Journal of Experimental Psychology:
Human Perception and Performance
2009, Vol. 35, No. 1, 264–280

0096-1523/09/$12.00 DOI: 10.1037/a0013482

The Role of Accent Salience and Joint Accent Structure in Meter Perception

Robert J. Ellis and Mari R. Jones
The Ohio State University

Previous research indicates that temporal accents (TAs; accents due to time changes) play a strong role in meter perception, but evidence favoring a role for melodic accents (MAs; accents due to pitch changes) is mixed. The authors claim that this mixed support for MAs is the result of a failure to control for accent salience and addressed this hypothesis in Experiment 1. Listeners rated the metrical clarity of 13-tone melodies in which the magnitude and pattern of MAs and TAs were varied. Results showed that metrical clarity increased with both MA and TA magnitude. In Experiment 2, listeners were asked to rate metrical clarity in melodies with combinations of MA and TA patterns to allow the authors to ascertain whether these two accent types combined additively or interactively in meter perception. With respect to the additive or interactive debate, the findings highlighted the importance of (a) accent salience, (b) scoring methods, and (c) conceptual versus statistical interpretations of data. Implications for dynamic attending and neuropsychological investigations are discussed.

Keywords: accent salience, meter perception

When a listener taps to a musical event, taps are typically in synchrony with a perceived "beat," reflecting a perceived temporal regularity in the musical surface (cf. Snyder & Krumhansl, 2001; Temperley, 2001). How these beats are organized in time is one of the functions of *meter*. Meter is fundamental to music; it is what makes a waltz a waltz and a march a march. It is also significant psychologically, in that the perception of meter depends on a listener's ability to recover temporal regularity from a sound pattern containing temporal irregularities. Furthermore, because of its periodic nature, an established metrical framework affords a basis for a listener to generate anticipations about timings of future events, thereby guiding attending in a dynamic fashion (Jones, 1976, 1990, 2001, 2004; London, 2001). Picking up on temporal regularity in a musical event facilitates listener performance in a number of tasks, including (a) judgments of event duration (e.g., Boltz, 1991, 1998; Jones & Boltz, 1989), (b) judgments of melodic phrase completeness (e.g., Boltz, 1989a), (c) discrimination of pitch changes (e.g., Jones, Boltz, & Kidd, 1982; Jones, Johnston & Puente, 2006; Jones, Moynihan, MacKenzie, & Puente, 2002), and (d) discrimination of time changes (e.g., Large & Jones, 1999). Even very young listeners are sensitive to differences between metrical frameworks (Hannon & Johnson, 2005; Hannon & Trehub, 2005, 2006).

An important contributor to meter is the relative timing of *accents*. An accent refers to a tone in a sequence that stands out from other tones along some auditory dimension (e.g., pitch, intensity, timbre, duration). More concretely, an accent is a "deviation from a norm that is contextually established by serial constraints" (Jones, 1987, p. 624); thus, an accent acquires its status from surrounding tones (Cooper & Meyer, 1960). Accented tones often correspond to stronger beats (Cooper & Meyer, 1960; Handel, 1989; Huron & Royal, 1996; Jones, 1987, 1993; Lerdahl & Jackendoff, 1983).

Two common types of accents are melodic accents (MAs; accents due to pitch relationships) and temporal accents (TAs; accents due to time relationships). The precise nature of MA and TA contributions to music perception continues to be an important research topic. One line of research has sought to determine whether MAs and TAs, taken together, make additive (e.g., Palmer & Krumhansl 1987a, 1987b) or interactive (e.g., Jones, 1987, 1993; Jones & Boltz, 1989) contributions to the listening experience. As Schellenberg, Krysciak, and Campbell (2000) have pointed out, the answer to this question has implications for low-level representations of pitch and duration as well as higher level representations of phrase grouping, expectancies, and emotion, making it a critical issue to the field of music psychology.

Alternatively, other research suggests that MAs contribute negligibly to music perception (e.g., Hannon, Snyder, Eerola, & Krumhansl, 2004; Huron & Royal, 1996; Snyder & Krumhansl, 2001). However, such findings create a logical problem: If MAs fail to reliably contribute to meter perception, then discussions of whether MAs and TAs combine additively or interactively in meter perception become moot. Thus, before addressing the issue of additivity versus interactivity (the focus of Experiment 2), we first examine the mixed evidence regarding the impact of MAs in meter perception. In reviewing this literature, we note that one overlooked reason for mixed outcomes involves the stimuli in-

---

Robert J. Ellis and Mari Riess Jones, Department of Psychology, The Ohio State University.

This research was completed as part of Robert J. Ellis's master's thesis and was generously supported by the Caroline B. Monahan Fund for Experimental Research Support in the Music Cognition/Perception Area within the Department of Psychology. We gratefully acknowledge comments from Devin McAuley and Molly Henry and extensive feedback from Peter Pfordresher.

Correspondence concerning this article should be addressed to Robert J. Ellis, 175 Psychology, The Ohio State University, 1835 Neil Ave., Columbus, OH 43201. E-mail: ellis.306@osu.edu

volved. Stimuli with TAs that are more pronounced than MAs may bias listeners toward TA information. Only when the salience of MAs is comparable to that of TAs can their relative contributions (within the same melody) be assessed. As an analogy, suppose that a study reports that subjects exhibit a preference for eating apples from one bowl over oranges from another, but neglects to report that the bowl of oranges was out of subjects' reach. This conclusion is meaningless: Only if the bowls of apples and oranges were equally accessible would the finding that subjects preferred one over the other be interesting.

Issues of additivity versus interactivity in perception were addressed in classic work by Garner (1970, 1974), who proposed that relationships between two dimensions exist along a continuum ranging from *separable* (i.e., additive) to *integral* (i.e., interactive). However, Garner also noted that when two constituent stimulus dimensions are not matched for salience, the more salient dimension can affect the perception of the less salient dimension, thereby obscuring these relationships (Garner & Felfoldy, 1970; see also Melara & Marks, 1993, 1994; Tekman, 1997, 1998, 2001). Our investigation of MA and TA salience fits within this broader context.

In what follows, we outline two opposing hypotheses about the respective roles of MAs and TAs in meter perception. The *joint accent structure hypothesis*, which posits roles for both MAs and TAs, is contrasted with a *temporal accent bias hypothesis*, which holds that meter perception depends primarily on TAs. Both hypotheses have received support, but because previous investigations have failed to systematically calibrate the salience of MAs and TAs in the stimuli employed, their conclusions remain tenuous. This leads us to formulate a third, an *accent salience hypothesis*, which makes explicit the relationship between serial change and accent salience.

Meter, Accent Types, and Accent Tokens

Meter involves a succession of strong and weak beats, equally spaced in time and organized into metrical frameworks (Lerdahl & Jackendoff, 1983; Justus & Bharucha, 2002). Two of the most common metrical frameworks in the Western classic tradition are duple and triple meter (cf. Huron, 2006, p. 195). Duple meter refers to a pattern of alternating strong (S) and weak (w) beats (SwSwSw . . .), frequently heard in marches. Triple meter, consisting of a strong beat followed by two weak beats (SwwSww . . .), is associated with waltzes. Meter implies temporal invariance in two levels: shorter (lower order) time spans marked by the inter-onset intervals (IOIs) between successive S and w elements; and longer (higher order) time spans, or metrical periods, marked by successive S elements (cf. Benjamin, 1984; Lerdahl & Jackendoff, 1983; Yeston, 1976).

Accents are used to mark strong beats and hence are important to meter (Cooper & Meyer, 1960; Handel, 1989; Huron & Royal, 1996; Jones, 1987; 1993; Lerdahl & Jackendoff, 1983). In theory, accents of different types correspond to defining acoustic dimensions (e.g., pitch, time). Our approach goes further than others in its explicit acknowledgment of the role of time in defining accents of all types. That is, accent *type* is taken to mean a family of salient local changes over time along a common dimension (e.g., pitch, loudness, duration). In auditory sequences, serial variations along each dimension (type) express *token* accents of that type (type/token terminology is useful here in distinguishing between general and particular; cf. Esposito, 1998; Levelt, 1989). For example, a salient pitch change might occur if three successive tones, each ascending by 1 semitone (ST), were followed by a pitch that ascended by 5 ST. A salient time change might occur if three successive short IOIs were followed by a long IOI.

The MA type refers broadly to any salient local serial change in pitch relationships and has three common tokens (cf. Huron & Royal, 1996; Jones, 1987, 1993). First, a *pitch-contour accent* depends on a temporal ordering of pitches; it results from a local change in the direction of pitch trajectory (e.g., ascending to descending), with accentuation on the inflection point (cf. Thomassen, 1982). Second, a *pitch-leap accent* falls on the second of two tones forming a pitch interval, when that interval is larger than preceding pitch intervals in a series (cf. Tekman, 1997, 1998; Thomassen, 1982). Third, a *tonal accent* arises from a serial shift in "stability" within a tonal context, that is, from the leading tone to the tonic (keynote; cf. Bharucha, 1984; Dawe, Platt, & Racine, 1993; Smith & Cuddy, 1989).

The TA type refers to any salient local serial change in time relationships and has two common tokens. A *rhythmic accent* (or *pause accent*; e.g., Jones & Pfordresher, 1997) results from a change within a serial pattern of IOIs (e.g., three 200-ms IOIs followed by a 600-ms IOI), and its location depends on the serial context (Jones, 1987; Jones & Pfordresher, 1997; Narmour, 1996; Povel & Essens, 1985; Povel & Okkerman, 1981). A *duration accent* occurs on a tone that has a longer duration (i.e., from tone onset to offset) than neighboring tones in a sequence (Woodrow, 1951; cf. Castellano, Bharucha, & Krumhansl, 1984).

The Role of MAs in Meter Perception: Two Contrasting Views

Dynamic Attending and Joint Accent Structure

One view that favors a role for both MAs and TAs in meter perception comes from *dynamic attending* theory (Jones, 1976, 2004; Jones & Boltz, 1989; Large & Jones, 1999; McAuley & Jones, 2003). This approach assumes that recurrent time spans within real-world events can synchronize attending via the mechanism of entrainment. Entrainment is the physical process whereby internal attending periodicities become attuned to salient recurrent stimulus time spans. Resulting attentional synchronies are possible at multiple time scales, and they are facilitated when time spans at different time scales are hierarchically nested. In auditory patterns, such time spans are marked by accents that arise from salient serial changes in pitch (MAs) and/or timing (TAs). When both MAs and TAs are present in a single pattern, they contribute to the emergence of a common higher order time structure, or *joint accent structure* (JAS; e.g., Boltz & Jones, 1986; Jones, 1987, 1993).

At a basic level, a JAS refers to the relative timings of different accent types within a melody. These properties are illustrated in Figures 1 and 2. In Figure 1, local serial changes in a single, isochronous, melodic line instantiate both pitch-leap accents (an MA token; large open circles) and duration accents (a TA token; large solid circles). In this example, the accent periods of both MAs and TAs are consistently double the lower order time spans (IOIs). This relationship results in invariant accent periods within a musical event. According to dynamic attending theory, invariant

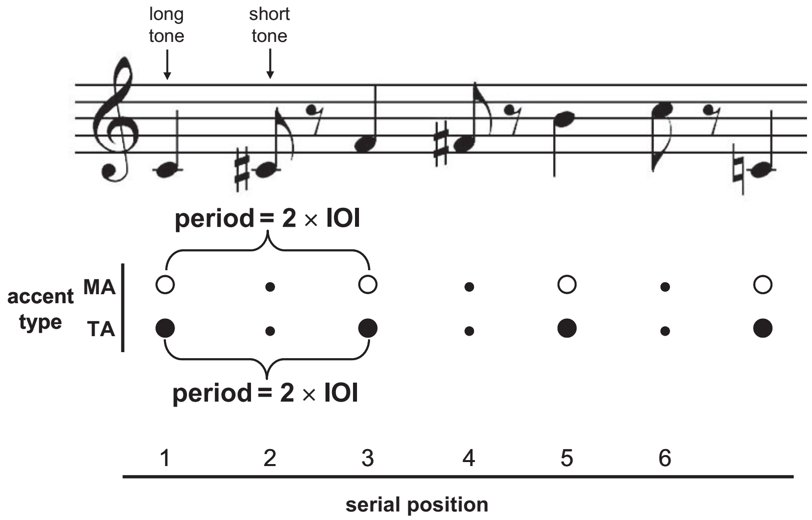

Figure 1. An example of the joint accent structure (JAS) for an isochronous melody. Melodic accents (MAs) temporal accents (TAs), and unaccented tones are indicated by the large open circles, large solid circles, and small circles, respectively. IOI = inter-onset intervals.

accent timing should promote more effective entrainment than variable accent timing (Large & Jones, 1999).

Figure 2 illustrates four JASs that emerge from combinations of duple (D = 2 × IOI) and triple (T = 3 × IOI) MA and TA patterns.[1] Throughout this article, we use the phrase *accent pattern* (either MA pattern or TA pattern) to refer to a temporal succession of accents and the phrase *meter* to refer to listeners' percepts. We use a shorthand notation to describe JASs, with each accent period (D, T) subscripted by its accent type (MA, TA); for example, $D_{MA}T_{TA}$ denotes a melody with a duple MA pattern and a triple TA pattern. Complexity refers to the overall irregularity of accent timing; JAS complexity depends upon constituent accent patterns (i.e., both MA and TA periods). An index of JAS complexity is given by the quotient (or ratio; e.g., Jones, 1987) of the larger accent period divided by the smaller accent period. In both Figures 1 and 2A, for example, the MA and the TA patterns are both duple ($D_{MA}D_{TA}$), and the accent period quotient is 1 (= 2/2), indicating low temporal complexity (high regularity). In Figure 2B ($T_{MA}T_{TA}$), both accent patterns are triple, and the quotient is again 1 (= 3/3), again indicating low complexity. In general, small integer quotients (1, 2, 3, 4) indicate low temporal complexity. These highly regular JAS patterns are termed *concordant* (Jones & Pfordresher, 1997). By contrast, more complex JASs emerge when duple and triple accent patterns are both present in a single melody, as in Figures 2C ($T_{MA}D_{TA}$) and 2D ($D_{MA}T_{TA}$). In these patterns, accent timing is less regular, as indexed by a noninteger accent period quotient, 1.5 (= 3/2). These JASs are termed *discordant*.

Dynamic attending theory implies that lower JAS complexity facilitates listener comprehension, due to more efficient entrainment. This view is supported by empirical evidence. When listening to melodies with a concordant JAS (compared with listening to discordant-JAS melodies), listeners (a) judge the melodies' endings as more complete (Boltz, 1989a, 1989b), (b) reproduce the melodies' durations more accurately (Boltz, 1998; Jones, Boltz, & Klein, 1993), (c) recall melodic structure more accurately (Boltz, 1991; Boltz & Jones, 1986; Deutsch, 1980; Drake, Dowling, & Palmer, 1991), (d) detect deviant tones more accurately (Dowling, Lung, & Herrbold, 1987; Monahan, Kendall, & Carterette, 1987), (e) identify constituent accent patterns more accurately (Keller & Burnham, 2005), and (f) synchronize to constituent tones more precisely (Jones & Pfordresher, 1997; Pfordresher, 2003). In addition, the temporal structure of a sequence can highlight the pitch or tonal structure of a melody (Jones et al., 1982; Laden, 1994).

In the dynamic attending approach, specific predictions are made about meter. In discordant-JAS melodies, temporal complexities result in less efficient entrainment, reducing a listener's ability to track successive tones in real time (Jones & Pfordresher, 1997; Pfordresher, 2003) and reducing metrical clarity. A key prediction is that concordant JASs will elicit faster and stronger perceptions of metrical clarity than discordant JASs, due to facilitation from regularly timed accents in concordant JASs and/or interference from conflicting accent periods in discordant JASs. This prediction assumes that MAs and TAs have roughly equal salience. We restate this as a *JAS hypothesis*: When multiple accent types or tokens are present in a melody, meter perception is determined by JAS properties. Thus, a concordant JAS will elicit the percept of a meter more clearly and immediately than a discordant JAS.

---

[1] In these patterns, all initial phase differences between accent patterns were 0; that is, both duple and triple accent patterns begin with an accent on Tone 1. Accent phasing has been considered elsewhere (Boltz & Jones, 1986; Jones, 1987, 1993; Pfordresher, 2003). Phase shifting two accent periods (that are otherwise concordant) leads to slightly worse performance (i.e., in melody reproduction or tapping synchronization) than nonshifted concordant melodies, but better performance than melodies with discordant JASs (whether phase-shifted or not), depending on the phase of a shift. In the present study, setting the initial phase to 0 was necessary to preserve global aspects of melodies (i.e., two ascending six-tone cells, with accents on Tones 1, 7, and 13) despite variation in constituent accent patterns over the session; see the Methods section of Experiment 1 for more details.

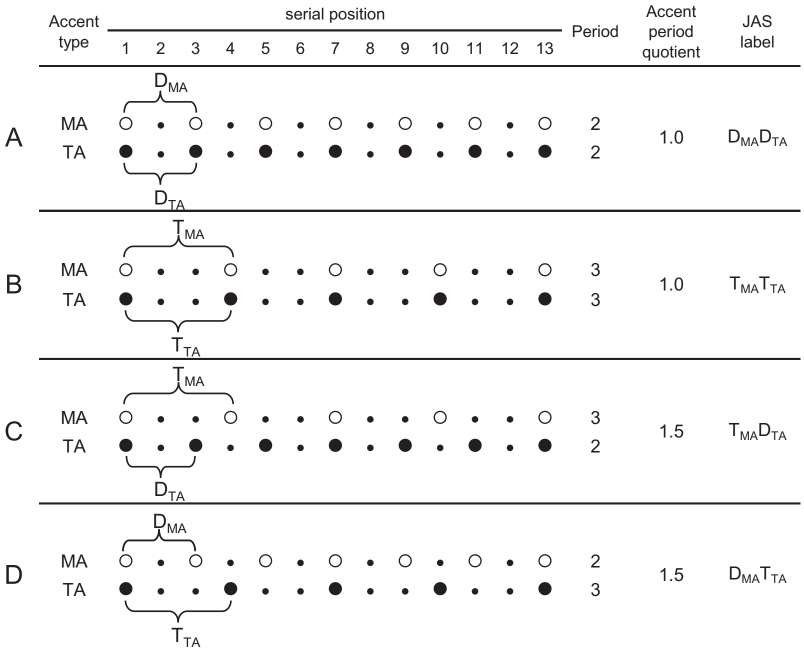

Figure 2. A schematic illustration of four joint accent structures (JASs) formed by combinations of accent periods (duple, triple) marked by melodic accents (MA) and temporal accents (TA). All tone onsets are isochronous. Concordant JASs appear in Panels A and B, and discordant JASs in Panels C and D.

We note that the dynamic attending approach is not the only theory in which MAs are posited to play a role in perception (cf. Dixon & Cambouropoulos, 2000; Temperley & Bartlette, 2002; Toiviainen and Snyder, 2003). However, dynamic attending theory differs from other approaches in its incorporation of (a) accents as serial changes, (b) JAS patterns, and (c) entrainment activities.

Temporal Accent Bias

Although the JAS hypothesis assumes that both MA and TA accent patterns can contribute equally to meter, this assumption has been questioned. In a variety of tasks, MAs have been found to contribute relatively little information compared with TAs. In an early study, Woodrow (1911) found that MAs were inconsistent in their ability to induce tone groupings; some listeners heard them as beginning a group, others as ending a group. Monahan and Carterette (1985) examined ratings of similarity for pairs of melodies that varied in both MA and TA properties and found that most subjects based their ratings on TA, not MA, properties. Snyder and Krumhansl (2001) asked subjects to “tap to the beat” of two versions of ragtime piano music (rhythms only vs. pitch plus rhythms); they found that the addition of pitch information did not facilitate beat finding. Hannon et al. (2004) found that listeners’ ratings of meter in folk songs relied on both MA (including pitch-leap, contour) and TA (rhythmic accents, duration accents) factors, but MAs failed to influence ratings significantly in melodies that contained both TAs and MAs. They concluded that TAs have a “primary role” in predicting meter (p. 968). It is noteworthy, however, that MA factors were more predictive in melodies that contained temporally regular MA patterns.

Additional evidence favoring the TA bias hypothesis comes from an analysis of musical scores. Randomly sampling a large corpus of folk tunes, Huron and Royal (1996) correlated temporal locations of MA and TA tokens with their respective metrical positions (as in scored time signatures). Only rhythmic accents yielded positive correlations between accent size and scored metrical position (their Experiment 1). Furthermore, even in melodies containing no TAs (their Experiment 2), neither pitch-leap nor pitch-contour accents significantly correlated with the strong metrical positions. Huron and Royal (1996) concluded that this “calls into question” the claim that pitch-leaps function as accents (p. 501).

Finally, a number of theoretical approaches also assume (explicitly or tacitly) that percepts of meter are biased toward temporal accent patterns (e.g., Desain & Honing, 1989; Drake & Bertrand, 2001; Johnson-Laird, 1991; Longuet-Higgins & Lee, 1982; Parncutt, 1994; Povel & Essens, 1985; for a review, see Clarke, 1999). Together with the empirical findings, these accounts can be summarized in a *TA bias hypothesis*: Metrical percepts are biased toward the pattern of temporal (rather than melodic) accents. When both MA and TA patterns are present, listeners are biased to use TA patterns when perceiving meter.

Accent Salience: A Neglected Concept

The JAS hypothesis assumes that both MAs and TAs are salient, whereas the TA bias hypothesis assumes that TAs are selectively salient for meter perception. The former approach considers MA salience necessary; the latter approach treats MA salience as irrelevant. No empirical investigations have systematically examined the relative salience of MAs and TAs. In some investigations, accent locations are defined a priori in experimenter-created stimuli. For example, MA locations might be defined as contour inflections or tones of a tonic triad, and TAs as the tones following a pause (i.e., a missing beat) in an otherwise isochronous sequence (Boltz, 1991; Boltz & Jones, 1986; Jones & Pfordresher, 1997; Pfordresher, 2003). Better performance with concordant JASs than discordant JASs provides only indirect support for MA salience, because the degree to which MA factors contribute to the JAS is left unexplored. To wit, Monahan and Carterette (1985) hypothesize that TA salience might have been greater than MA salience in their melodies, potentially explaining listeners' reliance on TA information in judging similarity.

A disparity in salience between MAs and TAs may be more pronounced in research that examines listener responses to original compositions (Boltz, 1989a, 1989b, 1989c, 1991; Hannon et al., 2004; Palmer & Krumhansl, 1987a, 1987b; Snyder & Krumhansl, 2001) or in analyses of metrical information within excerpts of "real" music (Dixon & Cambouropoulos, 2000; Huron & Royal, 1996; Temperley & Bartlette, 2002; Toiviainen & Snyder, 2003). Although dealing with naturalistic stimuli increases ecological validity, it denies the experimenter control of important independent variables such as accent magnitude. Failure to control for differential MA versus TA salience sets the stage for misleading conclusions. For instance, if a selected excerpt happens to contain TAs that are more salient than MAs, then a listener would likely ignore the latter and use only TAs to aid meter perception. On the surface, such a finding would appear to favor the TA bias hypothesis. But because the salience of constituent MAs and TAs (in isolation) is typically unknown, little can be claimed about their relative contributions when present together. Only when MAs and TAs are experimentally equated for salience can their relationship be validly assessed.

If accent salience depends upon serial change along a dimension, then a resolution of conflicting evidence about the role of MAs (vs. TAs) in meter perception (highlighted in contrasting predictions of the JAS and TA bias hypotheses) may turn on calibrating accent salience. We argue that these conflicting findings reflect a research orientation to accent structure in which differences in the salience of various accent types are overlooked or left uncalibrated. To address this, we offer an operational definition of accent salience involving two properties: (a) the magnitude of a change along a given acoustic dimension and (b) the number of serial changes associated with that tone.

In linking accent salience to the magnitude of a local serial change along some dimension, we should note an important qualification that involves global context. All local changes necessarily happen within a larger prevailing (global) serial context. Thus, the salience of any local serial change is modulated by the structural variability within a surrounding context. For instance, greater global variability in rhythmic patterns lowers the salience of local time changes (Jones & Yee, 1997; Large & Jones, 1999, Experiments 1 and 3). Similarly, a musical event containing a wide pitch range can attenuate the salience of embedded MAs, thereby leading to conclusions consistent with the TA bias hypothesis, a point we consider in the General Discussion. Because neither local nor global aspects of accent salience have been controlled in previous research, we manipulated the magnitude of local serial changes of two accent tokens (pitch-leap accents and duration accents) by holding at feasible minima the variability in pitch and timing within a global serial context (details in the Methods section of Experiment 1). All other things equal, accent salience should be systematically affected by the magnitude of these local serial changes.

Jones (1987, 1993) also argued that the accentuation potential of any tone depends upon the number of accents that coincide on that tone. Here we propose that accent salience monotonically increases with the number of local serial changes on a given tone. For example, if a tone is both lengthened in duration and a contour inflection, then it should have greater salience than a tone with only one of these two accent tokens. Because little evidence supports such claims, we sought to evaluate them in Experiment 2.

An *accent salience hypothesis* addresses both of these issues. It holds that the salience of a tone as an accent increases with (a) the magnitude of local serial changes associated with that tone and (b) the number of serial changes (in different dimensions) associated with that tone. Thus, a 5-ST pitch leap should have greater salience than a 2-ST pitch leap, and a 5-ST in combination with a 100-ms tone duration (vs. a 60-ms tone duration) should have greater salience than either a 5-ST pitch leap or a 100-ms duration alone.

Plan of Experiments

We addressed three main questions in this research. First, in Experiment 1, we asked, "In simple melodies, what magnitude of serial change (e.g., in a pitch leap, or a tone lengthening) makes an accent salient?" In this experiment, we tested the accent salience hypothesis prediction that accent salience increases with the local magnitude of individual serial changes. We manipulated the magnitude of pitch-leap accents (MAs) and duration accents (TAs) across different melodies that contained minimal global variation in serial structure (details in Experiment 1, Methods). We chose these particular accent tokens because their magnitudes are easy to quantify (i.e., as the number of semitones in the pitch leap or the duration of the tone in milliseconds) and hence to "titrate." Listeners heard 13-tone melodies with patterns of pitch-leap or duration accents and rated perceived metrical clarity.

Second, in Experiment 2, we asked, "Do concordant patterns, which contain two different coinciding accent types (MA plus TA), induce meter more clearly than patterns with having only a single accent type?" This experiment evaluated the other prediction of the accent salience hypothesis: that metrical clarity increases with the number of different (coinciding) accents. We tested this by having MAs and TAs occur either together or separately in different melodies.

Third, in Experiment 2, we also evaluated predictions of the JAS hypothesis versus those of the TA bias hypothesis using melodies that contained both MAs and TAs. These hypotheses offer different answers to the question, "Do melody and rhythm interact such that concordant JASs facilitate quicker and clearer metrical percepts than discordant JASs?" The JAS hypothesis answer is "Yes": a concordant JAS should facilitate meter perception, whereas a discordant JAS should hinder it. The TA bias hypothesis answer is "No": salient MAs in JASs will not interfere with meter perception, which depends solely upon the pattern of TAs.

Experiment 1: MA and TA Salience

In Experiment 1, we systematically manipulated the magnitude of local serial changes along two different dimensions (pitch and time) to assess the accent salience hypothesis. In different melodies, the two accent types (MA, TA) were embedded, respectively, as pitch-leap and duration tokens at serial locations specified by either duple or triple metrical frameworks. Our goal was to evaluate the accent salience hypothesis prediction that a listener's ability to differentiate duple from triple meter improves as accent magnitude (i.e., local serial change) increases. We controlled the variability of both the melodic and the temporal global serial context that surrounded local serial changes by creating melodies based on simple, ascending pitch trajectories that unfolded with uniform IOIs.

A related goal was to uncover MA and TA patterns of comparable salience levels in preparation for subsequent evaluations of the JAS and temporal bias hypotheses in Experiment 2. To calibrate accent salience ratings, we built upon the magnitude matching paradigm of J. C. Stevens & Marks (1980; cf. Marks & Gescheider, 2002). Unlike the cross-modality matching paradigm (J. C. Stevens & Marks, 1965; S. S. Stevens, 1959), wherein subjects adjusted the level of one dimension (e.g., brightness) to match the level of another (e.g., loudness), the magnitude matching paradigm required subjects to rate brightness and loudness separately, but on the same numerical scale. Stimulus values of lights and tones receiving the same ratings became x and y coordinates on the cross-modal function. In the present study, melodies with MA patterns or TA patterns were both rated on a common 6-point scale of metrical clarity. Magnitudes of MAs and TAs that produced similar ratings from subjects were considered comparably salient.

Method

Subjects. Twenty-eight Ohio State University (OSU) undergraduates in psychology (who participated for course credit) were randomly assigned in equal numbers to two presentation orders. Subjects had an average of 4.6 years of formal musical training ($SD = 3.3$; range = 0–12); all reported normal hearing.

Apparatus. Stimuli were programmed on a PC-compatible 200 MHz Pentium computer (Intel, Santa Clara, CA) running Version 6.0 of the MIDILAB program (Todd, Boltz, & Jones, 1989). The computer interfaced with a Roland MPU-401 MIDI processing unit (Roland Corp., Hamamatsu, Japan), which controlled a Yamaha TX81Z FM tone generator (Yamaha Corp., Hamamatsu, Japan) set to the "sine wave" voice. Stimuli were amplified by a Kenwood KA-5700 amplifier (Kenwood USA, Long Beach, CA) and delivered individually to subjects through AKG K-270 headphones (AKG Acoustics, Vienna, Austria).

Stimuli and conditions. Thirty-two unique, 13-tone sequences ("melodies") were created with either pitch-leap accents (an MA) or duration accents (a TA). In all melodies, IOIs between successive tones were always 500 ms. All melodies shared the same basic pitch contour structure: two identical, ascending 6-tone "cells" (Tones 1–6, Tones 7–12) followed by a 13th tone (a repeat of tones 1 and 7). Table 1 shows accent serial locations in duple and triple accent patterns. Accents in the duple pattern occurred at Tones 1, 3, 5, 7, 9, 11, and 13; in the triple pattern, they occurred at Tones 1, 4, 7, 10, and 13. These constraints preserved global pattern structure over the experiment; each 6-tone cell could either have three groups of 2 tones or two groups of 3 tones.

Our choice of pitch-leap and duration accents was also motivated by global pattern structure. While contour accents have been used in previous investigations of meter (e.g., Jones & Pfordresher, 1997; Pfordresher, 2003), using them as the MA here would introduce dramatic differences in melodic contour shape between duple and triple accent patterns. Using pitch leap accents allowed us to hold the locations of contour inflections (i.e., pitch peaks)

Table 1
A Schematic Illustration of the Isochronous Melodies Used in Experiments 1 and 2

	Serial position												
Accent pattern/magnitude	1	2	3	4	5	6	7	8	9	10	11	12	13
							Temporal accents[1]						
Duple	●	•	●	•	●	•	●	•	●	•	●	•	●
Triple	●	•	•	●	•	•	●	•	•	●	•	•	●
							Melodic accents[2]						
Duple													
2 ST	F5	F$^{\#}$5	**G$^{\#}$5**	A5	**B5**	C6	**F5**	F$^{\#}$5	**G$^{\#}$5**	A5	**B5**	C6	**F5**
3 ST	D$^{\#}$5	E5	**G5**	G$^{\#}$5	**B5**	C6	**D$^{\#}$5**	E5	**G5**	G$^{\#}$5	**B5**	C6	**D$^{\#}$5**
4 ST	C$^{\#}$5	D5	**F$^{\#}$5**	G5	**B5**	C6	**C$^{\#}$5**	D5	**F$^{\#}$5**	G5	**B5**	C6	**C$^{\#}$5**
5 ST	B4	C5	**F5**	F$^{\#}$5	**B5**	C6	**B4**	C5	**F5**	F$^{\#}$5	**B5**	C6	**B4**
Triple													
2 ST	F$^{\#}$5	G5	G$^{\#}$5	**A$^{\#}$5**	B5	C6	**F$^{\#}$5**	G5	G$^{\#}$5	**A$^{\#}$5**	B5	C6	**F$^{\#}$5**
3 ST	F5	F$^{\#}$5	G5	**A$^{\#}$5**	B5	C6	**F5**	F$^{\#}$5	G5	**A$^{\#}$5**	B5	C6	**F5**
4 ST	E5	F5	F$^{\#}$5	**A$^{\#}$5**	B5	C6	**E5**	F5	F$^{\#}$5	**A$^{\#}$5**	B5	C6	**E5**
5 ST	D$^{\#}$5	E5	F5	**A$^{\#}$5**	B5	C6	**D$^{\#}$5**	E5	F5	**A$^{\#}$5**	B5	C6	**D$^{\#}$5**

Note. Accent patterns are duple and triple. Magnitude is measured semitones (ST).

[1] Temporal accents (large circles) had a constant magnitude within a pattern: 80-ms, 100-ms, 120-ms, or 140-ms tone durations. Unaccented tones (small circles) had a duration of 60 ms.

[2] Melodic accents are highlighted in bold. Only those melodies with a pitch peak of C6 are shown in this table; patterns with a pitch peak of A5 were constructed similarly.

constant throughout the experiment. In a similar vein, using rhythmic accents to create duple versus triple accent patterns would have introduced variability in the IOI structure of melodies across the experiment. Duration accents preserve isochrony, helping to minimize global pattern variability.

Duration accents assumed one of four magnitudes over the course of the experiment: 80 ms, 100 ms, 120 ms, or 140 ms. Tones without accents had a duration of 60 ms. Interstimulus intervals (i.e., the time between the offset of one tone to the onset of the next) were adjusted to preserve sequence isochrony.

Pitch-leap accents also assumed one of four magnitudes: 2 ST, 3 ST, 4 ST, or 5 ST; tones without pitch-leap accents were 1 ST above preceding tones. Both the number and magnitude of a pitch leap increased the distance between the lowest and highest pitches of a melody, which ranged from 5 ST to 13 ST. Accordingly, the starting pitch of a melody was adjusted so that the highest pitch of that melody (the *pitch peak*) was either A5 or C6. Pitch peaks locations were constant over all melodies (Tones 6 and 12), and were not pitch-leap accents.

Tones 1, 7, and 13 did not have the same pitch-leap accent magnitudes as other accent locations. Tones 7 and 13 were low pitches, preceded by a pitch that could be up to 13 ST higher, making the direction of the leap different from other accents. Tone 1 could only form an interval with Tone 2, precluding it from being a pitch-leap accent. However, because Tones 1, 7, and 13 were common to both duple and triple frameworks, the presence of a (potentially larger) pitch accent at these locations could not be used to differentiate them.

Design. The primary design was 2 × 2 × 4, with accent type (MA, TA), accent pattern (duple, triple), and accent magnitude (four values) as within-subjects factors. All melodies were presented twice, in one of two presentation orders. Trial-to-trial constraints held that no three consecutive melodies had the same accent type, pattern, magnitude, or pitch peak.

Procedure. Subjects listened to recorded instructions while observing a task diagram. They were asked to attend to the grouping pattern of strong and weak tones within each melody. Subjects had up to 3.5 s following sequence cessation to press one of seven horizontally ordered buttons on a MIDILAB response box to indicate grouping clarity. From left to right, buttons were labeled as follows: *very clear groups of two*, *moderately clear groups of two*, *slight groups of two, neutral/can't decide, slight groups of three, moderately clear groups of three,* and *very clear groups of three*. Verbal, rather than numerical, labels were used to minimize intersubject variability in scale use (cf. Borg, 1982; Marks & Gescheider, 2002). Subjects were told that the task was subjective and were given no instructions regarding response speed other than to withhold a response until each melody finished.

Subjects received six practice trials with feedback. To help orient listeners, we used larger accent magnitudes in the practice trials than in the experimental trials (pitch-leap accent = 7 ST; duration accent = 180 ms); that is, melodies were either strongly duple or strongly triple. Responses were made on the seven-button box. Following a response, an LED screen on the response box displayed a single digit indicating the response that "most people might have given": 2 for *groups of two* and 3 for *groups of three*. Following practice, subjects heard each of the 32 melodies twice over the course of the experiment, without feedback.

Subjects were randomly assigned to one of two pseudorandom presentation orders. Within each order, no three successive melodies had the same pitch peak, accent type, or accent magnitude. At the end of the experiment, subjects completed a questionnaire on their musical background (training, musical preferences), task perceptions, and strategies.

Data reduction. Collapsed over the pitch peak variable, each melody was heard four times over the course of the experiment. To quantify a subject's perception of each melody, we mapped the seven MIDILAB buttons onto the integers from –3 to 3, from left to right. A signed clarity score (C score) was then calculated as the mean of the four responses to that melody. The sign of a C score identifies the meter choice: negative for duple and positive for triple. The number itself reflects the clarity of the identified meter, with 1 corresponding to *slightly clear* and 3 corresponding to *very clear*. A C score near 0 indicates that a subject either (a) consistently pressed the *neutral/can't decide* button or (b) pressed a combination of both *groups of two* and *groups of three* buttons, suggesting overall uncertainty about the meter. Each of the 28 subjects produced 16 C scores (2 accent types × 2 accent patterns × 4 accent magnitudes), resulting in a total of 448 data points.

Results

In all experiments, we report the eta-square values, calculated from sums of squares (SS) tables as SS_{effect}/SS_{total}, which indicates the proportion of variance uniquely explained by an effect (Keppel & Wickens, 2004).

Signed clarity scores. Signed C scores for MA melodies and TA melodies were analyzed separately, each with a 2 (accent pattern) × 4 (accent magnitude) repeated-measures analysis of variance (ANOVA). We used care in performing ANOVAs that included accent type as a factor, because such a design implies, a priori, certain equivalences between MA and TA magnitudes (i.e., assigning 2-ST pitch-leap accents and 80-ms duration accents as the first level in a unified accent magnitude variable). A main effect for accent type would thus be meaningless.

Figure 3 presents C score means as a function of accent type (MA, TA), accent pattern (duple, triple), and accent magnitude (four levels). A significant main effect appeared for accent pattern in both MA melodies, $F(1, 27) = 48.05$, $\eta^2 = .31$, $p < .001$, and TA melodies, $F(1, 27) = 125.41$, $\eta^2 = .51$, $p < .001$. Melodies with a duple accent pattern received negative scores (indicating that subjects heard tones grouped by twos) and melodies with a triple accent pattern received positive scores (tones grouped by threes). Accent pattern also interacted with accent magnitude in both ANOVAs. Scores to duple melodies became more negative, and scores to triple melodies more positive, as TA accent magnitude increased, $F(3, 60) = 224.69$, $\eta^2 = .26$, $p < .001$. Scores to triple melodies became more positive as MA magnitude increased, but scores to duple melodies were not modulated by MA magnitude, $F(3, 57) = 17.63$, $\eta^2 = .10$, $p < .001$. In general, differences between duple meter and triple meter increased with accent size.

If a given accent magnitude has the potential to differentiate duple from triple meter, then C scores for melodies with duple and triple patterns should be statistically different. We performed four planned comparisons from within the Accent Pattern × Accent Magnitude interaction (one for each level of accent magnitude) in

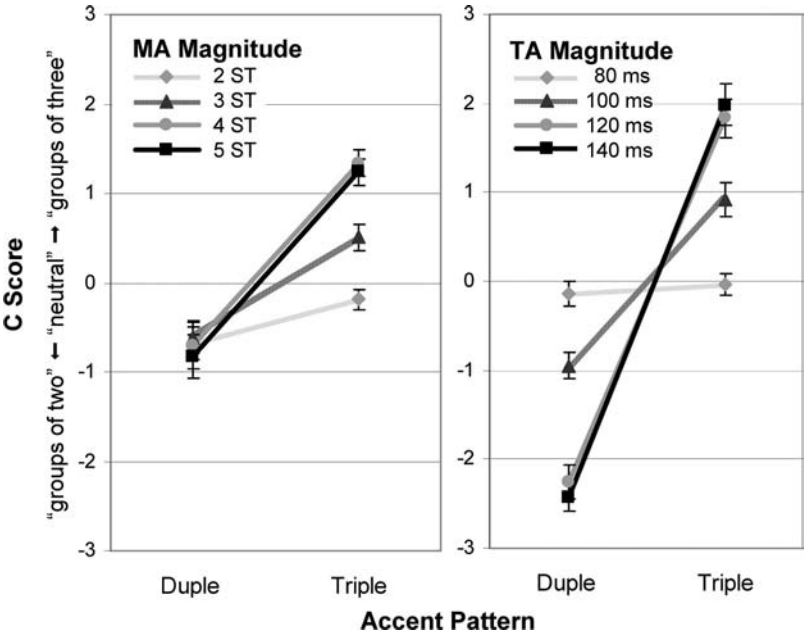

Figure 3. Experiment 1: Signed clarity score (C score) means as a function of accent type, accent pattern, and accent magnitude. Errors bars show standard errors. MA = melodic accent; TA = temporal accent; ST = semitone.

both ANOVAs. These comparisons revealed that the three largest MA magnitudes (3 ST, 4 ST, and 5 ST) and the three largest TA magnitudes (100 ms, 120 ms, and 140 ms) all significantly differentiated duple from triple meter (ps $<$ .01).

Absolute value of C scores. We also analyzed the data in a different manner, on the basis of the following logic. The C score scale treats duple and triple meter as polar opposites on a quasi-continuous variable. Such a scale, however, may not accurately reflect listeners' decision process. That is, rather than making a single decision (implied by the C scale), listeners may implicitly partition this task into two subtasks, the first emphasizing categorization ("Is the grouping pattern duple or triple?") and the second emphasizing perceptual clarity ("How clear is the grouping?"). Strong main effects for accent pattern (explaining 31% and 51% of the variance of ratings to MA and TA patterns, respectively) indicate that subjects were adept at the categorization portion of the decision. To isolate the metrical clarity portion, we took the absolute value of a signed clarity score (|C|). This reflects the fact that melodies rated as "C = –3" and "C = 3" reflect the same degree of metrical clarity (*very strong groups*) instantiated by different accent patterns. Thus, performance on the |C| scale isolates the clarity with which listeners perceived a meter (on a 4-point scale), from its category (duple, triple), with a higher score indicating greater clarity.

Two 2 (accent pattern) × 4 (accent magnitude) ANOVAs were conducted on |C| scores for MA melodies and TA melodies. Figure 4 presents the combined data as a function of these two factors for each accent type. A main effect for accent magnitude, with a strong linear (lin) trend component, emerged in both the MA (left panel) and TA (right panel) melodies, $F_{\text{lin}}(1, 27) = 36.92$, $\eta^2 = .58$, $p < .001$, and $F_{\text{lin}}(1, 27) = 103.55$, $\eta^2 = .79$, $p < .001$, respectively. Although an Accent Magnitude × Accent Pattern interaction was significant in both ANOVAs, in neither did it account for more than 3% of the variance; |C| scores increased with accent magnitude regardless of whether those accents appeared in duple or triple accent patterns.

One goal of Experiment 1 was to identify comparably salient accent magnitudes. Inspection of Figure 4 suggests that the 4- and 5-ST pitch-leap accent patterns (left panel) and the 100-ms duration accent pattern (right panel) elicited similar ratings of metrical clarity (|C| score means ~1.25) in both duple and triple accent patterns. We confirmed this by performing a 2 (accent type) × 2 (accent pattern) × 2 (accent magnitude) ANOVA and a Tukey's honestly significant difference (HSD) test on the three-way interaction. (It was necessary to include accent type as a factor, since

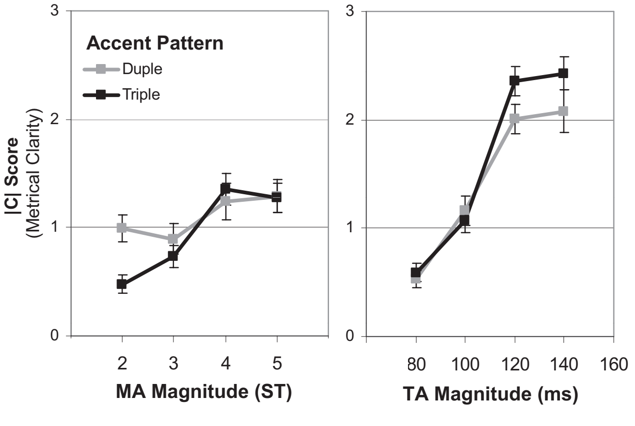

Figure 4. Experiment 1: Absolute value of a signed clarity score (|C| score) means as a function of accent type, accent pattern, and accent magnitude. Errors bars show standard errors. MA = melodic accent; TA = temporal accent.

we were interested in comparisons of all data points.) The |C| scores for the 4- and 5-ST MA and the 100-ms TA did not differ in either accent pattern ($ps > .8$).

We also conducted an experiment in which accent type was a between-subjects factor (versus as a within-subjects factor, reported here). The general pattern of findings was quite similar.

Discussion

Our main goal in Experiment 1 was to assess metrical salience as a function of accent type and magnitude. We found that magnitude increments of serial pitch and time changes increased perceived differences between duple and triple meter (C score analysis) and enhanced metrical clarity (|C| score analysis). These findings support the accent salience hypothesis.

On the whole, it could be argued that TA patterns conveyed more metrical information than MA patterns, as evidenced by the larger effect sizes for accent pattern (in C scores) and accent magnitude (in |C| scores) in TA melodies than MA melodies. Such findings would, on the surface, appear consistent with the TA bias hypothesis. We caution against such an interpretation of this "main effect" of accent type, however, because MA and TA magnitudes were chosen a priori; thus, the two magnitudes assigned to the same level in an ANOVA (e.g., 140 ms, 5 ST) cannot be directly compared. Instead, it is more appropriate to conclude that metrical clarity was higher for some TA magnitudes than for some MA magnitudes, and vice versa.

Finally, a related goal in Experiment 1 was to uncover MA and TA accents of comparable salience levels. On the basis of the analyses, we selected the 5-ST pitch-leap accent and the 100-ms duration accent as comparable in their ability to evoke meter.

Experiment 2: MAs and TAs Combined in Melodies

Experiment 2 was designed to compare predictions of the JAS and temporal bias hypotheses. These hypotheses feature different roles for MAs in meter. To assess them, we created melodies that contained both MA patterns and TA patterns in order to determine whether MAs had any impact on performance, and if so, whether that impact reflected an additive or interactive relationship with TAs. We developed a set of nine JASs by factorially crossing three MA patterns with three TA patterns. That is, in addition to the duple (D) and triple (T) accent patterns used in Experiment 1, a "neutral" (N) accent pattern was created, with accent locations on Tones 1, 7, and 13 (i.e., the accent locations common to both duple and triple accent patterns). This factorial approach is outlined in Table 2 in the shorthand notation introduced earlier (metrical period subscripted by accent type). In Table 2, the two concordant JASs ($\mathrm{D}_{\mathrm{MA}}\mathrm{D}_{\mathrm{TA}}$ and $\mathrm{T}_{\mathrm{MA}}\mathrm{T}_{\mathrm{TA}}$; cf. Figures 2A and 2B) are marked in bold, and the two discordant JASs ($\mathrm{T}_{\mathrm{MA}}\mathrm{D}_{\mathrm{TA}}$ and $\mathrm{D}_{\mathrm{MA}}\mathrm{T}_{\mathrm{TA}}$; cf. Figures 2C and 2D) are italicized. Table 2 also illustrates four "simple JASs"; these patterns contain either a duple or a triple accent pattern in one dimension and a neutral accent pattern in the other ($\mathrm{D}_{\mathrm{MA}}\mathrm{N}_{\mathrm{TA}}$, $\mathrm{T}_{\mathrm{MA}}\mathrm{N}_{\mathrm{TA}}$, $\mathrm{N}_{\mathrm{MA}}\mathrm{D}_{\mathrm{TA}}$, $\mathrm{N}_{\mathrm{MA}}\mathrm{T}_{\mathrm{TA}}$). Finally, a "neutral JAS" ($\mathrm{N}_{\mathrm{MA}}\mathrm{N}_{\mathrm{TA}}$) contained neutral accent patterns in both dimensions.

Table 2
Nine Joint Accent Structures Formed by Melodic Accent and Temporal Accent Patterns

	TA pattern		
MA pattern	Duple	Neutral	Triple
Triple	$\mathit{T}_{\mathit{MA}}\mathit{D}_{\mathit{TA}}$	$\mathrm{T}_{\mathrm{MA}}\mathrm{N}_{\mathrm{TA}}$	$\mathbf{T}_{\mathbf{MA}}\mathbf{T}_{\mathbf{TA}}$
Neutral	$\mathrm{N}_{\mathrm{MA}}\mathrm{D}_{\mathrm{TA}}$	$\mathrm{N}_{\mathrm{MA}}\mathrm{N}_{\mathrm{TA}}$	$\mathrm{N}_{\mathrm{MA}}\mathrm{T}_{\mathrm{TA}}$
Duple	$\mathbf{D}_{\mathbf{MA}}\mathbf{D}_{\mathrm{TA}}$	$\mathrm{D}_{\mathrm{MA}}\mathrm{N}_{\mathrm{TA}}$	$\mathit{D}_{\mathit{MA}}\mathit{T}_{\mathit{TA}}$

Note. Joint accent structure (JAS) patterns are labeled in shorthand with accent period (D = duple; N = neutral; T = triple) subscripted by the accent type instantiating it (MA = melodic accent; TA = temporal accent). Concordant JASs are marked in bold; discordant JASs are italicized.

The TA bias hypothesis leads to predictions about meter that contrast with those of both the JAS hypothesis and the accent salience hypothesis. First, the JAS hypothesis predicts that concordant melodies will elicit stronger (and faster) ratings of metrical clarity than discordant melodies, due to differences in temporal complexity of these joint accent patterns. In a C score analysis, this should result in more extreme ratings for concordant melodies (i.e., closer to either endpoint of the scale) and more neutral ratings for discordant melodies (i.e., closer toward $C = 0$). In a |C| score analysis, this should result in higher |C| values for concordant melodies and lower |C| values for discordant melodies. Similarly, response times (RTs) should be faster for concordant melodies than discordant melodies. The TA bias hypothesis, by contrast, does not predict that concordant JASs will facilitate meter perception; instead, C scores, |C| scores, and RTs should be similar for any two melodies with the same TA pattern regardless of the MA pattern (e.g., $\mathrm{D}_{\mathrm{MA}}\mathrm{D}_{\mathrm{TA}}$ and $\mathrm{T}_{\mathrm{MA}}\mathrm{D}_{\mathrm{TA}}$).

Second, the accent salience hypothesis predicts differences in performance between concordant JASs and simple JASs that have the same TA pattern (e.g., $\mathrm{D}_{\mathrm{MA}}\mathrm{D}_{\mathrm{TA}}$ vs. $\mathrm{N}_{\mathrm{MA}}\mathrm{D}_{\mathrm{TA}}$). According to the accent salience hypothesis, metrical clarity and response speed should increase with the number of co-occurring accent tokens. Thus, the presence of simultaneous MAs and TAs in concordant melodies (e.g., $\mathrm{D}_{\mathrm{MA}}\mathrm{D}_{\mathrm{TA}}$) should improve performance relative to melodies with TAs only (e.g., $\mathrm{N}_{\mathrm{MA}}\mathrm{D}_{\mathrm{TA}}$). The TA bias hypothesis does not predict this difference in performance: the MAs in concordant melodies are irrelevant and should not facilitate performance.

Method

Subjects. Twenty-five OSU undergraduates in psychology participated for course credit. They averaged 3.6 years of formal musical training ($SD = 2.6$; range = 0–9) and reported normal hearing. They were randomly assigned to one of four presentation orders ($ns = 6, 6, 6,$ and 7).

Apparatus. The apparatus was identical to that used in Experiment 1.

Stimuli and conditions. Twenty-seven melodies were created. All retained the same invariant IOI of 500 ms and pitch contour used in Experiment 1 melodies. However, in Experiment 2, each melody had both an embedded MA pattern and an embedded TA pattern (neutral, duple, or triple). The nine possible JASs are shown in Table 2. MA patterns were marked by 5-ST pitch-leap accents and TA patterns by 100-ms duration accents. Because a possible confounding of peak pitch with accent magnitude is not a factor in Experiment 2 melodies (due to a single pitch-leap magnitude), all melodies were anchored by their starting pitch (rather than by pitch peak as in Experiment 1). Equally often, the starting pitch was C4, D$^{\#}$4, or F$^{\#}$4.

Design. The primary design was 3 × 3 repeated-measures, with factors MA pattern (neutral, duple, triple), and TA pattern (neutral, duple, triple).

Procedure. The procedure was identical to that used in Experiment 1, with the following exceptions: (a) For each of the four presentation orders, the trial-to-trial constraints were that no three consecutive trials could possess the same starting pitch, the same MA pattern, or the same TA pattern; and (b) All melodies were heard three times over the course of the experiment.

Data reduction and analysis. Data from 5 subjects who failed (on the basis of their C scores) to follow instructions were excluded from further analysis.[2] Remaining data were analyzed as in previous experiments. RTs were recorded from the onset of the final tone of each melody; all melodies were exactly the same length. Because RT distributions are positively skewed, a few long response times can dramatically alter RT means (Ratcliff, 1993). To correct for this, we excluded RTs that were more than 2.5 standard deviations away from each subject's grand mean RT. On average, 7% of individual RTs were excluded for each subject.

Results

Signed C scores. We analyzed signed C scores using a 3 × 3 ANOVA, with factors MA Pattern (neutral, duple, triple) and TA Pattern (neutral, duple, triple). Figure 5 presents C score means as a function of these two factors. Significant effects were present for MA pattern, $F(2, 38) = 71.26$, $\eta^2 = .51$, $p < .001$; TA pattern, $F(2, 38) = 39.38$, $\eta^2 = .17$, $p < .001$; and their interaction, $F(4, 76) = 2.76$, $\eta^2 < .01$, $p < .05$. The MA factor accounted for three times more variance than the TA factor, suggesting that when both MA and TA patterns are present in the same melody, the TA pattern is not necessarily the stronger metrical marker. Compared with the main effects, the interaction explained a negligible amount of variance.

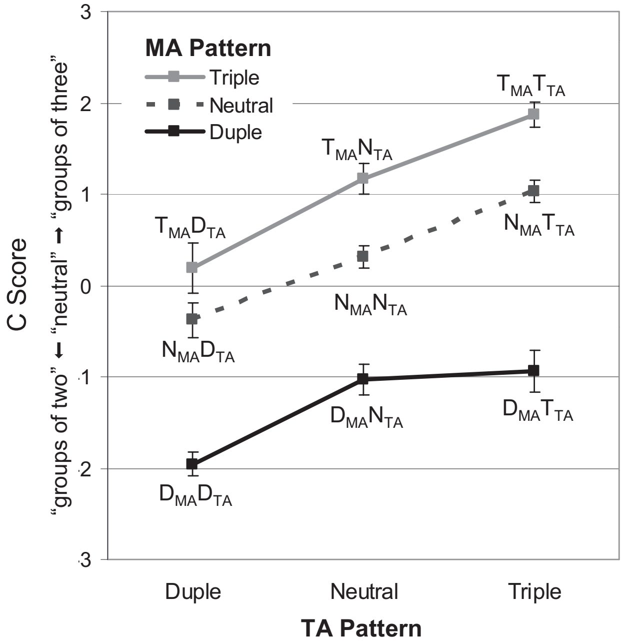

Figure 5. Experiment 2: Signed clarity score (C score) means as a function of melodic accent (MA) pattern and temporal accent (TA) pattern for the nine joint accent structures. Errors bars show standard errors. T = triple; N = neutral; D= duple.

In Figure 5, it is clear that concordant JASs ($T_{MA}T_{TA}$ and $D_{MA}D_{TA}$) received the most extreme C scores (1.87 and –1.96, respectively), whereas discordant JASs ($T_{MA}D_{TA}$ and $D_{MA}T_{TA}$) received less extreme scores (0.19 and −0.94, respectively). This evidence supports the JAS hypothesis but not the TA bias hypothesis.

Absolute value of C scores. The |C| scores isolate the clarity portion of a listener's decision from the categorization portion. According to predictions of the JAS hypothesis, concordant JAS melodies should elicit the percept of meter (D or T) more clearly (i.e., larger |C| values) than discordant JAS melodies. In addition, the accent salience hypothesis predicts that concordant JAS melodies should elicit clearer metrical percepts than comparable simple JAS melodies.

First, we submitted |C| scores to the same 3 × 3 ANOVA used earlier. Once again, MA pattern explained more variance than TA pattern, $F(2, 38) = 16.12$, $\eta^2 = .15$, $p < .001$; $F(2, 38) = 7.21$, $\eta^2 = .05$, $p < .005$, respectively. Compared with Experiment 1, however, the amount of explained variance by the two main effects was markedly less. Instead, the MA Pattern × TA Pattern interaction, $F(4, 76) = 19.73$, $\eta^2 = .16$, $p < .001$, accounted for substantially more variance in the |C| score ANOVA than the C score ANOVA (16% vs. < 1%).

Next, we directly tested a prediction of the JAS hypothesis that contrasts with that of the TA bias hypothesis by focusing upon the subset of concordant and discordant JASs. Within this subset, the JAS hypothesis predicts greater clarity (i.e., larger |C| values) for concordant JASs than for discordant JASs, whereas the TA bias hypothesis does not. That is, if the predicted dependency of melodic and temporal accent structures is present, it should emerge as an MA Pattern × TA Pattern interaction. Consistent with the JAS hypothesis, the 2 × 2 design revealed a main effect for neither factor, Fs < 1, η^2s < .01, but instead a pronounced MA Pattern × TA Pattern interaction, $F(1, 19) = 60.58$, $\eta^2 = .29$, $p < .001$. (This same 2 × 2 ANOVA on C scores, by contrast, resulted in only a slight interaction, with $\eta^2 = .01$.) As shown in the left panel

---

[2] Because data from this experiment tested several meter predictions of the JAS account, care was taken to eliminate subjects who incorrectly used the MIDILAB response box, under the following logic. Experiment 1 established that duple and triple accent patterns consistently elicited negative and positive C scores, respectively (or, at the very least, scores to triple-meter patterns that were more positive than scores to duple-meter patterns). We reasoned that a subjects' $T_{MA}T_{TA}$ scores should be more positive than $D_{MA}D_{TA}$ scores, thus $T_{MA}T_{TA} - D_{MA}D_{TA} > 0$. Because two adjacent response buttons are 1.0 C score units apart, the final inclusion criteria was restated as "include if $T_{MA}T_{TA} - D_{MA}D_{TA} \geq 1.0$." Five subjects failed to meet this criterion. In the original sample ($n = 25$), the mean value of $T_{MA}T_{TA} - D_{MA}D_{TA}$ was 2.73 ($SD = 2.03$). For the reduced sample ($n = 20$), the mean value of $T_{MA}T_{TA} - D_{MA}D_{TA}$ was 3.82 ($SD = 1.04$).

of Figure 6, concordant melodies received higher |C| scores than discordant melodies.

We also directly tested a different prediction, contrasting the accent salience and temporal bias hypotheses. This concerns a different subset of the nine JASs shown in Table 2: those melodies with identical accent periods in which MAs and TAs occur together and those in which only TAs appear. Specifically, $D_{MA}D_{TA}$ (MAs and TAs present) was contrasted with $N_{MA}D_{TA}$ (no MAs present); similarly, $T_{MA}T_{TA}$ was contrasted with $N_{MA}T_{TA}$. The accent salience hypothesis predicts greater metrical clarity for concordant JASs ($D_{MA}D_{TA}$ and $T_{MA}T_{TA}$) than for simple JASs ($N_{MA}D_{TA}$ and $N_{MA}T_{TA}$). By contrast, the TA bias hypothesis predicts no increased clarity for a TA pattern that is simultaneously marked by an MA pattern. A Tukey's HSD test on the MA Pattern × TA Pattern interaction in the 3 × 3 |C| score ANOVA revealed that concordant JAS melodies did indeed have higher |C| scores than simple JAS melodies ($ps < .001$). This finding is consistent with the accent salience hypothesis but inconsistent with the TA bias hypothesis.

Finally, there is the suggestion of a main effect for accent type in the right panel of Figure 6, indicating that the 5-ST MA may have lead to greater clarity than the 100-ms TA. However, a 2 (accent pattern) × 2 (accent type) ANOVA on the four simple JAS means revealed that the accent type effect was nonsignificant, $p > .05$, $\eta^2 < .06$.

Response times. The JAS hypothesis predicts faster RTs to concordant JAS melodies than to discordant JAS melodies. The accent salience hypothesis likewise predicts faster RTs to concordant JAS melodies than to simple JAS melodies.

We analyzed JAS predictions with respect to RT data using the same ANOVA designs that we used in analyzing |C| scores. Results are shown in the left panel of Figure 7. Mean RTs to melodies with concordant or discordant JASs are shown in the left panel of Figure 7. In the 2 × 2 ANOVA, neither MA pattern nor TA pattern main effects explained a meaningful proportion of the variance (both $\eta^2 < .01$). Rather, these factors interacted, $F(1, 19) = 15.35$, $\eta^2 = .06$, $p < .001$: concordant JAS melodies elicited faster RTs than discordant JAS melodies. (We also performed this same ANOVA on *z* scores for individual subject's response times. This boosted the η^2 value of the interaction from .06 to .15.)

Accent salience hypothesis predictions with respect to RT data were assessed via a Tukey's HSD test on the MA Pattern × TA Pattern interaction in the 3 × 3 ANOVA, $F(4, 76) = 3.49$, $\eta^2 = .03$, $p < .05$, and are shown in the right panel of Figure 7. The four simple JAS means ($N_{MA}D_{TA}$, $D_{MA}N_{TA}$, $N_{MA}T_{TA}$, and $T_{MA}N_{TA}$) did not differ from each other, nor did they differ from mean RTs to discordant JASs ($ps > .9$). However, all were significantly slower than the mean RT to $D_{MA}D_{TA}$ ($ps < .05$). Together, these findings are in greater agreement with the accent salience hypothesis than with the TA bias hypothesis.

Discussion

Four major findings appeared in Experiment 2. First, strong evidence emerged that MAs can contribute to meter perception. In fact, in the present case, the MA pattern (established by 5-ST pitch-leap accents) accounted for three times more variance in subjects' ratings than the TA pattern (established by 100-ms duration accents) in both C score and |C| score ANOVAs. This result is markedly different from previous findings (e.g., Hannon et al., 2004; Huron & Royal, 1996; Monahan & Carterette, 1985).

Second, we found clear support for the JAS hypothesis, which predicts that metrical clarity should be greater for melodies with a concordant JAS than for melodies with a discordant JAS. Concordant JASs had more extreme values in the C score analysis and higher values in the |C| score analysis than discordant JASs and simple JASs.

Third, we found the first clear evidence supporting a prediction of the accent salience hypothesis: that the number of co-occurring accents influences accent salience. Metrical clarity increased and RTs decreased as the number of coincidental accents increased from one (MA or TA patterns alone in simple JAS melodies) to two (MAs plus TAs in concordant melodies).

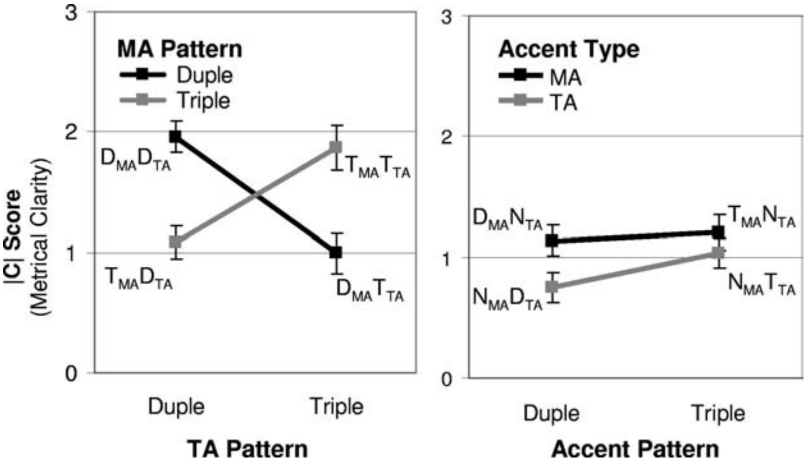

Figure 6. Experiment 2: Absolute value of a signed clarity score (|C| score) means for concordant and discordant joint accent structures (JASs) (left panel) and simple JASs (right panel). Errors bars show standard errors. $N_{MA}N_{TA}$ melodies had a mean |C| score of .44 (*SE* = .10). MA = melodic accent; TA = temporal accent; T = triple; N = neutral; D= duple.

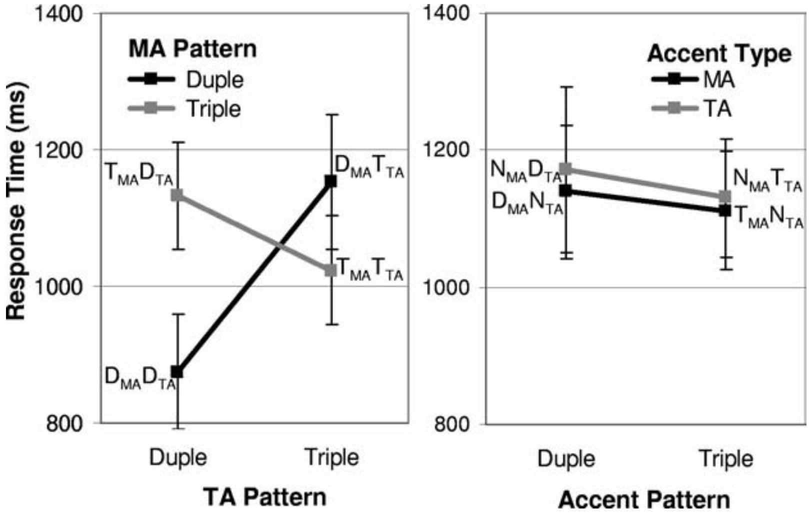

Figure 7. Experiment 2: Response time means to concordant and discordant joint accent structures (JASs; left panel) and simple JASs (right panel). Errors bars show standard errors. MA = melodic accent; TA = temporal accent; T = triple; N = neutral; D= duple.

Fourth, analyses of signed C scores versus |C| scores revealed that measurement scales can change the apparent relationship of MA and TA factors from additive to interactive, despite the same stimuli and factorial design. We suggested that the C scale reflects both a categorization decision (duple vs. triple) by its sign and a clarity decision (slightly, moderately, or very clear groups) by its value, whereas the |C| scale isolates the clarity decision. This finding underscores the need to assess the link between task and scoring to ensure that they reasonably reflect the phenomena of interest—in this case, categorization and clarity. Here, the apparent relation between MA and TA information in ratings to concordant versus discordant melodies changed as function of whether we were scoring categorization using C or metrical clarity using |C|. Analyses for the former indicate additivity (η^2 for interaction = .01), whereas those for the latter confirm interactivity (η^2 for interaction = .29). This implies that differences in scoring, which often differ as a function of task, may lead to markedly different conclusions about the apparent nature of MA and TA contributions to meter perception.

General Discussion

We assessed predictions of three hypotheses about the salience and timing of different accents in music-like events. Two hypotheses, the accent salience hypothesis and the JAS hypothesis, derive from a dynamic attending framework (Jones, 1987, 1993; Jones & Boltz, 1989). A third, the TA bias hypothesis, reflects a popular alternative view, with contrasting predictions on these issues. We first discuss implications of our findings regarding these hypotheses. We then consider the issues of independence of melody and rhythm, and cortical localization of melody and rhythm.

The Accent Salience Hypothesis

The accent salience hypothesis formulated here expands an earlier position outlined by Jones (Jones, 1987, 1993). It posits that there are three contributors to accent salience: (a) the effective magnitude of a local serial change, (b) the number of simultaneous changes (type or tokens) at a particular serial location, and (c) the global serial context. In Experiment 1, global context was controlled in order to evaluate the effects of serial change magnitude on the salience of pitch-leap accents and duration accents. This hypothesis received support, in the form of a strong linear relation between accent magnitude and perceived metrical clarity. That metrical salience increases with accent magnitude has been previously noted by Windsor (1993) in an examination of the categorical perception meter instantiated by intensity accents. The present set of experiments, however, is the first in which perceived meter has been systematically investigated with regard to the salience of pitch-leap and duration accents.

Experiment 1 also permitted calibration of the relative salience of different accent types (MA and TA), the logic being that some magnitude of a pitch-leap accent and some magnitude of a duration accent should elicit meter to a comparable degree. We noted that the practice of finding comparable accents before investigating their relative contributions to meter is largely absent from previous investigations. If accents are not comparable (i.e., if TA patterns are salient but MA patterns are not), then risks of biased or misleading conclusions are considerably increased.

In Experiment 2, we used some melodies that included both MAs and TAs and others that contained only one of these two accent types. This experiment provided the first evidence favoring another accent salience prediction, which holds that increasing the number of simultaneous changes (type or tokens) at particular serial locations increases overall accent salience. In this case, coincidental melodic-plus-temporal accents in concordant JAS melodies elicited greater metrical clarity and speedier responding than single accent patterns (either MA or TA), even when the underlying accent pattern was the same. This finding is not consistent with the TA bias hypothesis, which predicts no difference in performance as function of adding an MA accent to a TA accent pattern.

Experiment 2 also provided a surprising twist on traditional views of accent salience. In this experiment, individual accent magnitudes were fixed at salient levels (5-ST pitch-leap, 100-ms duration accents, from Experiment 1). Despite indications in Experiment 1 that these two accent tokens were of comparable salience, when they were combined in the same melodies, the MA factor explained three times more variance in listeners' meter ratings than did the TA factor. Such a finding casts doubt on the claim that temporal accents carry more weight during meter perception, limiting the generality of the TA bias hypothesis. In fact, the outcome of Experiment 2 tempts the formulation of an "MA bias hypothesis." We caution against either bias hypothesis, however. Instead, it is more plausible to assume that the relative contributions of different accent types depend not only on their magnitudes but also on the surrounding serial context (discussed later). In bias hypotheses, there is an incorrect assumption that a particular type of accent is inherently better suited to mark time, regardless of local accent magnitude or global serial context.

We note a parallel between our findings and those of Melara and Marks (1993), who manipulated the salience of word and color dimensions in a Stroop paradigm (Stroop, 1935; cf. MacLeod, 1991). When Melara and Marks made words less discriminable than colors (by reducing their font size), the classic Stroop effect disappeared: Word variation did not hinder performance in classifying colors. In fact, when colors were made more discriminable than words, a reverse Stroop effect was found: Irrelevant variation in color hindered performance in classifying words. The authors concluded that Stroop dimensions are not special or intrinsic in their effects. Rather, these effects vary as a function of the relative salience of stimulus dimensions.

In sum, our findings underscore the importance of calibrating accent salience in various experimental contexts. Alternative hypotheses endorsing biases toward one accent type versus another are misleading if not qualified by considerations of local change magnitude, accent co-occurrence, and global pattern context.

Caveats on Generality

Our conclusions regarding the contributions of melodic and temporal accents to meter perception are, of course, based on a rather tightly controlled stimulus set: 13-tone sequences in which global melodic contour structure and tempo (IOI) were matched and in which accent patterns were established by a single MA token (pitch leaps ranging from 2–5 ST) and/or a single TA token (duration accents ranging from 80–140 ms). We chose these accent tokens specifically because they allowed us to preserve global pattern structure and because it was easy to systematically manipulate their respective magnitudes. This may constrain the generality of our findings and raise questions about the extent to which they apply to more natural examples of music based on artistic (rather than experimental) goals. Composers vary the magnitudes and co-occurrences of accents to meet artistic objectives; according to the present account, these accents will bias listeners in various ways. However, the aim of the present research was not to assess compositional style but to ascertain, in a controlled experiment, whether average listeners can use certain MA tokens, singly or in combination, to infer meter. It appears they can.

Our definition of MAs is broader than pitch leaps, and for completeness, other accent types and tokens must also be explored. Such investigations may prove daunting, because how a melodic accent is defined is subjective, depending largely upon who does the defining (Huron & Royal, 1996). For example, in addition to pitch-leap accents and contour accents, music theorists have also proposed treble accents, bass accents, and registral extreme accents.

Also daunting is assessing the role of context variability in musical events. One consideration is that in Western music, a majority of pitch intervals tend to be relatively small (i.e., < 3 ST; Dowling & Harwood, 1986), suggesting low levels of overall melodic variability; in this case, sampling real musical excerpts would likely result in auditory patterns that contain few salient MAs (i.e., pitch leaps of 4 ST of more). Perhaps, then, the tendency of naturalistic research to find support for a TA bias hypothesis rests on MAs of salient magnitudes being less available in the corpus (i.e., Hannon et al., 2004). But another interpretation of Western music may also apply. Over the course of a composition, the range and variability of accumulated melodic changes may be greater than those of rhythmic ones, leading to a global serial melodic context that attenuates the salience of all but relatively large changes in pitch. If so, this too may contribute to ways in which a musical genre can, by virtue of introducing more or less global variability along one or another dimension, either increase or decrease the salience of certain accent types. Indirect evidence for this effect of global context is found in the present research if we compare the salience of MAs in Experiments 1 and 2. In Experiment 1, the MA melodies necessarily resulted in more global melodic variability than in Experiment 2. In Experiment 1, MAs varied between 2 ST and 5 ST over the course of session with corresponding variations in melodic structure; by contrast, in Experiment 2, the MAs did not vary over a session but remained fixed at 5 ST. These manipulations were necessary consequences of using pitch-leap accents; however, the results allow us to suggest why MAs may have become more salient in the global context of Experiment 2 than in the global context of Experiment 1.

In sum, although our stimuli have been constrained to permit causal inferences, these findings also offer the advantage of such experimental controls. We can safely conclude that neither MAs nor TAs (as manipulated here) are inherently better at communicating metrical information. Rather, if two accent tokens have comparable salience (within comparable global contexts), then even in more naturalistic environments they should function according to principles outlined by the JAS approach.

The JAS Hypothesis

The JAS hypothesis (e.g., Jones, 1987, 1993) addresses time patterns that are created by a variety of accent types within a given melody. We distinguish it from the accent salience hypothesis, which focuses upon the nature and strength of individual time markers (i.e., accents) rather than upon their relative timing. Assuming that multiple accent types or tokens are present and salient, the JAS hypothesis predicts that metrical clarity is affected by the temporal complexity of an emergent higher order time structure (i.e., the JAS). Concordant JASs have low temporal complexity (i.e., greater temporal regularity of accent periods, indexed by an integer accent period quotient) and should promote efficient entrainment of attending, thereby facilitating meter perception. Discordant JASs have greater temporal complexity (a noninteger accent period quotient) and hence should hinder meter perception.

Consistent with the JAS hypothesis, in Experiment 2, we found that concordant JASs elicited clearer and more rapid meter judgments than discordant JASs. The factorial design of Experiment 2 enabled us to assess the relative contributions of melodic and temporal accent information to meter perception. In one analysis (C scores), melodic and temporal factors were found to combine additively, while in another analysis (|C| scores), the two factors interacted. Despite this difference, these two analyses point toward a common finding: Meter is clearer for melodies with concordant JASs than those with discordant JASs. C scores were more extreme for concordant JASs than discordant JASs (which produced the two main effects in the C score ANOVA), and |C| scores were higher for concordant JASs than discordant JASs (which caused the interaction effect in the |C| score ANOVA).

MA and TA Information: Additive or Interactive?

Schellenberg et al. (2000, p. 156) note that whether the relationship between MA and TA information is additive or interactive is of “central” importance in understanding both lower and higher level music perception. As such, the issue has been explored across a variety of tasks, including ratings of melodic phrase completion (Palmer & Krumhansl, 1987a, 1987b) and phrase endings (Boltz, 1989a, 1989b), melodic similarity (Monahan & Carterette, 1985), melody duration (Boltz, 1998), melodic recall (Boltz, 1991; Boltz & Jones, 1986; Deutsch, 1980; Drake et al., 1991), detection of pitch changes (Dowling, Lung, & Herrbold, 1987; Jones et al., 1982; Monahan et al., 1987), and tapping synchronization (Jones & Pfordresher, 1997; Pfordresher, 2003; Snyder & Krumhansl, 2001). With such a large pool of data, the consensus might be summarized as follows: Melodic and temporal information combines additively in some tasks and interactively in others (cf. Krumhansl, 2000; Krumhansl & Iverson, 1992; Schellenberg et al., 2000).

The present data have three implications for the additive versus interactive debate. The first concerns the requirement of salient accents. With regards to meter in particular, evidence from prior research, which has been interpreted as favoring an additive relationship, often draws on the absence of salient MA information relative to TA information. Thus, although Garner (1970, 1974) famously distinguished between separable (i.e., additive) and integral (i.e., interactive) dimensions in perception, he noted that notions of additivity versus interactivity must be tempered by dimensional discriminability or salience (Garner & Felfoldy, 1970). If listeners are not sensitive to MAs, then discussions of whether MAs and TAs combine additively or interactively in perception are groundless.

The second implication concerns the importance of data scoring in circumscribing a pattern of results as additive or interactive. We have demonstrated that a strong additive relationship between MA and TA factors can be transformed into a strong interactive relationship between MA and TA factors via an absolute value transformation of our C score data. We have interpreted the transformed score, |C|, as isolating performance appropriate to a task in which subjects simply rate metrical clarity while ignoring metrical category. Although it has been noted previously that task differences may change the apparent relationship of MA and TA information (Krumhansl, 2000; Krumhansl & Iverson, 1992; Schellenberg et al., 2000), we believe that our analysis is the first to recover both relationships from the same data set.

The third implication concerns interpretations of statistical findings. We reported a strong statistical interaction in our |C| score analysis of concordant versus discordant JASs. Alternatively, we could have simply reported a main effect for JAS concordance: Concordant JASs have higher |C| scores than discordant JASs overall (cf. Figure 6). The constructs of *concordant* and *discordant* in our stimuli imply a de facto interdependency of MA and TA information (vis-à-vis the accent period quotient). In other words, the presence of a *statistical* interaction should be secondary to the *conceptual* interpretation of the data. A main effect for JAS concordance thus reflects a melodic–temporal interdependency that holds regardless of how data are parsed during analysis. In sum, the absence of a statistical interaction does not necessarily indicate the absence of a dependent relationship between MA and TA contributions to perception.

An Implication for Neuropsychological Investigations of Listener Deficits

A number of reports in the neuropsychological literature suggest that melodic and temporal information may be processed independently in the brain (for reviews, see Peretz & Morais, 1989; Peretz & Zatorre, 2005; Vignolo, 2003; Zatorre, Belin, & Penhune, 2002). Patients with left hemisphere (LH) damage to the auditory cortex exhibit deficits in rhythmic grouping but spared sensitivity to changes in pitch structure (e.g., pitch contour). Conversely, patients with damage to right hemisphere (RH) auditory cortex often exhibit a reverse pattern of sensitivities, showing loss of sensitivity to pitch contour but a spared sensitivity to rhythm. In addition, there is evidence of a dissociation of meter and rhythm, such that damage to the RH may interfere with listeners’ ability to produce metrical regularities (meter) but not with their sensitivity to temporal grouping (rhythm; e.g., Penhune, Zatorre, & Feindel, 1999; Peretz & Zatorre, 2005). Such findings have led to the development of a modular model of auditory processing (Peretz & Coltheart, 2003, their Figure 1) that separates the processing of pitch and time relations into parallel and independent subsystems. A pitch organization subsystem (comprising pitch interval and contour analyses) contains no architectural links to a meter analysis module, which is exclusively encapsulated within a temporal organization subsystem.

Additive relationships between pitch (MA) and time (TA) structures are considered consistent with modular accounts. Both imply independence of these dimensions. Thus, in the experiments we have reported here, a modular account leads to an expectation that JAS melodies with a salient MA pattern (e.g., duple vs. triple) should be processed in a pitch organization module and that a salient TA pattern should be processed independently in a separate meter module. However, logically it remains unclear how the MA pattern, if processed in the pitch organization module, might lead to percepts of duple or triple meter. Conversely, if the MA pattern is processed in a meter module, then it is unclear how pitch changes inherent in the melodic pattern can be effectively processed in this module. Indeed, our data do not support independence of pitch (melody) and time (meter) relations in that they show that listeners do rely on melodic patterns when perceiving meter. This presents a problem in applying the modularity approach to normal listeners.

We have already suggested that issues related to accent magnitude or calibration and to global pattern variability are at the heart of controversies over the role of MAs in meter perception in behavioral investigations. Similar points can be raised with regard to the impact of stimulus–design factors on neuropsychological findings. The neurological assessments of listeners cited earlier rest on diagnostic tests of meter that lack both (a) magnitude calibrations of MA and TA salience in test stimuli (as in Experiment 1) and (b) factorial designs that orthogonally cross melodic and rhythmic variables (as in Experiment 2). These stimulus–design limitations weaken causal inferences about the direct impact of pitch structure on a patient's perception of meter (as well as rhythm). Our findings reported here indicate that comparable accent salience must be established before relative accent contributions to meter can be determined.

Nevertheless, it is quite possible that patients with either RH or LH damage would not differentiate concordant from discordant JASs in the same manner as our listeners. The RH-damaged patients, being insensitive to pitch relationships, may well conform to predictions of a TA bias hypothesis; similarly, LH-damaged patients, being insensitive to time relationships, might perform according to a MA bias hypothesis. Although the latter outcome could be interpreted to mean that melodic and meter processing are both localized in the RH, it would not support a modularity claim that meter perception is independent of pitch structure, as suggested by Peretz and Coltheart (2003). That is, if LH patients successfully differentiated duple from triple meter in patterns containing salient, well-calibrated MAs, then meter perception would be dependent on melodic structure. A modular view seems to predict that neither healthy nor brain-damaged listeners can infer meter from melodic structure. We have shown that healthy listeners do not conform to this prediction; whether brain-damaged listeners do remains to be determined.

In summary, the present research suggests that for both healthy and brain-damaged listeners, issues of accent salience, contextual change and variability, and the factorial control of accent timing are important. Just as with healthy listeners, calibration of accent salience is important with brain-damaged listeners. Furthermore, with healthy listeners, we saw that when MAs have low salience in certain musical events, it becomes logically impossible to assess whether MAs interact with TAs. A parallel situation holds for brain-damaged listeners. If these listeners perceive either MAs or TAs as having greatly reduced salience, then it is logically impossible to evaluate whether MAs and TAs interact. In both instances, it would be impossible to fully test a modularity hypothesis of independence.

References

Benjamin, W. E. (1984). A theory of musical meter. *Music Perception, 1,* 355–413.

Bharucha, J. J. (1984). Anchoring effects in music: The resolution of dissonance. *Cognitive Psychology, 16,* 485–518.

Boltz, M. G. (1989a). Perceiving the end: Effects of tonal relationships on melodic completion. *Journal of Experimental Psychology: Human Perception and Performance, 15,* 749–761.

Boltz, M. G. (1989b). Rhythm and "good endings": Effects of temporal structure on tonality judgments. *Perception & Psychophysics, 46,* 9–17.

Boltz, M. G. (1989c). Time judgments of musical endings: Effects of expectancies on the "filled interval effect." *Perception & Psychophysics, 46,* 409–418.

Boltz, M. G. (1991). Some structural determinants of melody recall. *Memory & Cognition, 19,* 239–251.

Boltz, M. G. (1998). The processing of temporal and nontemporal information in the remembering of event durations and musical structure. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 1087–1104.

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

Borg, G. A. (1982). A category scale with ratio properties for intermodal and interindividual comparisons. In H.-G. Geissler & P. Petzold (Eds.), *Psychophysical judgment and the process of perception* (pp. 25–34). Berlin: Deutscher Verlag der Wissenschaften.

Castellano, M. A., Bharucha, J. J., & Krumhansl, C. L. (1984). Tonal hierarchies in the music of North India. *Journal of Experimental Psychology: General, 113,* 394–412.

Clarke, E. F. (1999). Rhythm and timing in music. In Deutsch, D. (Ed.), *The psychology of music* (2nd ed., pp. 473–500). New York: Academic.

Cooper, G. W., & 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 interference of metrical structure and perception of rhythm. *Perception & Psychophysics, 54,* 794–807.

Desain, P., & Honing, H. (1989). The formation of rhythmic categories and metrical priming. *Perception, 32,* 341–365.

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

Dixon, S., & Cambouropoulos, E. (2000). Beat tracking with musical knowledge. In W. Horn (Ed.), *Proceedings of the 14th European Conference on Artificial Intelligence*: *ECAI 2000*. (pp. 626–630). Amsterdam: IOS Press.

Dowling, W. J., & Harwood, D. L. (1986). *Music cognition.* San Diego: Academic Press.

Dowling, W. J., Lung, K. M.-T., & Herrbold, S. (1987). Aiming attention in pitch and time in the perception of interleaved melodies. *Perception & Psychophysics, 41,* 642–656.

Drake, C., & Bertrand, D. (2001). The quest for universals in temporal processing in music. *Annals of the New York Academy of Sciences, 930,* 17–27.

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.

Esposito, J. L. (1998). Type and token. In P. Bouissac (Ed.), *Encyclopedia of semiotics* (p. 622). New York: Oxford University Press.

Garner, W. R. (1970). The stimulus in information processing. *American Psychologist, 25,* 350.

Garner, W. R. (1974). *The processing of information and structure.* New York: Wiley.

Garner, W. R., & Felfoldy, G. L. (1970). Integrality of stimulus dimensions in various types of information processing. *Cognitive Psychology, 1,* 225–241.

Handel, S. (1989). *Listening: An introduction to the perception of auditory events* (pp. 383–459). Cambridge, MA: MIT Press.

Hannon, E. E., & Johnson, S. P. (2005). Infants use meter to categorize rhythms and melodies: Implications for musical structure learning. *Cognitive Psychology, 50,* 354–377.

Hannon, E. E., Snyder, J. S., Eerola, T., & Krumhansl, C. L. (2004). The role of melodic and temporal cues in perceiving musical meter. *Journal* *of Experimental Psychology: Human Perception and Performance, 30,* 956–974.

Hannon, E. E., & Trehub, S. E. (2005). Metrical categories in infancy and adulthood. *Psychological Science, 16,* 48–55.

Hannon, E. E., & Trehub, S. E. (2006). Tuning in to musical rhythms: Infants learn more readily than adults. *Proceedings of the National Academy of Sciences, 102,* 12639–12643.

Huron, D. (2006). *Sweet anticipation: Music and the psychology of expectation.* Cambridge, MA: MIT Press.

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

Johnson-Laird, P. N. (1991). Rhythm and meter: A theory at the computational level. *Psychomusicology, 10,* 88–106.

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. (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 and W. J. Dowling (Eds.), *Psychology and music: The understanding of melody and rhythm* (pp. 67-–92). Hillsdale, NJ: Erlbaum.

Jones, M. R. (2001). Temporal expectancies, capture, and timing in auditory sequences. In C. Folk & B. Gibson (Eds.), *Attraction, distraction, and action: Multiple perspectives on attentional capture* (pp. 191–229). New York: Elsevier.

Jones, M. R. (2004). Attention and timing. In J. Neuhoff (Ed.), *Ecological psychoacoustics* (pp. 49–85). New York: Academic Press.

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

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

Jones, M. R., Boltz, M. G., & Klein, J. M. (1993). Expected endings and judged duration. *Memory & Cognition, 21,* 646–665.

Jones, M. R., Johnston, H. M., & Puente, J. (2006). Effects of auditory pattern structure on anticipatory and reactive attending. *Cognitive Psychology, 53,* 59–96.

Jones, M. R., Moynihan, H., MacKenzie, N., & Puente, J. (2002). Temporal aspects of stimulus-driven attending in dynamic arrays. *Psychological Science, 13*, 313-–319.

Jones, M. R., & Pfordresher, P. Q. (1997). Tracking musical patterns using joint accent structure. *Canadian Journal of Experimental Psychology, 51,* 271–290.

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.

Justus, T. C., & Bharucha, J. J. (2002). Music perception and cognition. In H. Pashler (Series Ed.) & S. Yantis (Vol. Ed.)*, Stevens' handbook of experimental psychology* (3rd ed., Vol. 1, pp. 453–492). New York: Wiley.

Keller, P. E., & Burnham, K. (2005). Musical meter in attention to multipart rhythm. *Music Perception, 22,* 629–661.

Keppel, G., & Wickens, T. D. (2004). *Design and analysis: A researcher's handbook* (4th ed.). Upper Saddle River, NJ: Pearson Prentice Hall.

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

Krumhansl, C. L., & Iverson, P. (1992). Perceptual interactions between musical pitch and timbre. *Journal of Experimental Psychology: Human Perception and Performance, 18,* 739–751.

Laden, B. (1994). Melodic anchoring and tone duration. *Music Perception, 12,* 199–212.

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

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

Levelt, W. J., 1989. *Speaking: From intention to articulation (*pp. 83–86)*.* Cambridge, MA: MIT Press.

London, J. (2001). Rhythm. In S. Sadie & J. Tyrrell (Eds.), *The new Grove dictionary of music and musicians* (Vol. 21, pp. 277–309). London: Macmillan.

Longuet-Higgins, H. C., & Lee, C. S. (1982). The perception of musical rhythms. *Perception, 11,* 115–128.

MacLeod, C. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin, 109,* 163–203.

Marks, L. E., & Gescheider, G. A. (2002). Psychophysical scaling. In H. Pashler (Series Ed.) & S. Yantis (Vol. Ed.)*, Stevens' Handbook of Experimental Psychology* (3rd ed., Vol. 4, pp. 91–138). New York: Wiley.

McAuley, J. D., & Jones, M. R. (2003). Modeling effects of rhythmic context on perceived duration: A comparison of interval and entrainment approaches to short-interval timing. *Journal of Experimental Psychology: Human Perception and Performance, 29,* 1102–1125.

Melara, R., & Mounts, J. (1993). Selective attention to Stroop dimensions: Effects of baseline discriminability, response mode, and practice. *Memory & Cognition, 21,* 627–645.

Melara, R., & Mounts, J. (1994). Contextual influences on interactive processing: Effects of discriminability, quantity, and uncertainty. *Perception & Psychophysics, 56,* 73–90.

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, 41,* 576–600.

Narmour, E. (1996). Analyzing form and measuring perceptual content in Mozart's Sonata K. 282: A new theory of parametric analogues. *Music Perception, 13,* 728–741.

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.

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

Penhune, V. B., Zatorre, R. J., & Feindel, W. (1999). The role of auditory cortex in retention of rhythmic patterns in patients with temporal-lobe removals including Heschl's gyrus. *Neuropsychologia, 37,* 315–331.

Peretz, I., & Coltheart, M. (2003). Modularity of music processing. *Nature Neuroscience, 6,* 688–691.

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

Peretz, I., & Zatorre, R. J. (2005). Brain organization for music processing. *Annual Review of Psychology, 56,* 89–114.

Pfordresher, P. Q. (2003). The role of melodic and rhythmic accents in musical structure. *Music Perception, 20,* 431–464.

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.

Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin, 114,* 510–532.

Schellenberg, E. G., Krysciak, A. M., & Campbell, R. J. (2000). Perceiving emotion in melody: Interactive effects of pitch and rhythm. *Music Perception, 18,* 155–171.

Smith, K. C., & Cuddy, L. L. (1989). Effects of metrical 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,* 455–489.

Stevens, J. C., & Marks, L. E. (1965). Cross-modality matching of brightness and loudness. *Proceedings of the National Academy of Sciences, 54,* 407–411.

Stevens, J. C., & Marks, L. E. (1980). Cross-modality matching functions generated by magnitude estimation. *Perception & Psychophysics, 27,* 379–389.

Stevens, S. S. (1959). Cross-modality validation of subjective scales for loudness, vibration, and electric shock. *Journal of Experimental Psychology, 57,* 201–209.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18,* 643–662.

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

Tekman, H. G. (1998). Effects of melodic accents on perception of intensity. *Music Perception, 15,* 391–401.

Tekman, H. G. (2001). Accenting and detection of timing variations in tone sequences: Different kinds of accents have different effects. *Perception & Psychophysics, 63,* 514–523.

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,* 117–149.

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

Todd, R. E., Boltz, M. G., & Jones, M. R. (1989). The MIDILAB auditory research system. *Psychomusicology, 8,* 83–96.

Toiviainen, P., & Snyder, J. S. (2003). Tapping to Bach: Resonance-based modeling of pulse. *Music Perception, 21,* 43–80.

Vignolo, L. A. (2003). Music agnosia and auditory agnoisa. *Annals of the New York Academy of Sciences, 999,* 50–57.

Windsor, W. L. (1993). Dynamic accents and the categorical perception of metre. *Psychology of Music, 21,* 127–140.

Woodrow, H. (1911). The role of pitch in rhythm. *Psychological Review, 18,* 54–77.

Woodrow, H. (1951). Time perception. In S. S. Stevens (Ed.), *Handbook of experimental psychology* (pp. 1224–1236). New York: Wiley.

Yeston, M. (1976). *The stratification of musical rhythm.* New Haven, CT: Yale University Press.

Zatorre, R. J., Belin, P., & Penhune, V. B. (2002). Structure and function of auditory cortex: Music and speech. *Trends in Cognitive Sciences, 6,* 37–46.

Received November 15, 2006
Revision received May 14, 2008
Accepted June 26, 2008 ■