each were formed following the same criteria for musical experience used in Experiment 1. Using identical criteria for elimination as Experiment 1, data from 20 listeners were eliminated either due to incorrect claims of musical experience or failure to follow tapping instructions.

Stimulus Conditions

A new set of four melodies were transformationally generated from a base melody as in Experiment 1; although the base melody in Experiment 2 did not confine the timing of melodic contour accents to periods of four beats. In Experiment 2, contour inflections were organized that (1) a triple meter would be implied, and (2) a contour accent would never fall at a point where a pause would occur in the discordant patterns. As such, an alternating pattern of three and six beats was derived for the base melody in Experiment 2. In all, base melodies had 10 melodic contour accents. This produced a mean inter-accent period of 1080 ms (about 3.5 beats) for concordant patterns and melodic controls.

Concordant melodies were created by inserting a pause every six beats, starting on the fifth beat of the pattern, excepting the last tone. Thus, concordant melodies featured temporal accents that aligned with 5 of the 10 melodic accents. All temporal accents coincided with a melodic accent, therefore the mean inter-accent period for concordant patterns was the same as for melodic controls (1080 ms). For discordant melodies, temporal accents alternated between period of four and eight beats, creating a mean inter-accent period of 830 milliseconds (slightly under 3 beats). Figure 3c and 3d shows examples of some of the concordant and discordant patterns used in Experiment 2.

RESULTS AND DISCUSSION

The data from Experiment 2 were analysed following the same procedures as in Experiment 1.

Produced inter-accent periods: Mean and variance

As with Experiment one, our primary analyses relied on the mean and variance of each listener's produced inter-accent period to generate MD and CV scores. Table 3 shows group means and variances for each of the 15 melody × temporal conditions.

Mean Difference (MD). Figure 7 shows mean MD for the 15 conditions of Experiment 2. Again, participants' overestimations were significantly greater for discordant patterns (MD = 569.79 ms) than concordant patterns (MD = 236.36 ms), $F(1, 14) = 17.99$, $MS_e = 189{,}395$, $p < .001$. Not surprisingly, these two conditions also differed reliably in MD scores from their respective temporal control patterns, namely t6 (for concordant) and t4 (for discordant), $F(1, 14) = 107.25$, $MS_e = 44{,}816$, $p < .001$; and $F(1, 14) = 52.42$, $MS_e = 408{,}149$, $p < .001$, respectively. Finally, whereas discordant patterns produced significantly larger overestimations than did melodic control

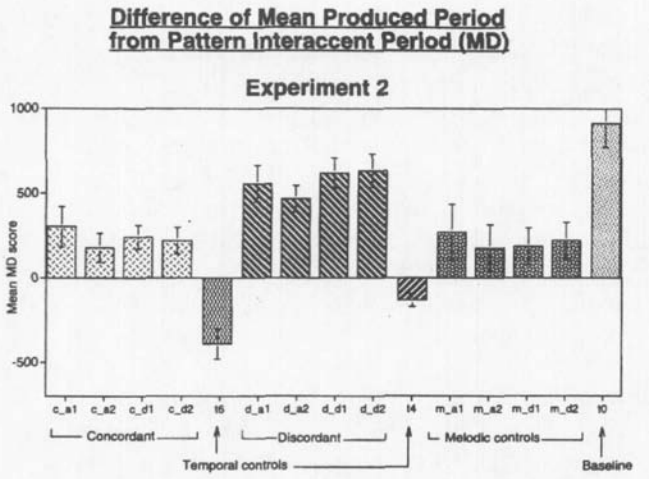

Figure 7. Mean MD scores for temporal accentuation and melodic instance conditions collapsed over counterbalance order and musical training in Experiment 2.

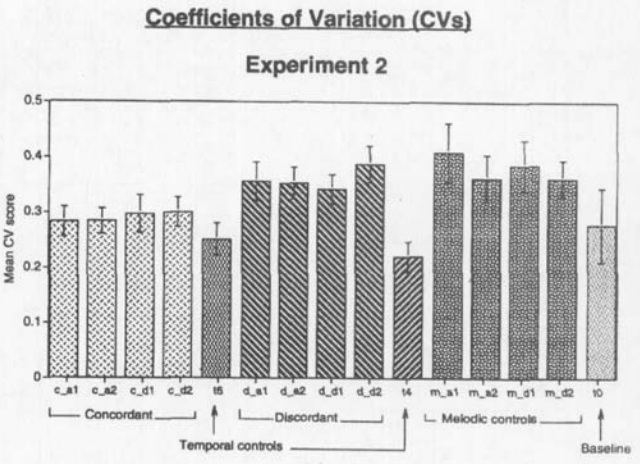

Figure 8. Mean CV scores for temporal accentuation and melodic instance conditions collapsed over counterbalance order and musical training in Experiment 2.

patterns, $F(1, 14) = 12.68$, $MS_e = 275{,}644$, $p < .01$, the comparison between concordant and melodic control patterns did not yield a similar difference.

An omnibus ANOVA confirmed these findings through main effects of melodic condition, $F(2, 22) = 13.42$, $MS_e = 169{,}599$, $p < .001$; temporal accent, $F(4, 44) = 4.27$, $MS_e = 123{,}653$, $p < .01$; and an interaction between these two effects, $F(8, 88) = 20.87$, $MS_e = 90{,}563$, $p < .001$.

As difference scores, the MD outcomes should be interpreted cautiously. For instance, unlike Experiment 1, where mean MD scores mirrored differences observed in mean produced period for concordant and discordant patterns, in Experiment 2 listeners did not differ significantly in mean produced period (1,342.5 ms vs. 1,438.5 ms respectively) as a function of JAS manipulations. This suggests that observed MD differences may indeed result from the subtraction of a smaller subtrahend (830 ms) in discordant than in concordant (1080). Ultimately, the differences obtained by MD in Experiment 2 are only valid if one accepts that listeners are attempting to reproduce an average time period that resembles the subtrahend invoked for a particular pattern. For discordant patterns, this assumption requires that listeners attempt to respond to both temporal and melodic accents. Forthcoming results, involving relative phase, is consistent with this assumption in showing that both kinds of accents affect the variability of responses.

Coefficient of Variation (CV). Do similar produced periods in this case yield similar variability? In Experiment 2, a planned comparison on concordant vs. discordant patterns indicates that the answer is "no." These results are shown in Figure 8. CVs for discordant patterns (mean CV = 0.361) were significantly higher than those for concordant patterns (mean CV = 0.291), $F(1, 15) = 9.72$, $MS_e = .016$, $p < .01$. Discordant patterns also yielded significantly higher CV's than their t4 temporal controls, $F(1, 15) = 13.73$, $MS_e = .018$, $p < .01$, although no difference was found between concordant patterns and their t6 temporal controls. The CV scores for melodic control patterns did not differ from discordant patterns, although they were significantly higher than CV's for concordant patterns, $F(1, 15) = 7.11$, $MS_e = .036$, $p < .05$.

The omnibus ANOVA provided converging support through main effects of temporal accent, $F(2, 24) = 6.03$, $MS_e = .020$, $p < .01$, and of melodic condition, $F(4, 48) = 4.01$, $MS_e = .021$, $p < .01$ although the predicted interaction between these two effects was not confirmed ($F < 1$). Two unpredicted effects were also found. One is a main effect of counterbalance order, $F(1, 12) = 13.66$, $MS_e = .027$, $p < .01$; one counterbalance order yielded higher CV scores than the other. A second interaction effect adds some qualification to the results. It is a three-way interaction among counterbalance order, temporal accent, and melodic condition, $F(8, 96) = 3.24$, $MS_e = .016$, $p < .01$. A number of trends contribute to this interaction. First, one counterbalance order shows a greater difference between concordant patterns and their t6 temporal control. Second, melodic control patterns, as a group, produce higher variability in one counterbalance order than in the other. Finally, in one counterbalance order, the baseline pattern yields higher variability than the other. None of these differences contradict our predictions, as each effect has to do with differences among control patterns. Finally, planned comparisons of concordant versus discordant patterns for raw variance scores parallel those for CV scores.

To summarize, periods produced by listeners in Experiment 2 were closer to the mean inter-accent period

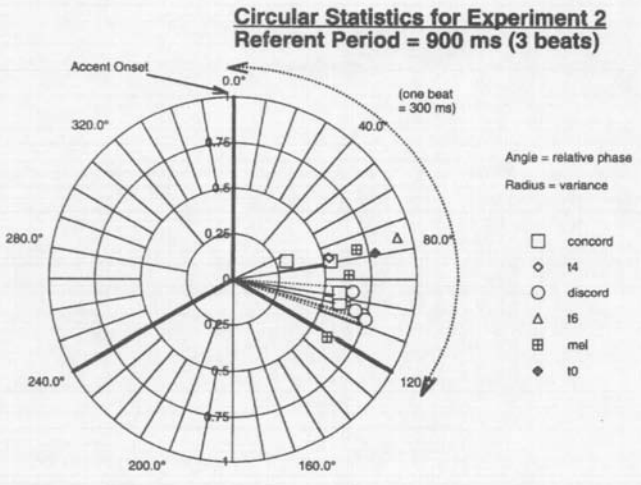

Figure 9. Circular statistics for Experiment 2. Mean angle reflects the mean relative phase of taps for temporal accent and melodic instance conditions. These are shown within a reference period of 900 milliseconds (3 beats = 360°; one beat, an IOI of 300ms = 120°; tone duration is 200ms = 80°. Mean angular variance is shown along radial spokes for each of 15 conditions (see legend).

of the attended-to pattern when the pattern was concordant than when it was discordant. It is possible that listeners attempted to produce similar periods for both kinds of patterns, and in so doing their produced periods for discordant patterns yielded a poor approximation of the mean inter-accent period of these patterns. Additionally, the produced periods of concordant JASs were less variable than those produced for discordant JASs. The distracting effect of the temporal accent period in discordant patterns appears to cause produced periods to be disproportionately more variable than in concordant patterns. This result was clearer in Experiment 2 than in Experiment 1.

Relative Phase: Mean and variance

In Experiment 2, it is possible to assess the mean and variance of relative phase of non-monotone patterns with respect to either of the two melodic accent periods (i.e., alternating patterns of 3 and 6) introduced into these sequences. These periods are themselves variable, meaning that either a three (900 ms) or six beat (1,800 ms) accent period may function as a reference period. We examined both, but we present analyses based on the three-beat melodic period for several reasons. First, the 900 reference frame falls between the average inter-accent periods of concordant (1,080 ms) and discordant (830 ms) patterns, whereas the 1,800 frame is closer to the concordant patterns (thus, differentially favouring the outcome we predict). Secondly, the mean produced period of concordant also favors the 1,800 ms frame whereas the 900 ms frame is closer to the mean produced period of the discordant patterns. It turns out that the 900 ms frame yields minimal angular variance scores for *both* concordant and discordant patterns (relative to both 1800 ms and 1200 ms). Thus, selecting the smaller of the two melodic accent periods as the reference frame introduces no intrinsic bias in favor of the concordant patterns. Finally, although our primary analysis relies on a three beat reference frame, where appropriate we examined discordant patterns both with this frame and with one based on the four beat t accents (wherein they proved more variable than in the 900 ms frame).

Mean Relative Phase. Mean relative phase values, shown as angular lags within the 900 ms polar coordinate system, appear in Figure 9 (1 degree = 2.5 secs; 120 degrees = 300 ms i.e. beat). As in Experiment 1, we found no support for our accuracy hypothesis with this measure. Moreover, no significant differences were found for mean relative phase, either in the planned comparisons or in the omnibus ANOVA.

Variance of Relative Phase. Angular variance for the 15 conditions is also shown in Figure 9 as radial distance from the center of the circle. Once again, discordant patterns yielded significantly more variability, on average, (var. = 0.705) than did concordant patterns (var. = 0.579), $F(1, 15) = 8.99$, $MS_e = .057$, $p < .01$. Discordant patterns also produced significantly lower variance than did their t4 temporal control pattern, $F(1, 15) = 38.16$, $MS_e = .016$, $p < .001$, whereas concordant patterns did not differ from their t6 temporal control. Melodic control patterns did not differ significantly from either of the JAS patterns in angular variance.

The omnibus ANOVA confirmed these results with a main effect of temporal accents, $F(2, 24) = 7.89$, $MS_e = .081$, $p < .01$; a main effect of melodic condition, $F(4, 48) = 5.69$, $MS_e = .025$, $p < .001$; and an interaction of the two, $F(8, 96) = 3.03$, $MS_e = .023$, $p < .01$. A further interaction of skill by counterbalance order, $F(1, 12) = 8.04$, $MS_e = .165$, $p < .05$, results from the fact that musicians in one counterbalance order tended to be somewhat more variable than in the other order whereas the opposite was observed for non-musicians.

To summarize, although the accuracy with which listeners aimed their taps at melodic accents remained unaffected by manipulation of JAS, as reflected by mean relative phase, they were more variable in aiming these taps within discordant than in concordant patterns. This is true in spite of more variable melodic accent periods in melodies of Experiment 2 relative to Experiment 1. Experiment 2 therefore offers strong evidence that variability of tapping time is affected by the complexity of a JAS. Evidence for this conclusion is shown both in the raw

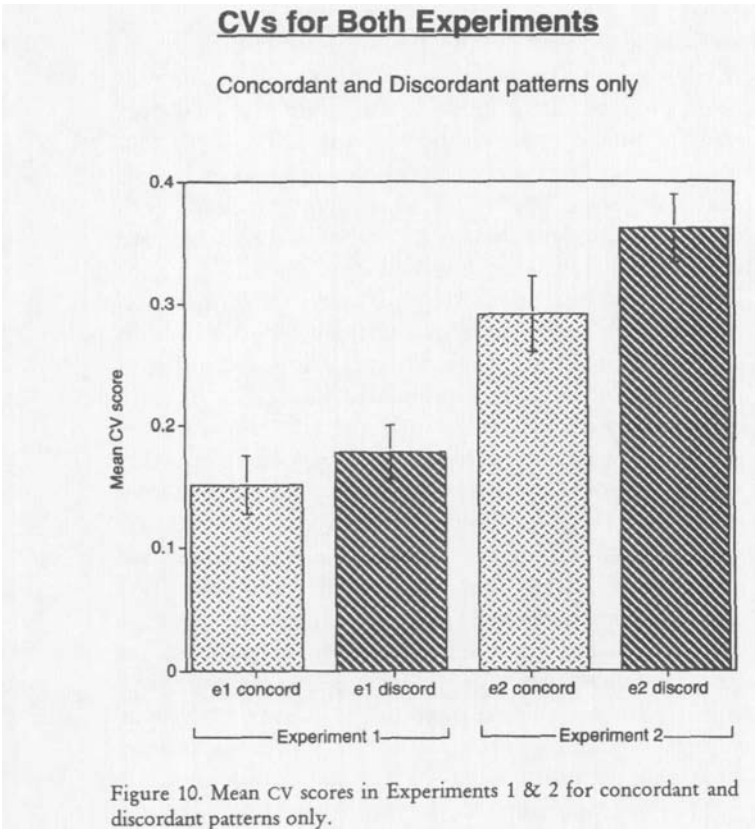

Figure 10. Mean CV scores in Experiments 1 & 2 for concordant and discordant patterns only.

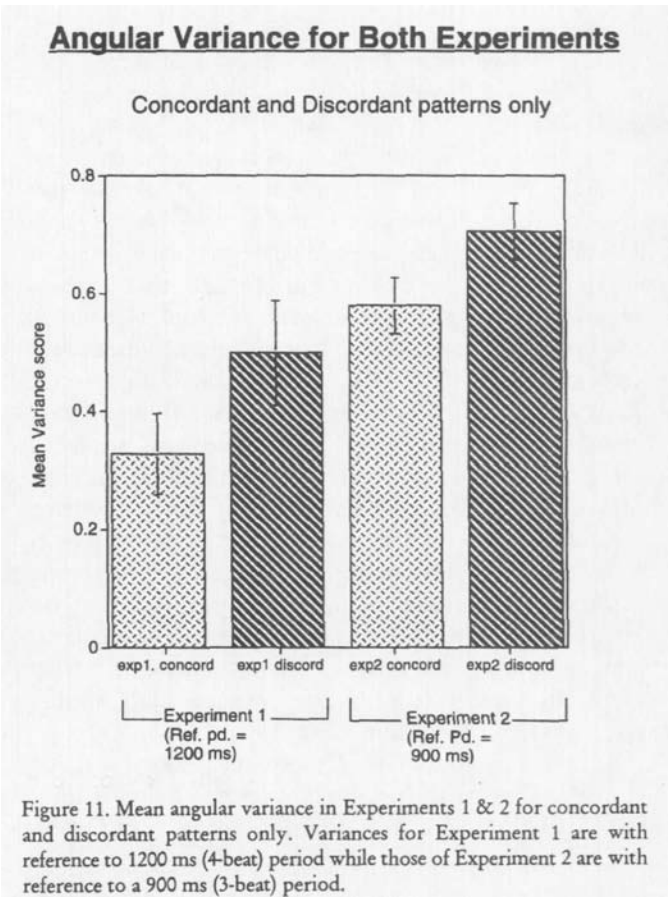

Figure 11. Mean angular variance in Experiments 1 & 2 for concordant and discordant patterns only. Variances for Experiment 1 are with reference to 1200 ms (4-beat) period while those of Experiment 2 are with reference to a 900 ms (3-beat) period.

time periods produced by listeners and by the relationship of these taps to a 1200 ms (4-beat) recurring period.

Experiment 1 versus Experiment 2

Finally, two analyses compared performance variability across the two experiments, using CV and angular variance scores, respectively. The design effectively crossed melodic accent timing (four beats of Experiment 1, three/six beats of Experiment 2) with two of the three temporal accent levels (four and six beats); counterbalance order and training variables were also included. For this analysis we extracted scores pertaining only to experimental patterns where joint accents appear i.e. we focus strictly on difference between concordant and discordant patterns. The CV and angular variance data appear in Figures 10 and 11, respectively. In both, interaction of melodic and temporal accent structures are evident.

Coefficient of Variation (CV) scores are shown in Figure 10 as a function of melodic accent timing (Experiment 1 with 4 beat m periods vs. Experiment 2 with 6 beat m periods) and temporal accent timing (4 beat t accents versus 6 beat t accents) that yielded different versions of concordant and discordant patterns in the two Experiments. The interaction of melodic and temporal accent manipulations was significant, $F(1, 28) = 11.76$, $MS_e = .014$, $p < .01$. Figure 11 presents a corresponding profile of angular variance. Here, the comparable result was also significant, $F(1, 28) = 17.35$, $MS_e = .091$, $p < .001$. Both measures show clearly that greater variability resulted when responding to events in Experiment 2 than in Experiment 1; that is, the main effect of melodic accent (Experiment) was significant in both analyses ($p < .0001$, and $p < .05$, respectively).

Generally, these findings support the hypothesis that concordant patterns produce less performance variability than discordant ones. One higher-order interaction qualifies this generalization with regard to CV scores. The aforementioned four way interaction described in Experiment 1 contributes to the finding that concordant patterns in that study, while yielding lower overall variability, did not produce significantly lower mean CV scores than the discordant patterns.

General Discussion

People are generally sensitive to manipulations of JAS in a selective attending tapping task. This is more evident in analyses of variabilities in response timing in both Experiments 1 and 2 than it is in accuracy of their timed productions.

When we consider the timing accuracy of taps with regard to onsets of melodic contour accents, no substantial

differences emerge as a function of JAS manipulations. We consider relative phase the best indicator of accuracy in this study and it yields no significant differences as a function of JAS manipulations. A second measure of accuracy, the mean difference (MD) between listeners' produced period and pattern inter-accent period, also provides some information about how well people mimicked the succession of inter-accent time intervals within each pattern. This measure suggests that listeners are better at doing this in concordant and control melodies than in discordant patterns. However, the MD measure is only meaningful if we can assume that listeners are actually responding to a pattern's sequence of inter-accent times. In Experiment 1, the modal length of produced periods for discordant patterns approximated 1200 ms (i.e. the four-beat period of concordant patterns), indicating that they were perhaps "tuning" out the six-beat **t** accent. In fact, in both experiments, there is evidence that the period of the melodic accent is a strong influence on performance. Nevertheless, our findings with respect to variability suggest that listeners were not attending merely to melodic accents, but rather were sensitive to the succession of different kinds of accents.

Although listeners' tapping accuracy vis à vis melodic accents suggests that they attempt to tap at similar locations for both concordant and discordant patterns, the variability that listeners display in aiming these taps changes reliably with JAS manipulations. This suggests that listeners were indeed attending to both temporal and melodic accents in a pattern. Moreover, the temporal accents in discordant patterns may have distracted listeners from the generally melodic accents to which they tapped.

Of the two variability measures we used, CV and variance of relative phase, the latter offers the strongest support for the hypothesis that concordant patterns are easier to track attentionally than discordant ones. In Experiment 2, listeners' coefficients of variation (CV scores) revealed less variability in produced inter-accent periods for concordant than discordant patterns, but corresponding differences in Experiment 1 were not significant. By contrast, variability of relative phase was significantly lower for concordant than discordant patterns in both experiments. Experiment 2 confirms that the four beat temporal accent period was not responsible for these effects in Experiment 1. Finally, when data from both experiments are combined (Experiment 1 versus 2), it is clear that both CV scores and relative phase variability are generally less in patterns with simpler JASs.

At a more general level, the finding of an interaction of melodic with temporal accentuation patterns as it appears in the two variability measures also suggests that distinctive melody and rhythm accents form higher-order interdependent accent patterns for listeners. At least in the present task, where the listener had to engage an unfolding auditory event actively, the data suggest that melody and rhythm are not treated independently of one another. In terms of an oscillator model of attending, the oscillators tracking **m** and **t** periodicities are linked in a dependent fashion, i.e., they function as coupled internal oscillators (Large & Jones, 1997). This was generally true for both musicians and nonmusicians. However, differences due to musical skill do emerge in several aspects of the data. For example, in one counterbalance order musicians performed with equivalently low angular variance in both concordant and discordant patterns. It is possible that in this condition musicians responded to a series of early encounters with concordant patterns and refined a selective attending strategy of listening for melodic accents that was carried over to subsequent discordant patterns. We did observe a profile of responding in a few musicians that was consistent with this interpretation. Accordingly, we do not rule out the possibility that training and prior exposure may enable some listeners to "decouple" internal attending rhythms that entrain to melodic and temporal accent periods when this is advantageous (as it may be with discordant musical patterns). However, even with musicians in our study this was not the typical response to discordant patterns.

The interpretation of interactions of melodic and temporal accentuations across experiments (Figures 10 & 11) raises another question. We have suggested that this pattern of findings indicates that the greater variability found with discordant patterns stems primarily from interfering effects of six beat temporal accents on attending to four beat melodic accents; in other words, it is not due to facilitating effects of concordant patterns. We speculate that perhaps two factors play a role in the lack of any facilitating effects of concordant patterns, as gauged by listeners performance on temporal control patterns. First, in both experiments the monotone control sequences provide a challenging baseline of performance because they promote quite stereotypic responding. That is, as a monotone control pattern unfolds a listener does not have to categorize differing pitches under the same distinction of "accent" versus "non-accent." These patterns are maximally predictable. Second, the monotone controls appeared in the last block of a session. A general learning effect, therefore, may also contribute to their low variability. These factors may all operate to lower tapping variability to control patterns and hence raise the threshold measuring any facilitating effects for concordant patterns.

Finally, because our task is a perceptual-motor one, we address the possibility that effects we have observed are strictly due to motor control constraints on synchronized rhythmic tapping (e.g., Summers et al., 1993; Wing &

Kristofferson, 1973a & b). Although Jagacinski et al. (1988) demonstrated that tapping to polyrhythms based on two different pitches is sensitive to pitch differences, synchronized tapping is less common than continuation tapping in existing research. In continuation tasks, the listener is supposed to continue synchronized tapping without the original sounds and preceding IOIs marked out by a monotone sequence. Thus, the patterns involved are similar to our monotone control patterns t4, t6, and t0, but tapping instructions are not. That is, in these studies listeners do not *selectively* tap to accents but rather must tap to all tone onsets. Our task promotes selective attending in time wherein listeners must ignore tones intervening between perceived accents. In this case, the nature of these accents (markers of higher order time intervals) matters, as is evident from our findings. Indeed, if accentuations did not matter then performance on melodic controls would be equivalent to baseline tapping, and performance on concordant and discordant patterns would be equivalent to their respective temporal controls. Finally, it is significant that in our task some listeners even ceased to tap because they found no accents. In other words, although we do not deny that a motor component exists, the data suggest that responses are driven primarily by what a listener is attending to and how it is organized in time.

The question of what constitutes a distinctive or prominent tone (i.e., an accent) for these listeners, in turn, raises the issue of accent strength. Listeners' performance with both melodic controls and experimental patterns confirms that melodic contour inflections function as salient accents for them. This interpretation views contour as supplying accents that function to mark out time intervals within a higher-order time structure. As such, it builds on the conventional description of contour as a particular sequence of "ups" and "downs" in pitch (e.g., Dowling, 1978). It singles out those contour patterns based on successive "ups" and successive "downs", i.e., ascending and descending trajectories, as bearers of accents; in turn, they mark out temporal groups (melodic accent periods). We should be cautious here, however, because as with all musical events, various melodic (and rhythmic) properties can subtly covary; thus, despite our attempts to manipulate primarily contour accents in reasonably sounding tonal melodies, it is possible that other structural properties contribute to a heightening (or lessening) of the impact of some contour accents. By contrast, the data with the monotone control sequences confirms more conventional views of rhythmic segmentation in showing that four beat periods, especially when marked by temporal accents, are exceptionally compelling. Finally, in considering patterns containing both melodic and temporal accents, we find no clear cut evidence that coincidences of these two accent types systematically change the salience of the tone on which both fall.

This research was sponsored in part by NSF grant SBR-9412656 awarded to the first author.

The authors thank Annie Jordan for data collection and helpful comments on an earlier version of this article. Thanks are also due to Rosalee Meyer for helpful comments and to Ralph Barnes for assistance with computer hardware and audio equipment. We are also indebted to Edward Large for alerting us to angular statistics.

Requests for reprints should be directed to Mari Riess Jones, Department of Psychology, 127 Townshend Hall, The Ohio State University, Columbus, Ohio, 43210.

References

Bigand, E. (1997). Perceiving musical stability: The effects of tonal structure, rhythm, and musical expertise. *Journal of Experimental Psychology: Human Perception and Performance, 23*, 808-822.

Bharucha, J.J. (1994). Tonality and expectation. In R. Aiello & J. A. Sloboda (eds.) *Musical perceptions* (pp. 219-239). New York: Oxford.

Bharucha, J. J., & Krumhansl, C.L. (1983). The representation of harmonic structure in music: Hierarchies of stability as a function of context. *Cognition, 13*, 63-102.

Bharucha, J. J., & Stoeckig, K. (1987). Priming of chords: Spreading activation or overlapping frequency spectra? *Perception & Psychophysics, 41*, 519-524.

Boltz, M. (1991). Time estimation and attentional perspective. *Perception & Psychophysics, 49*, 422-433.

Boltz, M., & Jones, M.R. (1986). Does rule recursion make melodies easier to reproduce? If not, what does? *Cognitive Psychology, 18*, 389-431.

Boltz, M., Marshburn, E., Jones, M. R., & Johnson, W. (1985). Serial pattern structure and temporal order recognition. *Perception & Psychophysics, 37*, 209-217.

Brown, H., & Butler, D. (1981). Diatonic trichords and minimal tonal cue-cells. *In Theory Only, 5*, 37-55.

Butler, D., & Brown, H. (1984). Tonal structure versus function: Studies of the recognition of harmonic motion. *Music Perception, 2*, 6-24.

Cooper, G., & Meyer L. B. (1960). *The rhythmic structure of music*. Chicago: University of Chicago.

Dawe, L. A., Platt, J. R., & Racine, R. J. (1993). Harmonic accents in inference of metrical structure and perception of rhythm patterns. *Perception & Psychophysics, 54*, 794-807.

Dawe, L. A., Platt, J. R., & Racine, R. J. (1994). Inference of metrical structure from perception of iterative pulses within time span defined by chord changes. *Music Perception, 12*, 57-76.

Deliege, I. (1987). Grouping conditions in listening to music: An approach to Lerdahl & Jackendoff's grouping preference

rules. *Music Perception, 4*, 325-359.

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*, p. 341-354.

Dowling, W. J., & Fujitani, D. S. (1971). Contour, interval, and pitch recognition in memory for melodies. *The Journal of the Acoustical Society of America, 49*, 524-531.

Drake, C. (1993). Reproduction of musical rhythms by children, adult musicians, and adult nonmusicians. *Perception & Psychophysics, 53*, 25-33.

Drake, C., Dowling, W. J., & Palmer, C. (1991). Accent structures in the reproduction of simple tunes by children and adult pianists. *Music Perception, 8*, 315-334.

Drake, C., & Palmer, C. (1993). Accent structures in music performance. *Music Perception, 10*, 343-378.

Fisher, N. I. (1993). *Statistical analysis of circular data*. Cambridge, UK : Cambridge University Press.

Holleran, S., Jones, M. R., & Butler, D. (1995). Perceiving implied harmony: The influence of melodic and harmonic context. *Journal of Experimental Psychology : Human Perception and Performance, 21*, 737-753.

Huron, D., & Royal, M. (1996). What is melodic accent? Converging evidence from musical practice. *Music Perception, 13*, 489-516.

Jagacinski, R. J., Marshburn, E., Klapp, S. T., & Jones, M. R. (1988). Tests of parallel verus integrated structure in poly-rhythmic tapping. *Journal of Motor Behavior, 20*, 416.

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. (1987). Dynamic pattern structure in music: Recent theory and research. *Perception & Psychophysics, 41*, 621-634.

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

Jones, M.R. (1993). Dynamics of musical patterns: How do melody and rhythm fit together? In T. J. Tighe & W. J. Dowling (Eds.) *Psychology and music: The understanding of melody and rhythm* (pp. 67-92). Hillsdale, NJ : Erlbaum.

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

Jones, M. R., Boltz, M., & Kidd, G. (1982). Controlled attending as a function of melodic and temporal context. *Perception & Psychophysics, 32*, 211-218.

Jones, M. R., & Ralston, J. T. (1991). Some influences of accent structure on melody recognition. *Memory & Cognition, 19*, 8-20.

Klapp, S. T., Hill, M. D., Ryler, J. G., Martin, A. E., Jagacinski, R. J., & Jones, M. R. (1985). On marching to two different drummers: Perceptual aspects of the difficulties. *Journal of Experimental Psychology: Human Perception and Performance, 6*, 814-827.

Krumhansl, C. L., & Kessler, E. J. (1983). Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychological Review, 89*, 334-368.

Large, E. W. , & Jones, M. R. (1997). *The dynamics of attending: How we track time varying events*. Manuscript submitted for publication.

Large, E. W., & Kolen, J. F. (1995) Resonance and the perception of musical meter. *Connection Science, 6*, 177-208.

Massaro, D. W. (1972). Perceptual images, processing time, and perceptual units in auditory perception. *Psychological Review, 79*, 124-145.

Monahan, C. B., Kendall, R. A., & Carterette, E. C. (1987). The effect of melodic and temporal contour on recognition memory for pitch change. *Perception & Psychophysics, 42*, 306-307.

Narmour, E. (1992a). The influence of embodied registral motion on the perception of higher-level melodic implication. In M. R. Jones & S. Holleran (Eds.), *Cognitive bases of musical communication* (pp. 69-90). Washington, DC : American Psychological Association.

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

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

Palmer, C., & Krumhansl, C. L. (1990). Mental representations for musical meter. *Journal of Experimental Psychology: Human Perception and Performance, 16*, 728-741.

Palmer, C., & Krumhansl, C. L. (1987). Pitch and temporal contributions to musical phrase perception: Effects of harmony, performance timing, and familiarity. *Perception & Psychophysics, 41*, 505-518.

Parncutt, R. (1994). A perceptual model of pulse salience and metrical accent in musical rhythms. *Music Perception, 11*, 409-464.

Peretz, I., & Morais, J. (1989). Music and modularity. *Contemporary Music Review, 4*, 277-291.

Povel, D. J. (1981). Internal representation 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.

Povel, D. J., & Okkerman, H. (1981). Accents in equitone sequences. *Perception & Psychophysics, 30*, 565-572.

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

Summers, J. J.,Rosenbaum, D. A., Burns, B. D., & Ford, S. K. (1993). Production of polyrhythms. *Journal of Experimental Psychology: Human Perception and Performance, 19*, 416-428.

Thomassen, J. M. (1983). Melodic accent: Experiments and a tentative model. *Journal of the Acoustical Society of America, 71*, 1596-1605.

Todd, R., Boltz, M., & Jones, M. R. (1989). The midilab auditory

research system. *Psychomusicology, 8*, 17-30.

Watkins, A. J. (1985). Scale, key and contour in the discrimination of tuned and mistuned approximations to melody. *Perception & Psychophysics, 37*, 275-285.

Wing, A. M., & Kristofferson, A. B. (1973a). The timing of interresponse intervals. *Perception & Psychophysics, 13*, 455-460.

Wing, A. M., & Kristofferson, A. B. (1973b). Response delays and the timing of discrete motor responses. *Perception & Psychophysics, 14*, 5-12.

Woodrow, H. (1909). A quantitative study of rhythm. *Archives of Psychology, 18*, 1-66.

Sommaire

La structure d'accent commun (SAC) utilise les relations temporelles entre les différents accents d'un ensemble mélodique comme indices de sa complexité. Deux expériences ont permis d'étudier l'effet d'un indice de SAC et le rapport entre l'accent et la période sur la capacité des sujets à suivre attentivement la mélodie. Un modèle de rapport entre l'accent et la période compare la récurrence périodique d'un type d'accent, tel l'accent mélodique (changement dans l'orientation de la courbe de niveau du ton, par exemple) à celle d'un autre, tel l'accent temporel (une tonalité suivant une pause, par exemple). Lorsque la récurrence périodique est en rapport simple avec la période d'accents temporels (1:1, 1:2., par exemple), les ensembles de ces expériences sont « concordants ». Ils sont « discordants », lorsque les périodes d'accents mélodique et temporel ont un rapport plus complexe entre la période et l'accent (3:2, par exemple). Au cours de l'Expérience 1, les ensembles concordants comportaient une période d'accent mélodique à quatre temps combinée à une période d'accent temporel à quatre temps. De même, au cours de l'Expérience 2, une période d'accent mélodique à six temps a été combinée à une période d'accent temporel à six temps. Pour ce qui est des ensembles discordants de l'Expérience 1, la période d'accent mélodique à quatre temps a été combinée à la période d'accent temporel à six temps, et pour l'Expérience 2, la période d'accent mélodique à six temps, à la période d'accent temporel à quatre temps. Chaque expérience comportait des ensembles de contrôle supplémentaires isolant chacun des types d'accent (mélodique et temporel).

Au cours de chacune des expériences, les participants retraçaient les ensembles lors d'un test d'écoute anticipée. Dans ce test, une mélodie était présentée deux fois. Lors de la première présentation, le sujet devait trouver les notes qui ressortaient. Pour la deuxième présentation, il devait prévoir le locus temporel de chaque accent et appuyer sur un capteur optique au début de chacun d'eux. Les données ont été analysées en fonction des intervalles produits lors d'un essai donné et la position du signal par rapport à la période d'accent dominante de cet ensemble. Nous avions prévu que les ensembles concordants donneraient des résultats moins variables et qu'ils refléteraient plus précisément la synchronisation avec les débuts d'accent que les ensembles discordants. Les résultats ont révélé que l'écoute était plus variable dans les ensembles SAC discordants que dans les ensembles concordants, et cela pour les deux mesures. Nous n'avons toutefois rien découvert quant à nos prédictions sur la précision. À la limite, nous suggérons qu'assister à un événement musical est déterminé par la structure temporelle de ce dernier tel que produit par le rythme relatif des éléments accentués. La structure SAC offre un moyen de caractériser ces relations de temps et peut ainsi permettre de mieux comprendre la complexité rythmique d'une mélodie.