
The Role of Melodic and Rhythmic Accents in Musical Structure
Author(s): Peter Q. Pfordresher
Source: *Music Perception: An Interdisciplinary Journal*, Vol. 20, No. 4 (Summer 2003), pp. 431-464
Published by: University of California Press
Stable URL: http://www.jstor.org/stable/10.1525/mp.2003.20.4.431
Accessed: 31-10-2016 19:06 UTC

Music Perception
Summer 2003, Vol. 20, No. 4, 431–464

The Role of Melodic and Rhythmic Accents in Musical Structure

PETER Q. PFORDRESHER
The University of Texas at San Antonio

Two experiments investigated the perception of melodic and rhythmic accents in musical patterns. Musiclike patterns were created in which recurring melodic and/or rhythmic accents marked higher order periods that, when both accents were present, could differ in terms of period and/or phase according to the construct of joint accent structure (M. R. Jones, 1987). Listeners were asked to indicate the location of accents in these patterns by tapping to tone onsets. Each experiment pursued two main questions. First, are accents, as manipulated, salient to listeners? Second, do listeners track higher order time spans formed by melodic and rhythmic accents in a way that shows a sensitivity to interrelationships between melody and rhythm? Results supported affirmative answers to these questions in analyses of tapping locations and time spans between taps, respectively. Furthermore, results suggested that accents function as temporal landmarks that listeners can use when tracking the time structure of musical patterns, and that the complexity of this time structure arises from higher order time spans marked by different types of accents.

Received May 30, 2002, accepted March 10, 2003

MUSICAL patterns are usually structured in ways that allow listeners to anticipate future events (e.g., an individual tone), although events often violate these expectations because they deviate from a melody's implied trajectory (e.g., Meyer, 1956; Narmour, 1990). The current research focuses on the way in which such deviations create accents on certain tones (Jones, 1987, 1993) and function as temporal landmarks that guide expectancies. Furthermore, such "phenomenal" accents may provide part of the input to more abstract, expectancy-based accents like metrical accents (Lerdahl & Jackendoff, 1983; see also Jones, 1987; Jones & Boltz, 1989; Jones & Pfordresher, 1997). The current research addressed two main is-

Address correspondence to Peter Q. Pfordresher, Department of Psychology and Institute for Music Research, The University of Texas at San Antonio, 6900 North Loop 1604 West, San Antonio, TX 78249. (e-mail: ppfordresher@utsa.edu)

ISSN: 0730-7829. Send requests for permission to reprint to Rights and Permissions, University of California Press, 2000 Center St., Ste. 303, Berkeley, CA 94704-1223.

sues relating to these possible roles for accents. First, the degree to which listeners hear accents at certain hypothesized locations was assessed. Second, the current experiments tested the degree to which temporal relationships among accents contribute to a musical pattern's higher order time structure by applying the construct of joint accent structure (Jones, 1987).

What Constitutes an Accent?

Figure 1 shows a pitch/time trajectory, in which successive pitches are represented by horizontal lines with numbers below to indicate serial order. It outlines the two kinds of accents that this article focuses on: melodic (**m**) accents arising from pivot points in melodic contour, and rhythmic (**r**) accents arising from pauses. Figure 1 also highlights disagreements in past research concerning the specific locations of these accents. The current study attempted to clarify accent locations by having listeners tap in synchrony with onsets of accented tones, based on memory for a melody from a single previous exposure. This technique resembles those used in other studies that have required listeners to synchronize with a regular period in melodies, such as "the beat" or the beginnings of measures (e.g., Drake, Jones, & Baruch, 2000; Drake, Penel, & Bigand, 2000; Large, Fink, & Kelso, 2002; Snyder & Krumhansl, 2001; Vos, van Dijk, & Schomaker, 1994). However, the present study differs critically from such past work in that listeners were not instructed to tap regularly; instead they were instructed to select *any* notes that "stand out from the rest" (cf. Cooper & Meyer, 1960).

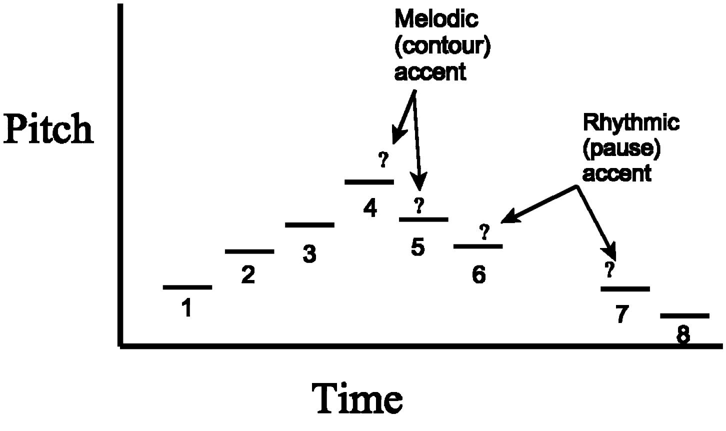


Fig. 1. Examples of melodic and rhythmic accentuation, using a pitch space vs. time coordinate system. Numbers beneath horizontal dashes indicate the serial position of successive tones in a melody. Arrows indicate possible locations of accents, with question marks indicating the uncertainty in accent locations found in past literature.

The specific locations at which structural deviations create accents are difficult to clarify experimentally, despite the intuitive appeal of previous suggestions (see Handel, 1989, for a review). This difficulty arises from the unfolding nature of music: once an accent is experienced, it has passed. It is not surprising, therefore, that many past experiments investigating the role of accents have relied on post-hoc inferences about which tone onsets were heard as accented. Perhaps because such procedures cannot specify the precise locations of accents, different studies have made different claims about where accents occur. One example involves pauses. Although the second tone clearly receives an accent when pauses segment groups of two tones (Povel & Okkerman, 1981), studies disagree about where accents occur when pauses segment groups of three or more tones. Some suggest that tones preceding pauses are accented (Drake & Palmer, 1993; Drake, Dowling, & Palmer, 1991; Fraisse, 1982), which would imply that the pause in Figure 1 would create an **r** accent on Tone 6. However, other researchers have speculated that tones both preceding and following the pause are accented (Povel & Essens, 1985, Tones 6 and 7 in Figure 1), or that the tone following a pause is accented (Jones, 1987; Jones & Pfordresher, 1997; Tekman, 2002; Tone 7 in Figure 1).

Similar disagreements have arisen with respect to accents formed by changes in melodic contour. Some researchers have claimed that the note just *preceding* a change in direction of pitch movement, the "pivot note," is heard as accented (e.g., Huron & Royal, 1996; Jones & Pfordresher, 1997; Thomassen, 1982; corresponding to Tone 4 in Figure 1). Others have claimed that the note *following* the pivot receives accentuation (e.g., Bigand, 1997; Boltz & Jones, 1986; corresponding to Tone 5 in Figure 1). Overall, evidence supporting the isolation of contour accents at pivot points is stronger, given the compelling evidence of Thomassen (1982), and this definition will be assumed as correct here, although further confirmation is warranted.

Two accent types were manipulated in the current study, although many kinds of accents are possible (Handel, 1989; Lerdahl & Jackendoff, 1983). Here, accent "type" refers to an accent that results from change along a single acoustic dimension. The current study used melodic (resulting from deviations from the context of pitch motion in a musical pattern) and rhythmic (resulting from deviations from the context of time spans marked by successive tone onsets and offsets) accent types. Melodic (**m**) accents were created by pairing changes in pitch direction with local increases in the semitone distance between successive pitches. Rhythmic (**r**) accents were created by inserting pauses in melodies. Accent types, particularly **r** accents, were chosen in part to disambiguate the locations of these accents. I compared the accentual effect of two kinds of melodic accent with one kind of rhythmic accent because past research suggests that rhythmic accents are more salient (Drake & Palmer, 1993; Drake et al., 1991; Halpern, Bartlett, & Dowling, 1998; Huron & Royal, 1996; Monahan & Carterette, 1985; Snyder & Krumhansl, 2001; but see Dawe, Platt, & Racine, 1993, 1994 for an exception when harmonic accents are pitted against rhythmic accents).

Accents and Higher Order Time Structure

A guiding hypothesis of the current research is that the time spans defined by successive accents contribute to the perceived higher order structure of a musical pattern. In other words, patterns are structurally simple when accents form predictable higher order rhythms (cf. Boltz, 1993; Essens, 1995; Essens & Povel, 1985; Jones & Boltz, 1989; Povel & Essens, 1985). Such possibilities were assessed in the current research through examinations of time spans between taps. Although listeners' instructions emphasized the selection of individual accents, rather than the extraction of a regular period, it was predicted that the time spans separating taps would reflect the regularity of a pattern's accent structure. Moreover, listeners may follow similar principles when tapping to higher order rhythms formed by recurring accents as they do when tapping to lower order rhythms formed by successive event onsets (e.g., Povel, 1981; Povel & Essens, 1985).

Accent structure, in general, refers to the contribution of time spans between successive accents to the higher order temporal structure of a musical pattern, assuming that these time spans are at least semiperiodic. Period and phase characterize the structure of these time spans (cf., Large & Jones, 1999; Povel & Essens, 1985; Repp, in press). *Accent period* in this paper refers to the number of beats separating recurring accents of a given type (**m** or **r**). Beats are defined as the shortest time spans between tone onsets in a pattern, for the present purposes (cf. Desain, 1992; Temperley, 2001). *Accent phase* refers to the position at which regular cycles of accents begin in relation to the start of a pattern. In-phase accent structures occur when the beginning of a pattern initiates a regular cycle of accents, whereas phase shifts occur when regular cycles start after the pattern begins (which may imply an anacrusis to the listener). Because of evidence that listeners may process melodic and rhythmic information separately (e.g., Herbert & Peretz, 1997; Palmer & Krumhansl, 1987a, 1987b; Peretz & Morais, 1989; Thompson, Hall, & Pressing, 2001), it is assumed that listeners initially process the time spans separating accents of one type (e.g., between **m** accents) separately from the time spans separating accents of another type (e.g., between **r** accents). Most musical patterns, of course, feature multiple accent types. In such cases, accent structures for single accent types form constituent structures, which combine to form a pattern's *joint accent structure* (JAS).

The construct of JAS (Boltz & Jones, 1986; Jones, 1987, 1993; Jones & Pfordresher, 1997) characterizes the time structure of a musical pattern as a function of relationships between time spans formed by each constituent accent type. Figures 2A–2D show examples of JASs distinguished by pe-

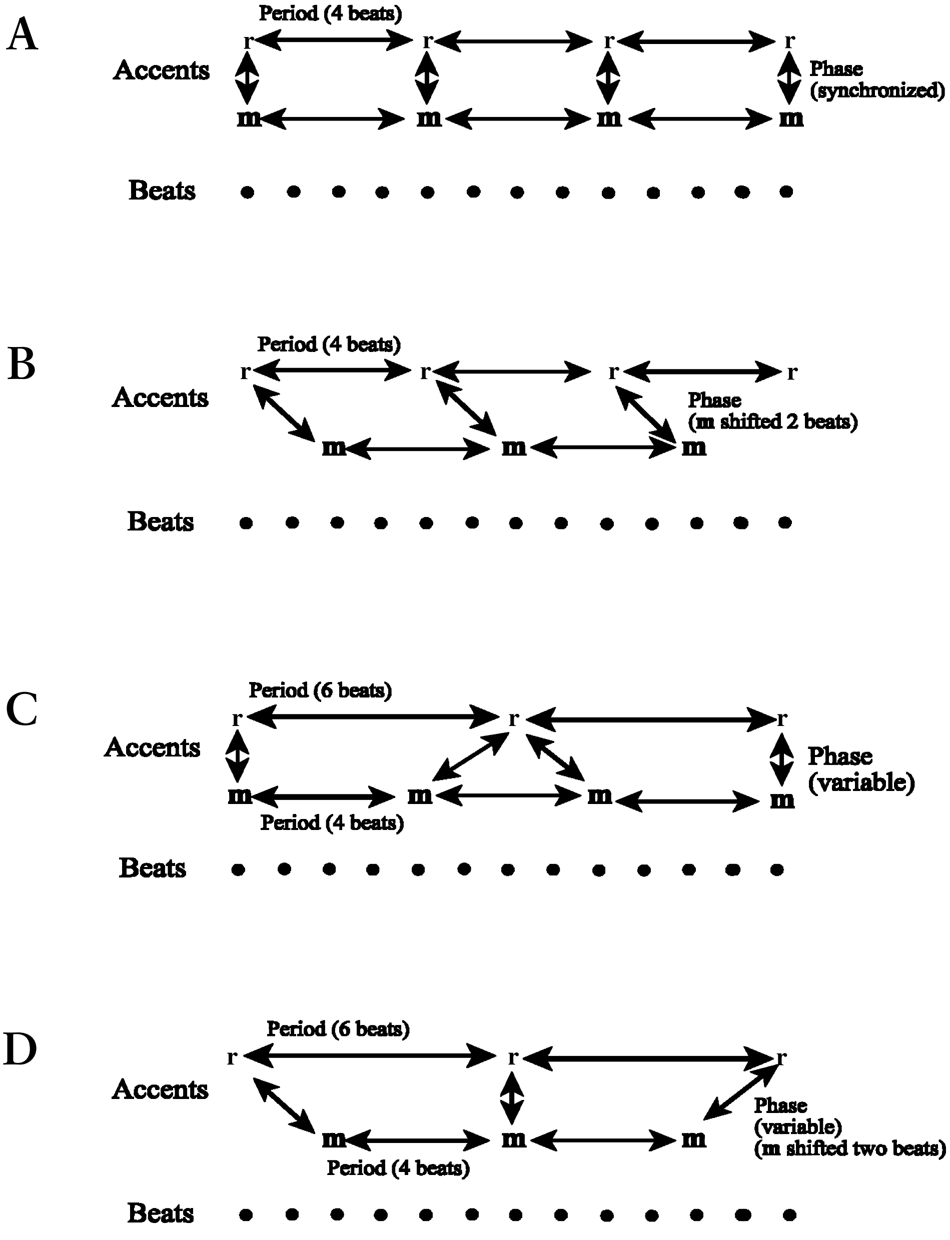


Fig. 2A–D. Schematic depictions of joint accent structure. See text for further description.

riod and phase relationships among constituent accent structures. Relationships between accent periods for each type are indexed by the ratio of one accent's period to the other, called *accent period ratio*. The constituent **m** and **r** accent periods for JASs outlined in Figures 2A and 2B both recur every 4 beats to form a simple (integer) accent period ratio 4:4, such accent period ratios form *concordant* patterns, whereas *discordant* patterns result from noninteger accent period ratios (Yeston, 1974 used the terms "consonant" and "dissonant" to describe similar relationships between pitch and rhythm). Figures 2C and 2D show discordant patterns that result when a 4-beat **m** accent period is combined with a 6-beat **r** accent period. *Accent phase* indexes the degree of synchrony between accent onsets of different types and relates to accent phasing of constituent accent structures. Each **m** accent in the concordant JAS in Figure 2A is synchronized with each **r** accent, whereas **m** accents in Figure 2B are phase-shifted relative to **r** accents by 2 beats. Similar **m** accent phase-shifting occurs in JASs of Figures 2C and 2D, although accent phase relationships become variable in discordant structures.

It was predicted that the variability of tapping responses to accents will reflect the complexity of a pattern's JAS. Accent period ratio (rather than accent phase) was predicted to provide the best JAS index of pattern complexity. These predictions stem in part from the theoretical idea that higher order pattern structure results primarily from embedded periods that form frequency ratios (cf. Jones, 1976; Yeston, 1974). Phase relationships, on the other hand, may direct attention on a local level while not affecting the overall perceived structure as strongly. Furthermore, some evidence from past research suggests that listeners are more sensitive to conflicting melodic and rhythmic relationships when these conflicts are based on period ratios. Studies of memory for music, for instance, document interference when time spans between pauses form periods that do not match the period suggested by melodic organization (Boltz & Jones, 1986; Deutsch, 1980). Furthermore, Jones and Pfordresher (1997) found more variable tapping to discordant patterns in a study of tapping responses to JAS patterns (similar to the current study), although they did not manipulate accent phase. Conversely, studies that vary phase relationships between the melodic and rhythmic structure of musical patterns appear to support independence of melodic and rhythmic information (Palmer & Krumhansl, 1987a, 1987b). In some cases, results that have suggested independence may be influenced by the relative salience of melodic and rhythmic information, with rhythm contributing more strongly to perceived temporal structure (e.g., Drake & Palmer, 1993; Monahan & Carterette, 1985; Snyder & Krumhansl, 2001).

This article reports two experiments that examined tapping responses to patterns with JASs, as well as control patterns that included constituent accent structures. JAS patterns varied relationships between these constituent structures according to both period and phase. Two primary hypotheses were explored. First, it was hypothesized that participants would tap at predicted accent locations (melodic contour pivots, and before and/or after pauses). Second, it was hypothesized that accent period ratio would determine the degree to which participants produced regular time spans between taps.

Experiment 1

In Experiment 1, participants tapped to accents in patterns with different accent structures. JAS patterns were created by combining the melodic and rhythmic accent structures of control patterns. In melodic control patterns, melodic (**m**) accents occurred every 4 beats when present (in this article, 1 beat = 300 ms). Rhythmic control patterns were monotonic and featured recurrent pauses (generating rhythmic, **r**, accents) every 4 beats or 6 beats. Concordant JAS patterns resulted when **m** accents were combined with the 4-beat **r** accent period, and discordant JAS patterns resulted when **m** accents were combined with the 6-beat **r** accent period. Phase relationships between accents in JAS patterns were varied by shifting each **m** accent by 2 beats on half the patterns, which generated phase conditions p1 and p2. Figure 3 shows examples of JAS patterns in music notation. Positions of **r** accents are not shown in Figure 3 because of the uncertainty regarding the accenting nature of pauses.

Also included in Experiment 1 were control patterns that consisted of only **m** or only **r** accent patterns, the former containing no pauses and the latter containing no pitch changes. These were used to assess the degree to which the integration of **m** and **r** accents in a JAS pattern affects listeners' responses, as well as the degree to which accent contributes independently (e.g., Herbert & Peretz, 1997; Jones & Pfordresher, 1997; Monahan, Kendall, & Carterette, 1987; Palmer & Krumhansl, 1987a, 1987b; Tekman, 2002). A final baseline pattern contained a sequence of monotonic pitches without pauses.

METHODS

Subjects

Twenty-four subjects from the Columbus, Ohio, area were recruited; eight were eliminated for failure to follow instructions (e.g., failure to tap to certain patterns, or consistently failing to tap to tone onsets).[1] Of the remaining subjects, eight were musically experi-

---

[1] The task was long and unfortunately required more vigilance than many participants were willing to offer. It was clear, however, that the remaining participants attended to the task. Furthermore, follow-up analyses demonstrated that responses of excluded subjects qualitatively resembled those reported here, on trials for which taps were produced.

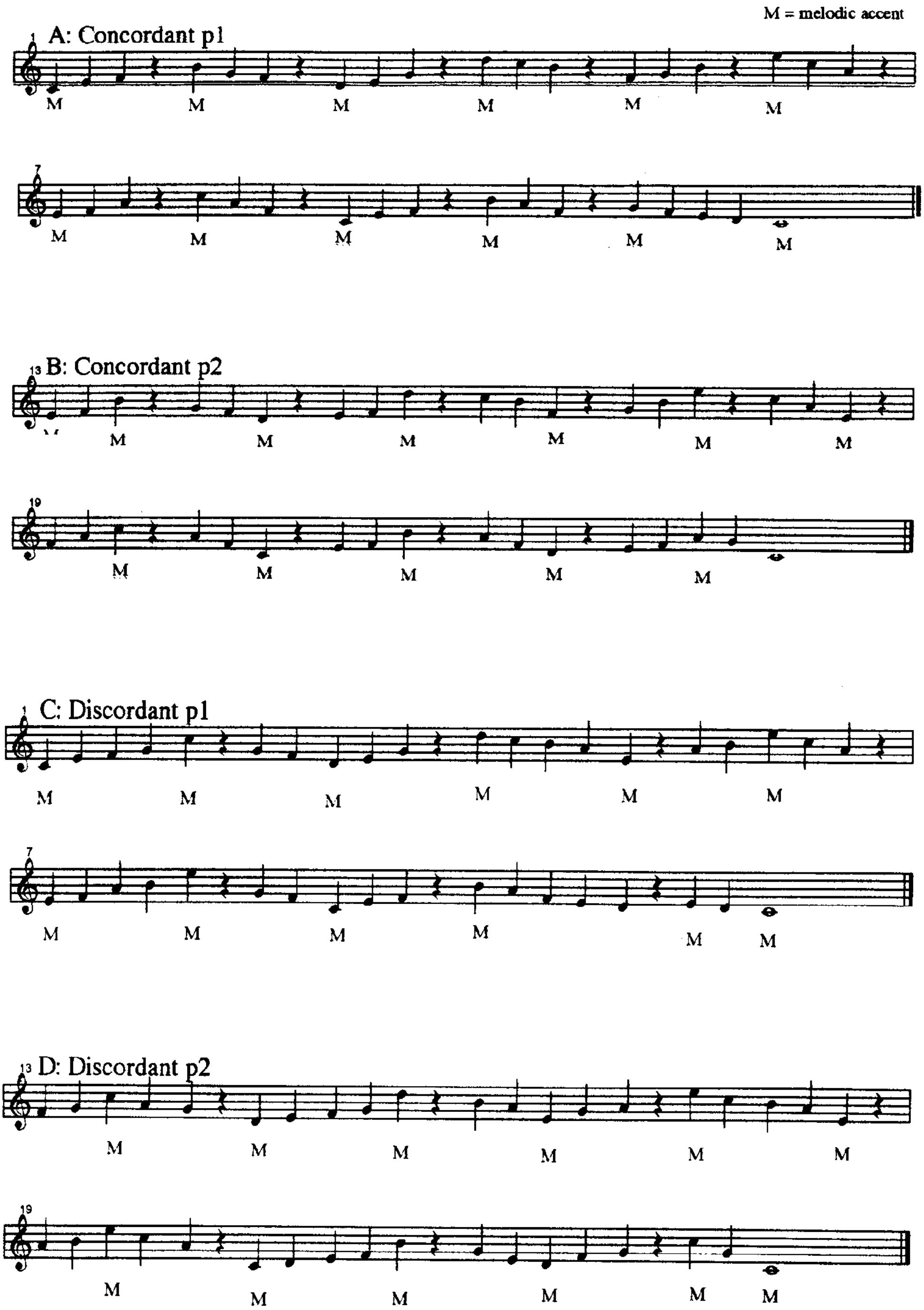


Fig. 3A–D. Notated example joint accent structure patterns for Experiment 1. Panels A and B show p1 and p2 concordant patterns, respectively. Panels C and D show p1 and p2 discordant patterns, respectively.

enced (5 years or more of formal training on a musical instrument, M = 13 years, range = 10 years) and eight were not (fewer than 5 years of musical training, M = 1 year, range = 4 years). Experienced listeners received nominal payment for participation; others received course credit in an introductory psychology course at the Ohio State University.

Apparatus and Stimulus Generation

All melodies were created using the 5.0 version of MIDILAB (Todd, Boltz & Jones, 1989), which ran on 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 was used to generate and organize sound patterns in experimental sessions and to collect data. Sound signals were transmitted to a separate testing room, amplified using a Rane HC-6 headphone console, and presented over AKG-K270 headphones at a comfortable listening level. Instructions were recorded and played on cassette tape. Tapping responses were produced on a box with a circular pad centered within a semicircle of four optical sensors. The size of the resting pad and optical sensors was similar to the tip of an adult's index finger. At the beginning of a session, subjects chose a single optical sensor to tap with the index finger of their dominant hand, which rested on the central pad between taps.

Design and Conditions

Nine different pattern types (see Table 1) resulted from crossing three levels of melodic (phase) structure (p1, p2, no **m** accents) with three levels of rhythmic accent period (pauses every 4 or 6 beats, no **r** accents). This produced four JAS conditions (16 patterns), four control conditions (10 patterns), and one baseline condition. The following discussion first describes the generation of melodic and rhythmic control patterns, then the JAS patterns that resulted from combining these controls. A table listing all events in all patterns used for this study is available from the author.

Melodic Controls

Transformations of an initial base melody generated the set of 8 melodic control patterns. Three expert listeners (one professor of music theory and two music graduate stu-

TABLE 1
Pattern Types in Each Experiment

	Melodic Accent Phase		
Rhythmic Accent Period	**p1**	**p2**	**No Melodic Accent**
Experiment 1			
4 beat	**Concordant p1**	**Concordant p2**	r4 Control
6 beat	**Discordant p1**	**Discordant p2**	r6 Control
No rhythmic accents	Melodic control p1	Melodic control p2	Baseline
Experiment 2			
6-beat	**Concordant p1**	**Concordant p2**	r6 Control
4-beat	**Discordant p2**	**Discordant p2**	r4 Control
No rhythmic accents	Melodic control p1	Melodic control p2	Baseline

NOTE—Joint accent structure patterns are shown in bold.

dents in composition at the Ohio State University) listened to these melodies and verified the salience of accent locations. The base melody was constructed by using pitches from the C-major diatonic set, and included 45 tones comprising 300 ms interonset intervals (IOIs) formed by 200-ms tone durations and 100-ms interstimulus intervals. Recurring **m** accents in the base melody were generated from correlated contour pivots and enlarged changes in pitch interval (relative to other pitch changes) that occurred every 4 beats, or 1200 ms (see Tone 4 in Figure 1). Every pair of successive **m** accents differed in pitch, so that short-term memory would not contribute to perceived accents. A total of 12 **m** accents occurred per pattern.

Two transformations of this base melody generated four of the eight melodic control patterns. One transformation reversed the pitch interval distribution; thus, a base melodic sequence of E_4-G_4-B_4-C_5, with pitch intervals (in semitones) of +3, +4, +1, becomes E_4-F_4-A_4-C_5 with intervals of +1, +4, +3 semitones. The second transformation inverted the direction of changes between successive pitches, such that pitch intervals of +3, +4, and +1 would become -3, -4, and -1. These transformational constraints were occasionally relaxed if a sequence sounded distractingly “unmusical.”

The next four melodic controls were created by phase shifting all **m** accents in each melody by 2 beats. This was done by eliminating the first two tones and adding two tones at the end. The first **m** accent in these phase-shifted melodies therefore occurred on Tone 3 rather than Tone 5. Patterns with phase-shifted melodic accents are denoted by “p2,” whereas the original four patterns are denoted “p1.”

Rhythmic Controls and Baseline

Two rhythmic controls comprised monotone (C_4) pitch sequences with pauses (a silent IOI) every 4 (r4) or 6 (r6) beats. A final baseline control sequence was an isochronous monotone string of C_4 tones with no pauses, without **m** or **r** accents. Rhythmic and baseline controls comprised 45 beats (300 ms cycles including tones and pauses), as did melodic controls.

JAS sequences

Sixteen JAS patterns resulted from crossing accentual characteristics of melodic and rhythmic controls. Concordant patterns were created by replacing every fourth tone in one of the melodic controls with a pause. Discordant JAS patterns were created by replacing every sixth tone in melodic controls with a pause. When a rest was positioned directly before an **m** accent, the pitch interval preceding the **m** accent in the melodic control was preserved. The pitch class of inflection tones in melodic controls were therefore often altered when converted into a JAS pattern.[2] Phase relationships among accents depended on the kind of melodic control that was combined with rhythmic controls (p1 or p2).

Procedure

Patterns were presented twice in a trial. The second presentation comprised an octave-raised version of the pattern. Trials began with a high-pitched warning tone, followed by a C-major triad that signaled the onset of each pattern. Participants were instructed to “hear out” between 5 and 15 accents while listening to the first presentation, where an accent was: “*any tone that stands out from the others around it.*” They were told to tap selectively during the second presentation to onsets of tones perceived as accented. At no point were listeners explicitly instructed to tap regularly. Participants were tested either individually or in sets of two; participants run in sets of two were not able to view each other.

---

[2] In Jones and Pfordresher (1997), the pitch class of inflection points that followed a pause was retained and the pitch interval preceding the inflection point often changed. Similar results were obtained in each study.

In order to limit the number of trials in a session, listeners were exposed to only half of patterns containing **m** accents. Half the subjects were presented with all instances of the base melody and its reversal transformation (including the different accent period ratio and phase variations). The other half were presented with the inversion transformations of these melodies. Thus a listener heard 15 patterns in a session: each of the 12 melodic and JAS patterns were repeated 5 times, each of the remaining three patterns (r4, r6, baseline) were repeated 10 times.

A session comprised seven blocks of trials. Twelve practice trials in Block 1 included melodic controls and JAS patterns. Blocks 2 through 5 each comprised 20 trials of JAS patterns; **m** accent phase was blocked in an alternating fashion across these blocks, and the order of blocks was counterbalanced across subjects. Concordant and discordant patterns were presented randomly within these blocks. Block 6 comprised 25 trials: 15 trials of rhythmic control and baseline patterns (intermingled) followed by 10 trials of melodic controls (p1 and p2 intermingled). Block 7 (25 trials) was similar to Block 6, except that the rhythmic and baseline controls appeared before melodic controls. Within these grouping constraints, all patterns were randomly intermingled according to two different orders. An experimental session (with five rest breaks) lasted approximately 1.5 hours (a total of 102 trials).

RESULTS

Each trial's data comprised a series of tap onset times produced by listeners. Taps occurring earlier than 1400 ms into the pattern (preceding the offset of the eighth tone) and ones later than 12,000 ms (the onset of the 41st tone) were eliminated to remove artifactual variability from beginning and ending taps. Preliminary analyses revealed that many listeners did not tap at all to baseline control patterns (which contained no accents). Because this behavior doesn't contradict instructions, these trials were dropped from all analyses in both experiments, although the participants were retained. Results are presented in two main sections, based on the two primary hypotheses. The first section considers locations of individual taps, the second considers time spans between taps.

Locations of Individual Taps

This section examines whether listeners tapped at predicted accent locations. In order to simplify this analysis, each tap was categorized relative to the nearest beat onset. Preliminary analyses indicated that almost all taps anticipated tone onsets by about 20 ms (cf. Aschersleben & Prinz, 1995). Each sequence was then divided into quarters, such that each tap was categorized as marking one of the 12 recurring beats as accented. For instance, a tap that anticipates the penultimate D_4 of the concordant p1 pattern (see Figure 3A) would be categorized as marking Beat 8 of the 12-beat cycle as accented. Figures 4A–4D display the total number of taps occurring at beat onsets using this categorization system, for each pattern type. Notation from the first quarter of the concordant p1 pattern is shown at the top as an example. Labels "Mp1" and "Mp2" indicate locations of melodic accents for responses to pattern from p1 and p2 **m** accent phase

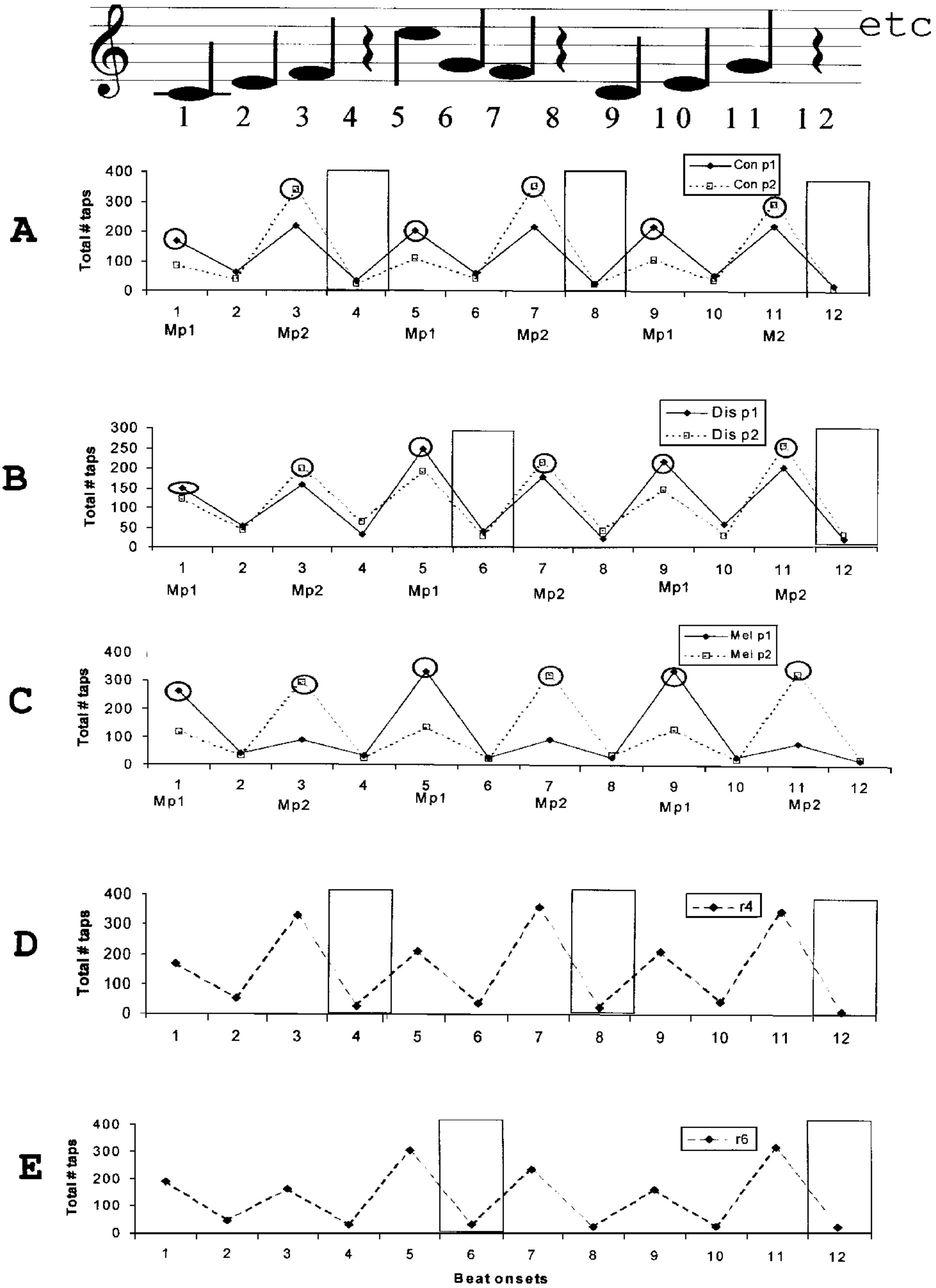


Fig. 4A–E. Frequency histograms showing locations of taps for patterns in Experiment 1. Tapping locations are rounded to the nearest beat within recurring 12-beat cycles (3600 ms) for each pattern. Locations of predicted melodic accents are indicated by Mp1 and Mp2 for melodic p1 and p2 conditions, respectively, and taps coincident with these responses are highlighted by circles. Rectangles outline beats at which pauses occurred. See the text for further details.

conditions, respectively. Circles highlight tapping responses to these melodic accents.

Taps generally coincided with predicted locations of accents, although listeners occasionally appeared to hear additional accents at subdivisions of larger accent periods. Responses to JAS patterns are shown in Figure 4A (concordant p1, p2) and 4B (discordant p1, p2). Both concordant patterns (Figure 4A) elicited taps at tones both preceding and following a pause, although listeners generally preferred the former. This result suggests that pauses create primary accents on tones preceding pauses and secondary accents on tones following pauses (cf. Povel & Essens, 1985). Listeners also showed a tendency to tap to locations where **m** accents occurred. This modulated the degree to which listeners favored tones that preceded pauses and resulted in a more equal distribution of taps before and after pauses in p1 patterns, for which **m** accents fell on tones following pauses. Responses to discordant patterns (Figure 4B) resembled those to concordant patterns in general, although responses to **m** accents were reduced, and listeners occasionally tapped to locations at which no hypothesized accents occurred (e.g., Beat 3 in the p1 discordant pattern). Because different phase conditions were blocked across a session, it is unlikely that this effect reflects a carryover from p1 to p2 patterns (cf. Jones & Yee, 1997; McAuley & Kidd, 1998). The salience of accents in JAS patterns was further confirmed by responses to control patterns, shown in Figures 4C-4E. Listeners responded to positions of **m** accents (Figure 4C), responded to tones both preceding and following pauses while favoring the former (Figures 4D and 4E), and heard accents at subdivisions of the longer **r** accent periods in r6 patterns (Figure 4E).

Table 2 shows the percentage of total taps for each accent type and pattern type, to confirm the observations of data shown in Figures 4A-4E.

TABLE 2
Percentage of Taps Falling at Certain Accent Types in Experiment 1

Pattern	Before Pause	After Pause	Contour	Joint	Exclusive	Any	(Chance)
Con-p1	44	39	39	39	44	83	42
Con-p2	67	21	67	67	21	88	42
r4	57	33	N/A	N/A	N/A	89	33
Dis-p1	32	24	44	28	43	72	50
Dis-p2	32	24	48	34	37	71	50
r6	39	27	N/A	N/A	N/A	66	50
Mel-p1	N/A	N/A	67	N/A	N/A	67	25
Mel-p2	N/A	N/A	63	N/A	N/A	63	25
Across patterns	46	28	55	43	36	75	
(Chance)	16	16	19	10	13	40	

Joint accents, which are relevant only for JAS patterns, occur when an **m** accent falls either before or after a pause, whereas exclusive accents (also relevant for JAS patterns) occur for **m** accents that are not adjacent to pauses, or for tones adjacent to pauses that are not marked by **m** accents. In general, listeners tapped to accents more often than would be predicted by chance, preferred tones preceding pauses, and preferred tones marked by **m** accents. A particularly important observation is that listeners favored joint accents over exclusive accents in JAS patterns, despite the fact that exclusive accents were more prevalent (see chance estimates, which reflect the number of accents for a given type divided by the total number of events in a pattern). Not surprisingly, the tendency to favor joint accents specifically reflects responses to concordant patterns, for which accents appear to have been more salient. On average (across phase conditions), 86% of taps coincided with some kind of accent for concordant patterns, whereas 72% of taps coincided with some kind of accent for discordant patterns, although accents were more prevalent in discordant patterns (see chance estimates). Further analyses confirmed that the types of preferred tapping locations shown here reflect patterns of responses for individual participants. Most listeners positioned the majority of taps on tones preceding pauses for concordant and rhythmic control patterns, and they did so more often when these tones were marked by melodic accents in concordant patterns. Melodic accents were similarly favored by most subjects in melodic control patterns. Discordant patterns elicited more variable responses, with similar numbers of subjects favoring accented or nonaccent locations.

Time Spans Between Taps

The previous analysis examined locations of individual taps but did not examine time spans defined by successive taps (intertap intervals). For instance, although taps to p1 concordant patterns congregated at the first and third beats of each 4-beat cycle, the previous analysis does not establish whether subjects tapped more often at *both* positions (producing 2-beat periods), or if they more commonly chose one of the two positions (producing 4-beat periods). Three analyses, summarized in this section, examined the number of beats elapsing between successive taps for different patterns: distributions of produced intertap intervals, the variability of intertap intervals, and the number of trials on which listeners tapped a regular "beat." These analyses addressed the degree to which accent structures suggest a higher order period that listeners track while responding to accent locations.

Distributions of intertap intervals, shown in Figures 5A–5C, were generated by categorizing each intertap interval into one of 10 bins. As in analyses of tapping locations, the time of individual taps was rounded to the

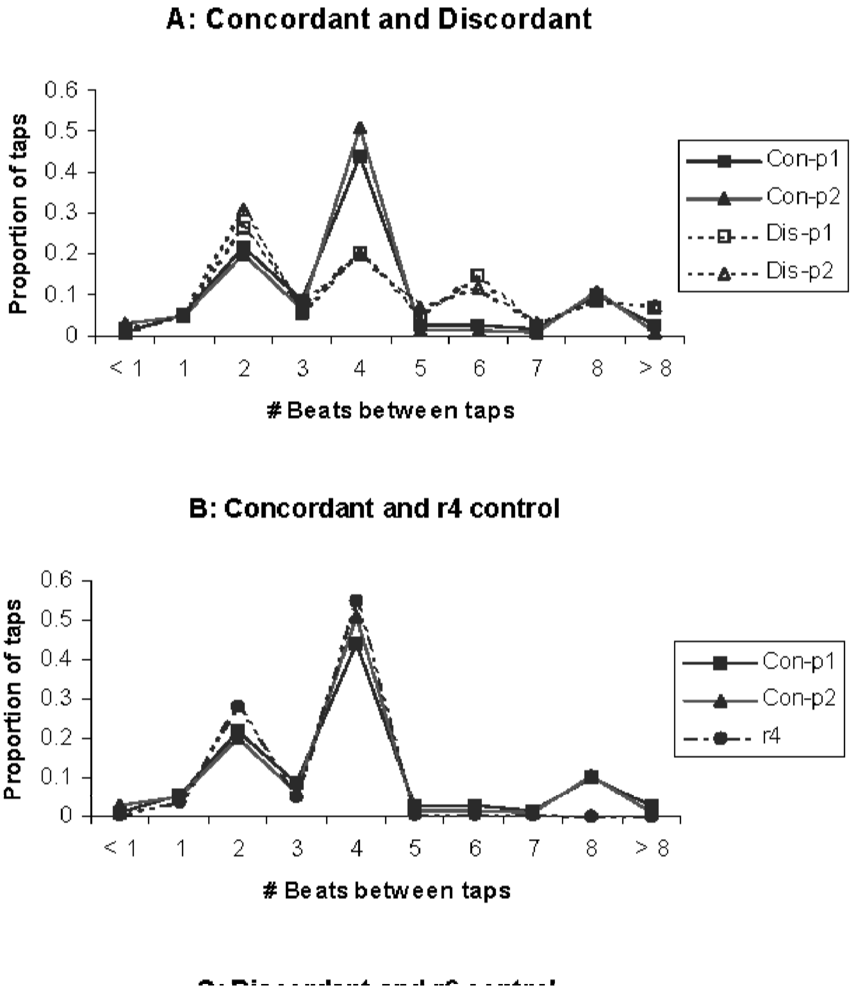


Fig. 5A–C. Distribution of intertap intervals in Experiment 1. Panel A shows all JAS patterns (concordant and discordant, crossed with p1 and p2), panel B shows both concordant patterns and the r4 rhythmic control, and panel C shows both discordant patterns and the r6 rhythmic control.

nearest beat onset. Eight bins were used for beats separated by 1 to 8 beats (300 to 3600 ms), and the remaining bins were used for intervals either less than 1 beat (< 150 ms) or greater than 8 beats (^{3}2550 ms). Proportions of intertap intervals were then generated by summing the number of taps in each bin over all trials within each pattern type (e.g., all concordant p1 patterns = 10 trials for each pattern type) and individual. Figures 5A–5C show the mean proportions across subjects for JAS and rhythmic control patterns. Melodic control patterns are not shown in order to simplify the presentation; their profiles were nearly identical to those of concordant patterns (r = .99).

These data were analyzed parametrically in three analyses of variance (ANOVAs) that separately compared JAS conditions (concordant vs discordant, for both phase conditions, see Figure 5A), concordant (p1 and p2) patterns versus the r4 control pattern (Figure 5B), and discordant (p1 and p2) patterns versus the r6 control (Figure 5C). Each omnibus ANOVA was followed by planned comparisons associated with certain bins (i.e., 2, 4, 6 and 8 beats, the evenly divisible beat periods). Planned comparisons averaged over melodic accent phase (p1, p2), which was not predicted to contribute. The critical α level was adjusted to .01 for all analyses.

Concordant and discordant JAS conditions are shown in Figure 5A. Concordant patterns encouraged more 4-beat intertap intervals whereas discordant ones encouraged more 6-beat intervals. This led to a significant interaction between accent period ratio and bin, $F(9, 108) = 8.31$, MSE = .020, $p < .00001$, in the omnibus ANOVA, which used a 10 (bin) × 2 (accent period ratio: concordant, discordant) × 2 (phase: p1, p2) × 2 (block order) × 2 (musical training) design. No other significant effects emerged. Planned comparisons confirmed that more 4-beat intervals were produced for concordant than discordant conditions, $F(1, 15) = 16.12$, $MSE = 0.69$, $p < .01$. By contrast, listeners produced more 6-beat intervals for discordant than concordant patterns, $F(1, 15) = 12.79$, $MSE = .016$, $p < .01$. Note that the peaks at 6 and 4 beats rule out the possibility that most listeners tracked every accent attracting taps in Figure 4.

Figures 5B and 5C compare concordant (p1, p2) with r4 patterns and discordant (p1, p2) with r6 patterns, respectively. Each related ANOVA incorporated a 10 (bin) × 3 (pattern type: p1, p2, rhythmic control) × 2 (block order) × 2 (musical training) design. Figure 5B shows highly similar responses across both phase conditions of concordant patterns and r4 patterns. The omnibus ANOVA did not reveal the critical interaction between pattern type and bin. Furthermore, the only planned comparison that revealed a difference involved the production of slightly more 6-beat intervals in concordant patterns (phase conditions combined) than in the r4 control, $F(1, 15) = 10.77$, $MSE = .003$, $p < .01$, although the actual number of 6-beat intervals involved was less than chance in both cases. Figure 5C, on the other hand, reveals different patterns of responding for discordant patterns and the r6 control pattern, confirmed by a significant pattern type by bin interaction, *F*(18, 216) = 4.84, *MSE* = .008, $p < .00001$. Planned comparisons confirmed that more 4-beat intervals were produced for the discordant than for the r6 pattern, *F*(1, 15) = 10.82, *MSE* = .102, $p < .01$, and that marginally more 6-beat intervals were produced for the r6 than discordant patterns, *F*(1, 15) = 6.49, *MSE* = .059, $p = .02$.

The second analysis of intertap intervals focused on variability and tested the prediction that variability would increase with pattern complexity, as indexed by accent period ratio. Coefficients of variation (CVs) were used to index variability (CV = mean IOI/SD per trial) because variability generally increases for longer intertap intervals (e.g., Ivry & Hazeltine, 1995, Wing & Kristofferson, 1973), and preliminary analyses revealed longer intertap intervals for discordant than concordant patterns. Figure 6 displays mean CV scores for the eight pattern types, averaged over trials and participants. CVs were greater for discordant patterns than for concordant patterns and all control patterns. This observation was tested in a 3 (rhythmic accent period: 4, 6, melodic control) × 2 (phase: p1, p2) × 2 (musical training) × 2 (counterbalance order) mixed factorial ANOVA, followed by planned comparisons between rhythmic control patterns and their relevant JAS pattern (r4 with concordant patterns, r6 with discordant patterns). The only significant effect in the ANOVA was a main effect of rhythmic accent period, *F*(2, 24) = 18.17, *MSE* = .012, $p < .0001$, which confirmed higher variability in discordant patterns. Planned comparisons showed that discordant patterns' CVs exceeded the r6 rhythmic control, *F*(1, 15) = 22.8, *MSE* = .011, $p < .001$, whereas concordant and r4 control patterns did not differ significantly ($p = .06$).

A final issue that is relevant to time spans between taps concerns the degree to which individual listeners tapped to a regular higher order time

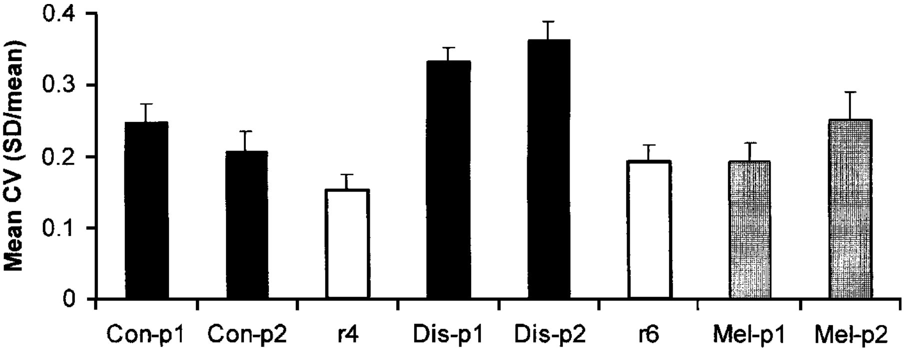


Fig. 6. Mean coefficients of variation for Experiment 1. Error bars show between-subject standard errors.

span, or "beat," when listening to musical patterns. Listeners were not instructed to do this; nevertheless, regularly recurring accents in patterns may have encouraged such tendencies. Table 3 shows the number of trials for each subject and pattern type that were characterized by tapping a regular higher order beat (total possible trials per cell = 10). Trials were categorized as "beat-tapping" trials if at least three intertap intervals matched the same beat period, and not more than two intervals deviated from this period (visual inspection of trials suggested that listeners often used two "warm-up" intertap intervals before attaining a regular pulse). These data reveal variability across subjects and pattern types; individuals clearly did not always respond by tapping a regular interval. Nevertheless, beat-tapping responses were more common for patterns with simpler accent structures. An ANOVA, using the same design as for CVs, followed by the same planned comparisons, verified these observations. Discordant patterns yielded fewer beat tapping trials than concordant patterns and melodic controls, leading to a main effect of rhythmic accent period, $F(2, 24) = 11.68$, $MSE = 9.187$, $p < .01$. Planned comparisons revealed fewer beat tapping trials for concordant than r4 patterns, $F(1, 15) = 10.28$, $MSE = 6.425$, $p < .01$, and fewer beat tapping trials for discordant than r6 patterns, $F(1, 15) = 18.15$, $MSE = 4.929$, $p < .05$. Although musicians produced slightly more beat tapping trials than nonmusicians, this difference was not significant.

TABLE 3
Number of Trials on Which Subjects Tapped a Regular Interval in Experiment 1 (Max = 10)

Training	Con-p1	Con-p2	r4	Dis-p1	Dis-p2	r6	Mel-p1	Mel-p2	Means
Nonmusicians									
s1	2	6	9	0	0	3	1	0	2.63
s2	0	0	8	0	0	0	0	1	1.13
s3	10	7	10	5	2	7	7	6	6.75
s4	0	0	8	1	0	6	0	1	2.00
s5	7	7	8	1	1	6	3	5	4.75
s6	8	10	4	1	0	4	7	9	5.38
s7	8	9	8	7	6	2	7	8	6.88
s8	3	5	7	6	3	7	7	4	5.25
Musicians									
s9	6	7	6	6	7	9	8	10	7.38
s10	8	9	10	4	7	9	6	0	6.63
s11	7	5	9	0	0	6	7	8	5.25
s12	4	3	9	2	0	2	9	3	4.00
s13	2	6	8	0	0	1	5	1	2.88
s14	2	9	10	0	0	9	8	10	6.00
s15	2	3	10	0	0	5	0	0	2.50
s16	9	8	8	1	1	8	6	2	5.38
Means	4.88	5.88	8.25	2.13	1.69	5.25	5.06	4.25	

DISCUSSION

Two main conclusions were supported by Experiment 1. First, listeners heard accents at predicted locations. Listeners usually tapped to positions marked by melodic contour pivots and to tones adjacent to pauses. For patterns with pauses, listeners primarily responded to the tone preceding the pause, although a secondary **r** accent appeared on tones following pauses. Furthermore, listeners appeared to hear accents on tones more often when tones were marked by coupled **m** and **r** accents (Boltz, 1993; Dawe et al., 1994; Jones, 1987; Jones & Boltz, 1989; Monahan et al., 1987; Tekman, 1997), but only when these coupled (joint) accents appeared in JAS patterns with a simple accent period ratio (concordant patterns).

Second, time spans between tapping responses were influenced by period ratios between recurring **m** and **r** accents and were unaffected by phase relationships between accents. Discordant patterns elicited more variable tracking of accents in patterns overall (CVs), and elicited a distribution of intertap intervals that differed from their constituent control patterns (r6 and melodic controls, the latter of which resembled concordant patterns). Concordant patterns elicited less variable tapping and a distribution of intervals that resembled their constituent controls. Importantly, listeners did not track *all* accents but rather selected certain accents and ignored others in a way that reflected the global accent structure of patterns. Although listeners did not tap a regular period most of the time, they nevertheless did this more often for simpler patterns (control patterns and concordant JAS patterns). One implication of these results is that local changes determine a field of temporal landmarks, but the global organization of these landmarks (specified by accent period ratios) influences which ones are chosen or passed over as a listener tracks an ongoing melody.

Experiment 2

Experiment 2 incorporated a new set of melodies, but was otherwise identical to Experiment 1. The accent structures of new stimuli, examples of which are shown in Figure 7, addressed two alternative interpretations for the results of Experiment 1. First, it is possible that variability of tapping responses to discordant patterns exceeded concordant patterns in Experiment 1 because time spans between locations of successive accents of any type (e.g., the time between an **m** accent and the next **r** accent) were more variable in discordant patterns (see Figure 2). In Experiment 2, the variability of time spans between successive accents in concordant and discordant patterns was better equated than in Experiment 1 by making **m** accent periods (which all JAS patterns share) more variable. Second, it is possible that concordant patterns in Experiment 1 were simpler to track

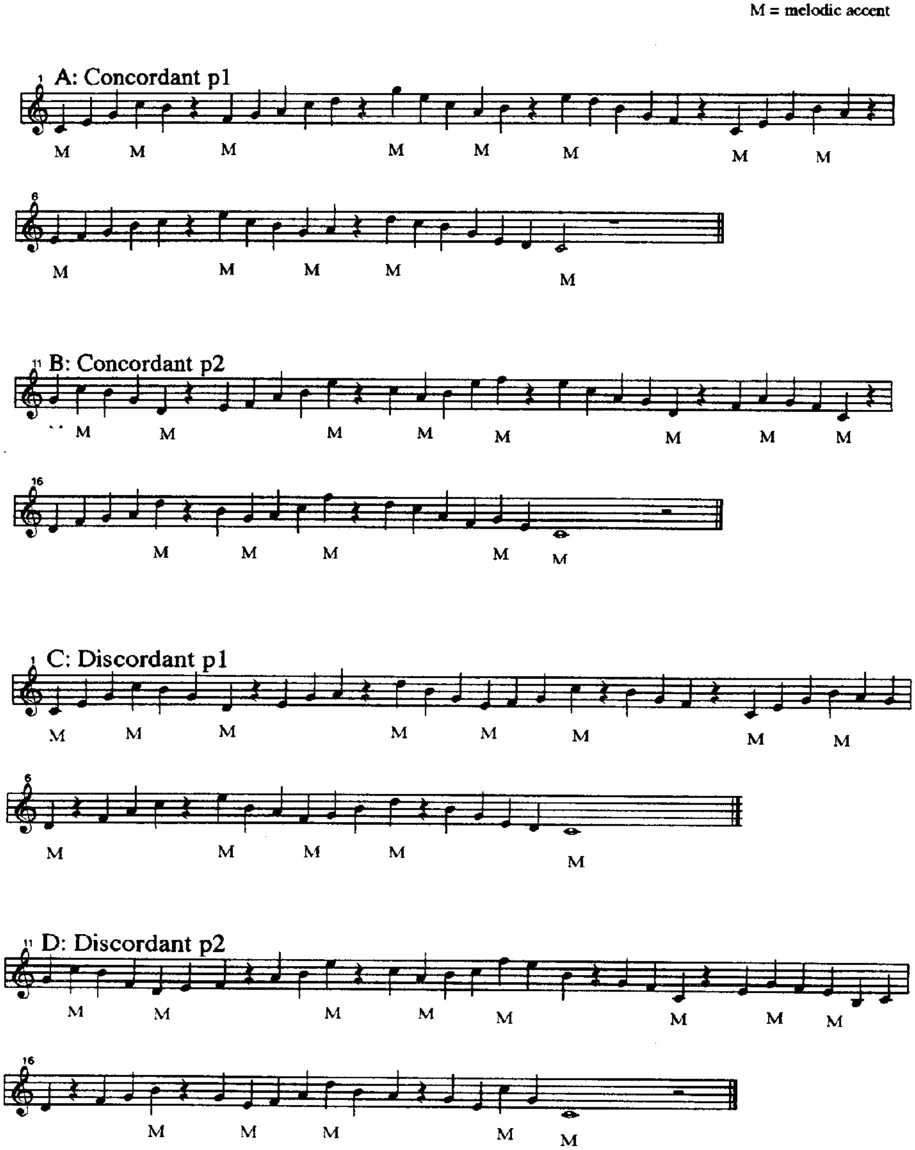


Fig. 7A–D. Example patterns for Experiment 2 in musical notation.

because they suggest a binary organization (Fraisse, 1982; Jones, 1990; Palmer & Krumhansl, 1990; Smith & Cuddy, 1989). In Experiment 2, concordant JAS patterns were designed to suggest a dominant ternary time structure (6 beats between pauses and alternating 3/6 **m** accent cycles), whereas discordant patterns resulted from combinations of ternary **m** accent periods and binary **r** accent periods.

METHODS

Experiment 2 was identical to Experiment 1 with the following exceptions.

Subjects

Thirty-five new subjects volunteered for Experiment 2 in return for credit in an introductory psychology course at the Ohio State University. Of these, data from 19 were eliminated for failure to comply with instructions. Half the remaining 16 subjects were musically experienced (*M* = 8.3 years of training on a primary instrument, range = 6 years) and the other half were less experienced (*M* = 1.9 years of formal musical training, range = 4 years).

Stimulus Conditions

A new set of four melodies were generated from a base melody as in Experiment 1; in Experiment 2 **m** accents formed a recurring cycle comprising one 6-beat accent period followed by two 3-beat periods (see Figure 7). In all, base melodies had 12 melodic contour accents. In order to present an appropriate number of accent cycles, melodies were slightly longer in Experiment 2 than Experiment 1 (49 events). The creation of melodic patterns and **m** accents was otherwise identical to Experiment 1.

Rhythmic accent periods were again defined by rests; rhythmic controls (r4, r6) were identical to those of Experiment 1. In concordant patterns, rests determined **r** accents every 6 beats, starting on the sixth beat. Rests in discordant patterns alternated between periods of 4 and 8 beats in order to equate the number of rests between the two conditions, and to equate the average time span between accents of any type in concordant and discordant patterns.

RESULTS AND DISCUSSION

The data from Experiment 2 were analyzed by using the same procedures as in Experiment 1. Data from trials with baseline sequences were again excluded because many listeners did not tap to these.

Locations of Individual Taps

The locations of individual taps were defined within recurring 12-beat spans, as in Experiment 1, and are shown in Figure 8. As in Experiment 1, these analyses revealed a preference for positions marked by accents. However, a less consistent relationship between tapping locations and accents than was found in Experiment 1, perhaps resulting from the greater variability in accent periods for patterns in Experiment 2.

Responses to JAS patterns are in Figure 8A and 8B. Concordant patterns (Figure 8A) revealed a strong tendency for listeners to favor tones either preceding or following pauses, with a preference for the former. The tendency to prefer tones that precede pauses was again enhanced for concordant p2 patterns, in which such tones were marked by **m** accents, relative to concordant p1 patterns. Listeners largely ignored **m** accents that did not converge with primary or secondary **r** accents; because of this, circles in the

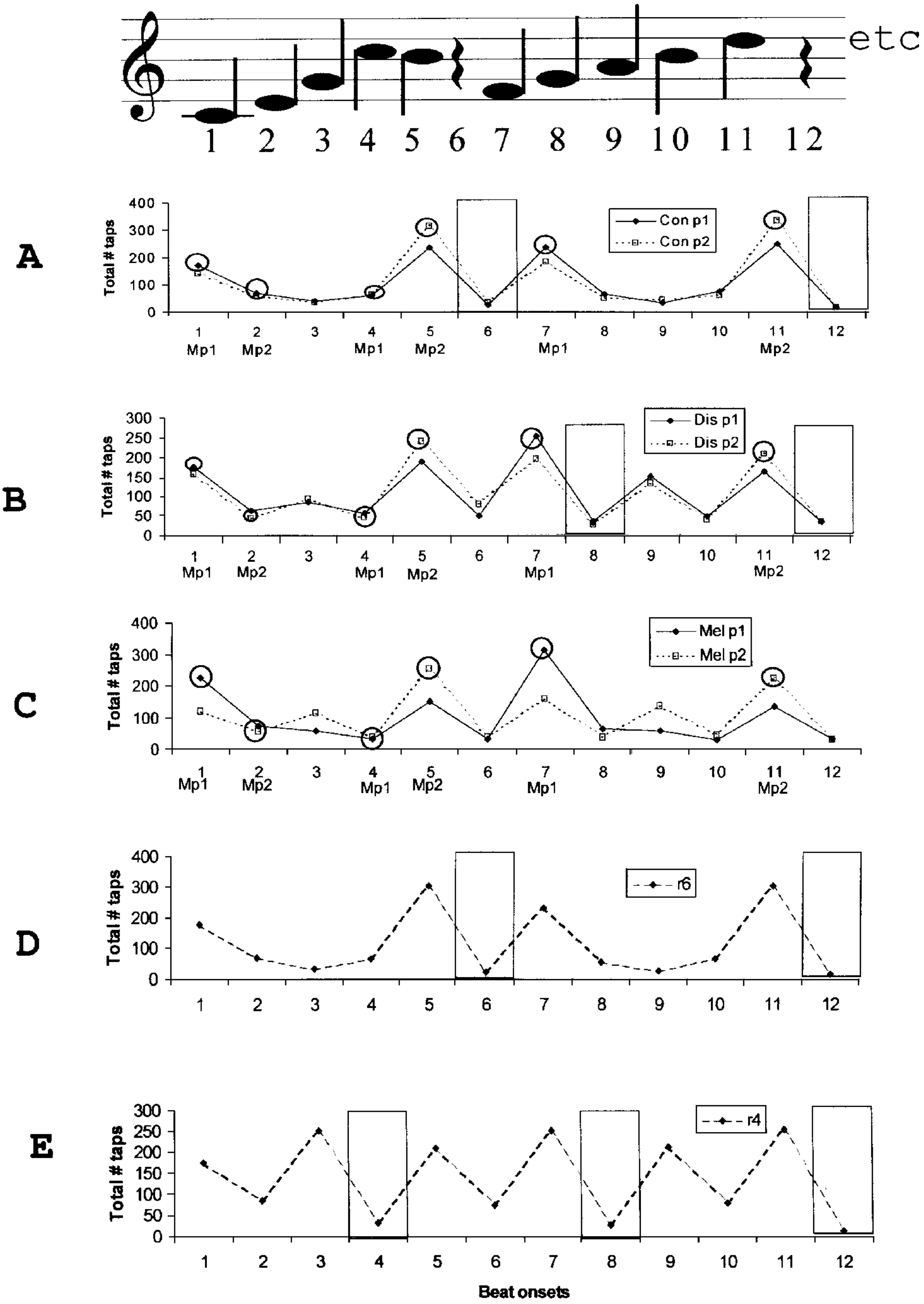


Fig. 8A–E. Frequency histograms showing locations of taps for patterns in Experiment 2. Tapping locations are rounded to the nearest beat within recurring 12-beat cycles(3600 ms) for each pattern. Locations of predicted melodic accents are indicated by Mp1 and Mp2 for melodic p1 and p2 conditions, respectively, and taps coincident with these accents are highlighted by circles. Rectangles outline beats at which pauses occurred.

figure often surround both phase conditions. Responses to discordant patterns (Figure 8B) likewise showed a tendency to favor tones occurring before or after pauses. One exception to this observation is Beat 5, which was not adjacent to a pause but nevertheless attracted many taps, especially when it was marked by an **m** accent (discordant p2).

As in Experiment 1, responses to control patterns (see Figure 8C-8E) resemble JAS patterns. Listeners tapped at accented locations in melodic controls, with the exception of accents that fell on Beats 2 and 5. Responses to rhythmic control patterns resembled those of Experiment 1, except that listeners did not tap to subdivisions of the higher order accent period for the r6 pattern (beats 3 and 9) as was found in Experiment 1. This difference probably results from the dominance of 6-beat accent periods across the experimental session in Experiment 2.

Table 4 shows the percentage of taps falling at different accent types for different patterns in Experiment 2. Again, the majority of taps fell on some kind of accent, though less often than was found in Experiment 1 (Experiment 2 = 68% across patterns, Experiment 1 = 75%). One apparent difference from Experiment 1 was that taps fell on joint accents (coincidences of **m** and **r** accents) less often than on exclusive accents in JAS patterns overall. This unexpected result mostly reflects responses to discordant p2 patterns.

Time Spans Between Taps

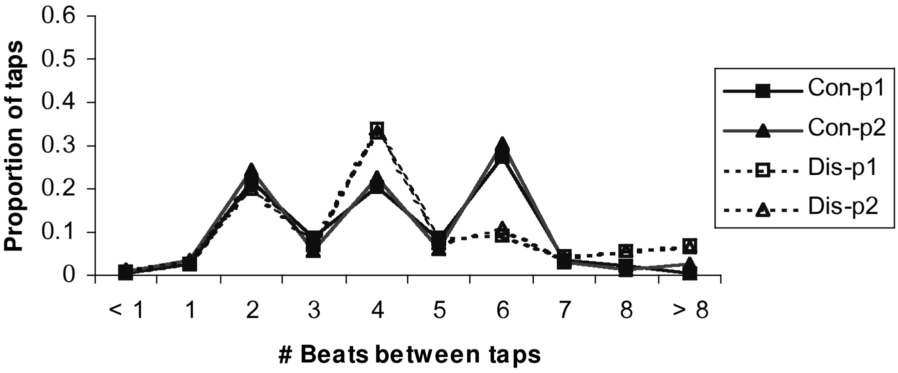
The higher order periods tracked by listeners were assessed as in Experiment 1. Figures 9A–9C present the distributions of intertap intervals, with responses binned and presented in the same way as in Experiment 1. Distributions for melodic controls again greatly resembled those for concordant patterns ($r = .97$) and are not shown. These data were again analyzed in

TABLE 4

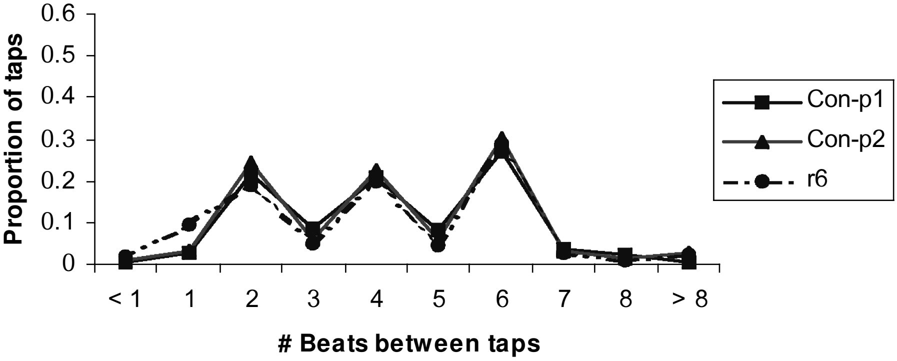
Percentage of Taps Falling at Certain Accent Types in Experiment 2

Pattern	Before Pause	After Pause	Contour	Joint	Exclusive	Any	(Chance)
Con-p1	38	33	37	33	43	76	42
Con-p2	49	25	53	49	29	78	42
r6	45	30	N/A	N/A	N/A	75	33
Dis-p1	32	25	38	33	28	62	42
Dis-p2	31	23	38	16	60	76	50
r4	46	36	N/A	N/A	N/A	81	50
Mel-p1	N/A	N/A	48	N/A	N/A	48	25
Mel-p2	N/A	N/A	42	N/A	N/A	42	25
Across patterns	40	29	43	33	53	68	
(Chance)	14	14	19	7	19	39	

A: Concordant and Discordant



B: Concordant and r6 control



C: Discordant and r4 control

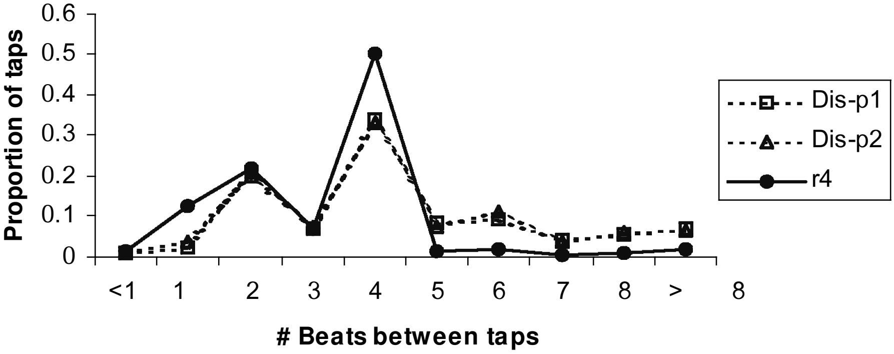


Fig. 9A–C. Distribution of intertap intervals for Experiment 2.

three analyses, using the same ANOVA designs and planned comparisons, with an adjusted α level of .01.

Figure 9A presents data from concordant and discordant patterns and both phase conditions. The two accent period ratio conditions (concordant, discordant) produced reliably different distributions of intertap intervals, leading to an interaction of accent period ratio with bin, $F(9, 108) = 13.82$, $MSE = .007$, $p < .00001$. Intertap intervals mostly spanned 6 beats for concordant patterns and 4 beats for discordant patterns. Planned comparisons verified the reliability of differences in 4-beat interval production, $F(1, 15) = 9.97$, $MSE = .023$, $p < .01$, as well as with 6-beat intervals, $F(1, 15) = 30.40$, $MSE = .019$, $p < .0001$. Listeners failed to produce many three-beat intervals in general, which concurs with examinations of tapping locations. As in Experiment 1, concordant patterns did not differ from their constituent (r6) rhythmic control, as shown in Figure 9B. ANOVAs and planned comparisons revealed no significant differences among these profiles. However, discordant patterns did differ from their r4 rhythmic control (Figure 9C). Pattern type (discordant p1, discordant p2, r4) interacted with bin, $F(18, 216) = 4.14$, $MSE = .008$, $p < .01$. Planned comparisons confirmed that listeners produced more 6-beat intervals for both discordant patterns than for r4 controls, $F(1, 15) = 26.96$, $MSE = .069$, $p < .01$. The apparent difference between r4 and discordant patterns at the 4-beat bin was not significant.

The second analysis of intertap intervals again examined variability using CV; these data are shown in Figure 10. Discordant CVs again exceeded those of concordant and control patterns. The same statistical design was used as in Experiment 1; the ANOVA revealed a main effect of rhythmic accent period, $F(2, 24) = 3.94$, $MSE = .011$, $p < .05$. A further ANOVA, excluding melodic controls, confirmed that concordant and discordant conditions differed significantly, $F(1, 12) = 4.85$, $MSE = .013$, $p < .05$. Planned

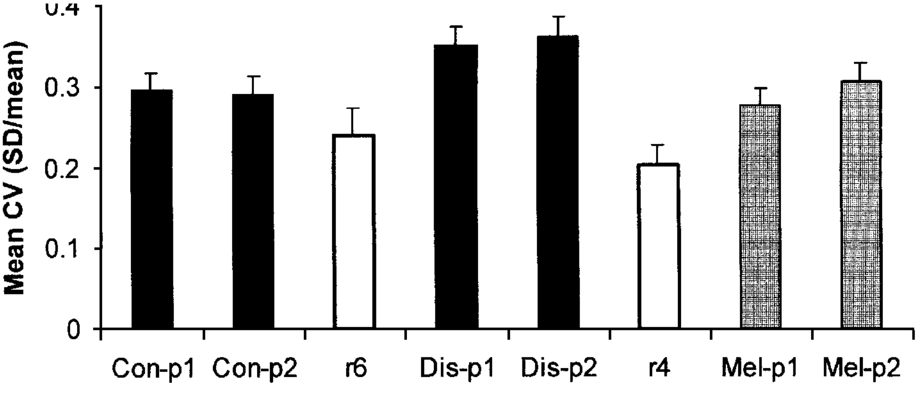


Fig. 10. Mean coefficients of variation (CV) for Experiment 2.

comparisons involving the JAS patterns and rhythmic controls revealed no significant difference between concordant patterns and the r6 control, whereas discordant patterns yielded higher CVs than the r4 control, *F*(1, 15) = 23.50, *MSE* = .011, $p < .001$.

The third analysis again examined the degree to which individual listeners tapped a regular higher order beat, by counting trials on which subjects tapped a regular intertap interval. Table 5 shows these data. These analyses reveal some interesting departures from the data of Experiment 1. First, very few trials elicited regular beat tapping, which most likely reflects the greater variability in accent period for patterns in Experiment 2. Another interesting departure from Experiment 1 is that r6 patterns elicited beat tapping for many trials (63% of trials, across subjects), whereas less beat tapping was found for the r4 pattern (39%). This may result from the fact that the r4 rhythmic pattern was associated with more variable discordant patterns in Experiment 2. As in Experiment 1, an ANOVA and planned comparisons like those used for CVs were run on these data. In Experiment 2, only planned comparisons revealed effects of pattern structure, with more beat tapping for r6 than for concordant patterns, *F*(1, 15) = 88.82, *MSE* = 5.024, $p < .001$, and for r4 than for discordant patterns, *F*(1, 15) = 56.54, *MSE* = 5.301, $p < .001$.

TABLE 5

Number of Trials on Which Subjects Tapped a Regular Interval in Experiment 2 (Max = 10)

Training	Con-p1	Con-p2	r6	Dis-p1	Dis-p2	r4	Mel-p1	Mel-p2	Means
Nonmusicians									
s1	2	1	8	4	5	0	4	3	3.38
s2	0	0	7	2	4	6	4	3	3.25
s3	1	0	6	1	6	4	2	1	2.63
s4	1	0	4	2	3	2	0	0	1.50
s5	0	3	9	2	3	4	2	1	3.00
s6	0	0	6	0	1	9	0	3	2.38
s7	7	7	8	7	2	6	5	5	5.88
s8	1	0	7	2	4	8	0	5	3.38
Musicians									
s9	0	1	9	1	0	9	2	1	2.88
s10	0	0	10	0	0	0	3	6	2.38
s11	2	0	4	4	2	0	6	1	2.38
s12	0	1	6	0	3	5	0	0	1.88
s13	1	0	0	3	5	2	0	1	1.50
s14	0	0	6	0	0	2	4	3	1.88
s15	2	6	4	2	3	3	5	7	4.00
s16	1	2	6	0	1	2	0	0	1.50
Means	1.13	1.31	6.25	1.88	2.63	3.88	2.31	2.5	

EXPERIMENT 1 VERSUS EXPERIMENT 2

Finally, the influence of JAS on the variability of intertap intervals was assessed by comparing CV scores across JAS melodies for both experiments. This design crossed experiment (1 versus 2, which distinguishes **m** accent period), accent period ratio (concordant, discordant within each experiment), accent phasing (p1, p2), and both musical training conditions. Mean CV scores for experimental patterns appear in Figure 11, averaged over musical training. Only accent period ratio significantly affected CV, with concordant patterns producing less variable performance than discordant ones across experiments, $F(1, 28) = 20.99$, $MSE = .013$, $p < .0001$. No other main effects or interactions emerged, although responses to concordant patterns were slightly more variable in Experiment 2 than in Experiment 1, for concordant patterns.

General Discussion

Two experiments explored the perception of melodic (**m**) accents generated by changes in melodic contour and rhythmic (**r**) accents generated by pauses. Musiclike patterns were constructed with regularly recurring **m** and/or **r** accent periods. In JAS patterns, which featured both accent types, relationships between these higher order accent periods were varied in terms of period and phase. The salience of these accents and the effect of their higher order organization was assessed through analyses of tapping responses, when listeners were instructed to tap at accented tone onsets.

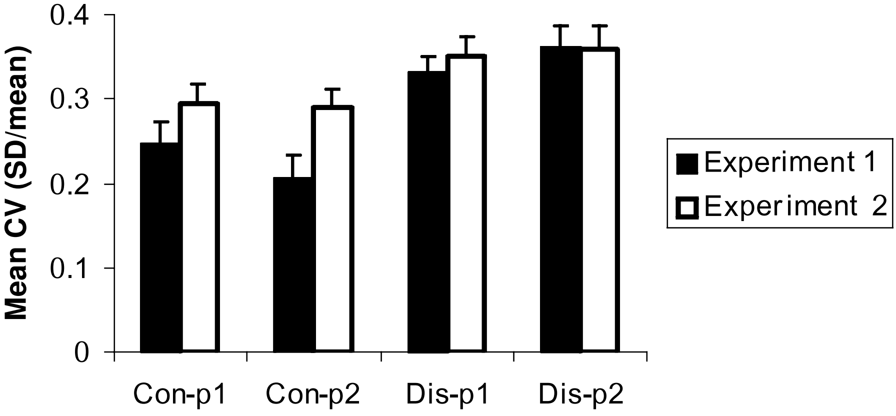


Fig. 11. Mean coefficients of variation (CV) in Experiments 1 and 2 for joint accent structure patterns.

Two main conclusions emerged, with respect to the primary hypotheses of the current research. First, previous claims about the structural determinants of accents were generally confirmed. Analyses of taps to **m** accent locations confirmed that listeners tapped to pivot points in the melodic contour (Thomassen, 1982). Analyses of taps to **r** accent locations revealed a new result: pauses appear to create a primary accent on tones preceding pauses and a secondary accent on tones following pauses (cf. Povel & Essens, 1985). Accent structure also appears to influence accent salience. Listeners were more likely to respond to accented locations for concordant patterns and the simpler patterns of Experiment 1, especially when combined **m** and **r** accents appeared in concordant patterns. Second, time spans between taps were influenced solely by accent period ratios. In analyses of both the distribution and variability of intertap intervals, responses to concordant patterns resembled their constituent control conditions, whereas discordant patterns differed from their constituent controls. Analyses of tapping variability confirmed that listeners tapped less regularly to discordant patterns, even in Experiment 2 for which listeners appeared to produce a more diffuse distribution of tapping periods for concordant than discordant patterns. In no case did musical experience qualify these results, which suggests that the manipulations of accents used here influence very basic responses to music (cf. McAuley & Semple, 1999; Trainor, Desjardins, & Rockel, 1999).

The following discussion will focus on the applicability of these findings to two major topics in music cognition: the relationship between melody and rhythm, and structural bases of higher order time structure in music.

MELODY AND RHYTHM

Much debate has focused on the relationship between melody and rhythm in music. The current research similarly focused on this issue, by examining the salience of **m** and **r** accents, as well as the way in which these accents may interact within the framework of JAS. Overall, listeners responded to both types of accents, but appeared overall to favor **r** accents. Melodic contour pivots, combined with locally greater melodic interval changes, attracted taps in both melodic control and JAS patterns; however, responses to **m** accents were reduced in favor of **r** accents for Experiment 2's more variable patterns. It is significant that **m** accents had any effect on listener responses, given that some research has discredited the salience of melody with respect to temporal structure in music (Drake & Palmer, 1993; Huron & Royal, 1996; Snyder & Krumhansl, 2001; Vos et al., 1994). In general, responses to combined **m** and **r** accents in JAS patterns resembled responses to accents in control patterns. Discordant patterns appeared to reduce the salience of accents overall, implying that higher order JAS does influence the salience of local accents.

The current research speaks to the issue of whether melody and rhythm contribute independently or interactively to musical pattern structure (see Krumhansl, 2000, for a review). The current data, as well as the theoretical framework of JAS, suggest a combination of independence and interaction. Theoretically, the perception of a musical pattern's JAS is initially guided by a sensitivity to periodic recurrences of accents within musical dimensions of pitch and time, which are then combined to form period and phase relationships between these dimensions. The current data support such a view. If melody and rhythm were entirely integrated, then tapping responses to concordant and discordant patterns would have been similarly variable in Experiment 2, for which the variability of time spans between all accents was similar. On the other hand, if melody and rhythm were processed independently, then the complexity of an accent period ratio should not strongly influence variables like coincidences of individual taps with accents. These observations resemble some previous claims that independence and interaction may figure in different stages of processing for patterns (Peretz & Morais, 1989; Thompson et al., 2001), although an explicit stage model is not proposed here.

Because the current research focuses explicitly on accents, the present results may supplement the roles for melody and rhythm proposed in some other studies. The role of pauses in music and auditory sequences, for instance, is often seen as one of segmentation and grouping (e.g. Bregman, 1990; Deutsch, 1980). Yet the present data support the observation that pauses in music also impart structural markers via accents (e.g., Jones, 1981; Boltz & Jones, 1986). Similarly, many regard pitch contour as instantiating an ordinal trajectory that serves as a factor in melody recognition and the perception of similarity across melodies (e.g. Dowling, 1978; Quinn, 1999; Schmuckler, 1999). The present research adds to these suggestions, by singling out prominent parts of that shape that may serve as attentional targets.

HIGHER ORDER TIME STRUCTURE

Another issue of importance in the literature involves the structural characteristics of music that guide the perception of higher order time structure, lending organizational properties to music such as meter, the "beat," and phrase structure (Lerdahl & Jackendoff, 1983).[3] This study focuses explicitly on (phenomenal) accents, which provide only part of the input to this complex process. Nevertheless, the current data do suggest that accent structures arising from the relationship between melody and rhythm help to direct the listener's attention to a pattern's higher order structure. Other

[3] Many researchers, such as Lerdahl and Jackendoff (1983) argue for a separation between meter and grouping, although many concede that the two concepts often overlap in practice (e.g., Temperley, 2001).

studies have similarly supported the importance of melody and rhythm in meter perception (Large, 2000; Steedman, 1977; Temperley & Bartlette, 2002), although some studies question the salience of melodic information (Snyder & Krumhansl, 2001), and specifically melodic accents (Huron & Royal, 1996).

Although the present study did not examine meter perception directly, analyses of intertap intervals suggests that JAS relates to the perceived meter of a musical pattern. Although listeners were not instructed to track a higher order "beat," some subjects did so, and did so more often for simpler patterns. Overall, accent period ratios, rather than phase relationships, affected intertap intervals, which suggests that higher order accent relationships contribute more to perceived structure than localized phase relationships. This functional distinction between phase and period makes sense, given that a listener's goal in tracking a pattern's higher order structure is typically to isolate an invariant period. Interestingly, a similar dissociation between phase and period has been found in a recent study that investigated relationships between musical accents and periodic visual motion (Lipscomb, in press).

The perception of accents also appears to be influenced by factors beyond those manipulated here. It may be that periodic grouping tendencies, possibly relating to the interpretation of metrical accents, help induce accentuation. Often listeners hear accents that result from the formation of binary groups (Bolton, 1894; Fraisse, 1982; Povel & Okkerman, 1981), which may explain the general avoidance of 3-beat intertap intervals and the common avoidance of many **m** accent locations in Experiment 2. Some avoidances of **m** accent locations in Experiment 2 may also result from a reluctance for listeners to hear accents at two adjacent locations (cf. Liberman & Prince, 1977), given that **m** accents at Positions 2 and 4 (which were rarely tapped to) were adjacent to secondary and primary **r** accent locations, respectively.

One theoretical view compatible with these observations explains monitoring of sequences in real time through activities of internal attending oscillators, capable of entraining to prominently marked time spans at several structural levels (Large & Jones, 1999; Large & Palmer, 2002; McAuley, 1995). The Large and Jones model focuses on low-level time spans marked by successive tone onsets; attentional oscillators respond to "when" onsets of significant markers occur in time (phase parameter) and to marked time spans (period parameter). Similarly, the self-sustaining characteristic of oscillators may explain the tendency for listeners to hear accents as a result of periodic grouping. These current models do not address the role of higher order structure as determined by accents, but may be extended to such a level (e.g., Large, 2000). The periodicities among recurring accents of each type (**m** or **r**) could drive an attending oscillator attuned to that acoustic dimension, with the combined activity of each oscillator reflecting integrated attending to time spans influenced by accent period ratios, as suggested by the current data.

The current study has supported these conclusions through responses to stimuli that possessed certain characteristics of musical structure, that were manipulated in a controlled manner. It constitutes an early and simplified effort to understand highly complex characteristics of musical structure. Given that experimental control is not a common goal of musical composition, the data presented here probably lack the subtlety and complexity of responses to music that is composed for aesthetic purposes, the examination of which is a necessary next step. Nevertheless, experimental control in the patterns used here is valuable, in that the current data show that accents *can* guide the perception of musical pattern structure when not put into conflict with possible metrical interpretations of music (cf. Huron & Royal, 1996). Ultimately, two important issues arise regarding the role of accents in the experience of music's temporal structure: (1) *can* accents contribute to a listener's experience if manipulated in certain ways, and (2) *do* accents function this way in the common practice of music composition? The current study suggests that the answer to the first issue, at least, is "yes."[4]

References

Aschersleben, G., & Prinz, W. (1995). Synchronizing actions with events: The role of sensory information. *Perception & Psychophysics, 57*, 305–317.

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

Bolton, T. L. (1894). Rhythm. *American Journal of Psychology, 6*, 145–238.

Boltz, M. G. (1993). The generation of temporal and melodic expectancies during musical listening. *Perception & Psychophysics, 53*, 585–600.

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

Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound.* Cambridge, MA: MIT Press.

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

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.

---

[4] This research was sponsored in part by NSF grant BCS-9809446 awarded to Mari Riess Jones. Many thanks are due to Mari Riess Jones for her support and guidance during this project. I also thank Caroline Palmer, Heather Moynihan, Ralph Barnes, Devin McAuley, Michael Brady, Edward Large, and an anonymous reviewer for helpful comments on an earlier version of this article. Annie Jordan and Susan Toth assisted with data collection. Thanks are also due to Ralph Barnes for assistance with computer hardware and audio equipment.

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

Desain, P. (1992). A (de)composable theory of rhythm perception. *Music Perception, 9*, 439–454.

Deutsch, D. (1980). The processing of structured and unstructured tonal sequences. *Perception & Psychophysics, 28*, 381–389.

Dowling, W. J. (1978). Scale and contour: Two components of a theory of memory for melodies. *Psychological Review, 85*, 341–354.

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

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., Jones, M. R., & Baruch, C. (2000). The development of rhythmic attending in auditory sequences: Attunement, referent period, focal attending. *Cognition, 77*, 251–288.

Drake, C., Penel, A., & Bigand, E. (2000). Tapping in time with mechanically and expressively performed music. *Music Perception, 18*, 1–24.

Essens, P. (1995). Structuring temporal sequences: Comparison of models and factors of complexity. *Perception & Psychophysics, 57*, 519–532.

Essens, P., & Povel, D. J. (1985). Metrical and nonmetrical representations of temporal patterns. *Perception & Psychophysics, 37*, 1–7.

Fraisse, P. (1982). Rhythm and tempo. In D. Deutsch (Ed.), *The psychology of music* (pp. 149–181). New York : Academic Press.

Halpern, A. R., Bartlett, J. C., & Dowling, W. J. (1998). Perception of mode, rhythm, and contour in unfamiliar melodies: Effects of age and expertise. *Music Perception, 15*, 335–356.

Handel, S. (1989). *Listening: An introduction to the perception of auditory events.* Cambridge, MA: MIT Press.

Herbert, S. & Peretz, I. (1997). Recognition of music in long-term memory: Are melodic and temporal patterns equal partners? *Memory & Cognition, 25*, 518–533.

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

Ivry, R., & Hazeltine, E. (1995). Perception and production of temporal intervals across a range of durations: Evidence for a common timing mechanism. *Journal of Experimental Psychology: Human Perception and Performance, 21*, 3–18.

Jones, M. R. (1976). Time, our lost dimension: Toward a new theory of perception, attention, and memory. *Psychological Review, 83*, 323–355.

Jones, M. R. (1981). Music as a stimulus for psychological motion: Part I. Some determinants of expectancies. *Psychomusicology, 1*, 34–51.

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 : Lawrence Erlbaum.

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

Jones, M. R., & Pfordresher, P. Q. (1997). Tracking melodic events using joint accent structure. *Canadian Journal of Experimental Psychology, 51*, 271–291.

Jones. M. R., & Yee, W. (1997). Sensitivity to time change: The role of context and skill. *Journal of Experimental Psychology: Human Perception and Performance, 23*, 693–709.

Krumhansl, C. L. (2000). Rhythm and pitch in music cognition. *Psychological Bulletin, 126*, 159–179.

Large, E. W. (2000). On synchronizing movements to music. *Human Movement Science, 19*, 527–566.

Large, E. W., & Jones, M. R. (1999) The dynamics of attending: How people track time-varying events. *Psychological Review, 106*, 119–159.

Large, E. W., & Palmer, C. (2002). Perceiving temporal regularity in music. *Cognitive Science, 26*, 1–37.

Large, E. W., Fink, P., & Kelso, J. A. S. (2002). Tracking simple and complex sequences. *Psychological Research, 66*, 3–17.

Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music.* Cambridge, MA : MIT Press.

Liberman, M., & Prince, A. (1977). On stress and linguistic rhythm. *Linguistic Inquiry, 8*, 249–336.

Lipscomb, S. D. (in press). The perception of audio-visual composites: Accent structure alignment of simple stimuli. *Selected Reports in Ethnomusicology.*

McAuley, J. D. (1995). *Perception of time as phase: Toward an adaptive-oscillator model of rhythmic pattern processing.* Unpublished doctoral dissertation, Indiana University at Bloomington.

McAuley, J. D., & Kidd, G. R. (1998) Effect of deviations from temporal expectations on tempo discrimination of isochronous tone sequences. *Journal of Experimental Psychology: Human Perception and Performance, 24*, 1786–1800.

McAuley, J. D., & Semple, P. (1999). The effect of tempo and musical experience on perceived beat. *Australian Journal of Psychology, 51*, 176–187.

Meyer, L. B. (1956). *Emotion and meaning in music.* Chicago : University of Chicago Press.

Monahan, C. B., & Carterette, E. C. (1985). Pitch and duration as determinants of musical space. *Music Perception, 3*, 1–32.

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

Narmour, E. (1990). *The analysis and cognition of basic musical structures: The implication-realization model.* Chicago : University of Chicago Press.

Palmer, C., & Krumhansl, C. L. (1987a). Independent temporal and pitch structures in determination of musical phrases. *Journal of Experimental Psychology: Human Perception and Performance, 13*, 116–126.

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

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

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.*

Quinn, I. (1999). The combinatorial model of pitch contour. *Music Perception, 16*, 439–456.

Repp, B. H. (in press). Phase attraction in sensorimotor synchronization with auditory sequences: Effects of single and periodic distractors on synchronization accuracy. *Journal of Experimental Psychology: Human Perception & Performance.*

Schmuckler, M. A. (1999). Testing models of melodic contour similarity. *Music Perception, 16*, 295–326.

Smith, K. C., & Cuddy, L. L. (1989). Effects of metric and harmonic rhythm on the detection of pitch alterations in melodic sequences. *Journal of Experimental Psychology: Human Perception and Performance, 15*, 457–471.

Snyder, J., & Krumhansl, C. L. (2001). Tapping to ragtime: Cues to pulse finding. *Music Perception, 18*, 455–490.

Steedman, M. J. (1977). The perception of musical rhythm and metre. *Perception, 6*, 555–569.

Tekman, H. G. (1997). Interactions of perceived intensity, duration, and pitch in pure tone sequences. *Music Perception, 14*, 281–294.

Tekman, H. G. (2002). Perceptual integration of timing and intensity variations in the perception of musical accents. *The Journal of General Psychology, 129*, 181–191.

Temperley, D. (2001). *The cognition of basic musical structures.* Cambridge, MA: MIT Press.

Temperley, D., & Bartlette, C. (2002). Parallelism as a factor in metrical analysis. *Music Perception, 20*, 117–150.

Thomassen, J. M. (1982). Melodic accent: Experiments and a tentative model. *Journal of the Acoustical Society of America, 71*, 1596–1605.

Thompson, W. F., Hall, M. D., & Pressing, J. (2001). Illusory conjunctions of pitch and duration in unfamiliar tone sequences. *Journal of Experimental Psychology: Human Perception and Performance, 27*, 128–140.

Todd, R., Boltz, M., & Jones, M. R. (1989). The midilab auditory research system. *Psychomusicology, 8*, 17–30.

Trainor, L. J., Desjardins, R. N., & Rockel, C. (1999). A comparison of contour and interval processing in musicians and nonmusicians using event-related potentials. *Australian Journal of Psychology, 51*, 147–153.

Vos, P. G., van Dijk, A., & Schomaker, L. (1994). Melodic cues for metre. *Perception, 23*, 965–976.

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