Music Perception
Summer 2002, Vol. 19, No. 4, 529–550

Cognitive Constraints on Metric Systems: Some Observations and Hypotheses

JUSTIN LONDON

This paper is a music-theoretic discussion of various studies on rhythmic perception and performance and their ramifications for discussions of musical meter. Meter is defined as a stable and recurring pattern of hierarchically structured temporal expectations. Metrical patterns, although related to the pattern of interonset intervals present in the musical surface, are distinct from that pattern. Studies of subjective rhythmization, spontaneous tempo, pulse perception, durational discrimination, and so forth are discussed with respect to their implications for meter. Not only do there seem to be upper and lower bounds for musical meter (from ≅100 ms to ≅6 s, depending on context), but there also appear to be important thresholds within this range (around 200–250 ms, 500–700 ms, and 1.5–2.0 s). Interactions between beats (i.e., interonset intervals between expectancies occurring at the rate perceived as the tactus), beat subdivision, and changes in tempo are discussed, and it is hypothesized that beat perception may require (at least potentially) the perception of a concomitant level of subdivision. The interactions between beat interonset interval, subdivision interonset interval, and various thresholds may also explain (in part) some of the differences in the expressive and/or motional character of rhythmic figures (duplets versus triplets) at different tempos. Last, a broader discussion of systematic relationships in larger metrical systems with respect to tempo is given. It is shown that the choice of tempo systematically constrains the number and kind of metric patterns that are available to the listener.

Received July 5, 2000, accepted August 21, 2001.

SINCE the nineteenth century, researchers have gathered a great deal of information regarding the temporal ranges in which we can hear a series of events as a rhythm, that is, as a coordinated and connected temporal pattern. The ways we can perceive, perform, and anticipate an extended series of more or less regularly occurring events within that range have also been studied. The ranges and phenomena investigated include the following.[1]

1. Summaries of much of this research may be found in Royal (1995), Palmer (1997), van Noorden and Moelants (1999), Clarke (1999), and Krumhansl (2000).

Address correspondence to Justin London, Department of Music, Carleton College, Northfield, MN 55057. (e-mail: jlondon@carleton.edu)

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

- The range of *subjective rhythmization*: the longest and shortest interonset periods for continuous, isochronous (and otherwise undifferentiated) stimuli that we tend to group into twos or threes;
- The range of *spontaneous tempo*: the longest and shortest periods in which we are able produce a steady beat;
- The values for *preferred tempos*: the rate at which we are most comfortable at producing a steady beat;
- The range in which we are most likely to hear a *pulse* or *tactus;*
- The *indifference interval*: a period that we tend to judge as neither too long or too short;
- Our sensitivity to *changes of tempo* at different initial rates and in different contexts;
- Our sensitivity to *differences of duration* relative to the magnitude of the durations involved;
- The shortest and longest durations musicians tend to produce in rhythmically palpable patterns (i.e., apart from trills, violin tremolos, vibratos, etc.);
- The extent and limits of the *psychological present.*

These investigations are particularly important for theories of musical meter. They strongly suggest that there are perceptual and cognitive limits on the temporal range of musical meter and, perhaps more significantly, that there may be important differences between our grasp and performance of certain rhythms within those limits. Likewise, an understanding of the metrical implications of various experimental stimuli and studies will engender better analysis and interpretation of experimental results.

Before going further, a caveat is required. It is probably impossible to come up with robust, absolute values for perceptual and cognitive limits for musical meter, and likewise it is impossible to determine hard and fast values for the various temporal thresholds that are described and discussed here. The values reported are largely from experiments that used nonmusical (or perhaps "quasi-musical") stimuli and contexts; for the most part, this research lacks ecological validity relative to real-life listening situations. Moreover, this research has shown that various thresholds, acuities, and so on are heavily dependent upon task, stimulus, context, and so forth. For example, research has convincingly shown that the perception of duration and accent interacts with pitch (Hirsh, Monohan, Grant, & Singh, 1990; Jones, Jagacinski, Yee, Floyd, & Klapp, 1995; van Noorden, 1975). And of course, there are often significant differences with respect to these various thresholds from subject to subject. Nonetheless, it seems reasonable to presume that in any given context the absolute temporal value of the component elements in a rhythmic pattern (and of overall pattern itself) will serve as a constraint on metric possibilities, and such constraints can affect our subjective sense of its movement and shape. More precisely, once a periodicity at a given temporal interval has been established in both the musical patterning and in the mind of the listener, one can then consider the implications for related periodicities, which is to say how that periodicity may fit into a more extended metric framework (e.g., Is the given periodicity apt to be heard as a tactus? Can it be subdivided into smaller periodicities, and if so, which ones?). Thus with the understanding that proposed thresholds may be fuzzy, or shift according to context, such thresholds nonetheless exist and hence are relevant to our conception and understanding of musical meter.

Meter: Some Definitions

Meter is often defined as a regular pattern of alternating strong and weak beats, but such definitions usually leave the terms "regular," "beat," and the criteria for "strong" versus "weak" undefined. Meter is also viewed, at least by most music theorists and psychologists, as being distinct from rhythm, where rhythm involves the phenomenal pattern of durations (more precisely, interonset intervals or "IOIs") and dynamic accents. It is acknowledged that the same melodic pattern may be heard in a number of different metric contexts (see Figure 1). Following the work of Jones and her colleagues (e.g., Jones, 1987, 1990, 1992), I will define meter as a stable, recurring pattern of temporal expectations, with peaks in the listener's expectations coordinated with significant events in the temporally unfolding musical surface. Following Large and Jones (1999), metrical articulations or time points may be read as the temporal locations of the peaks in a pattern of attentional pulses; a metrical pattern involves the hierarchic organization of these moments of greater (or lesser) attentional energy. Initially, meter involves synchronizing one's attention to regular features of the musical surface, but once established, a metrical pattern may be main-

Fig. 1. The same pitch/durational pattern in two different metric contexts.

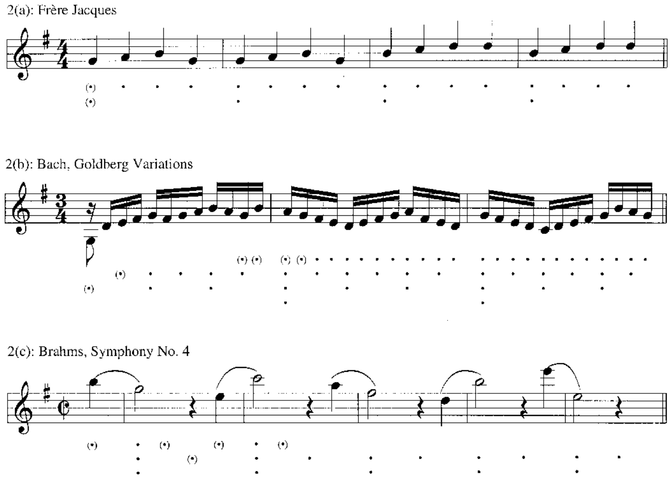

Fig. 2. Metric (attentional) structure versus the rhythmic surface of the music. (a) Frère Jacques, (b) J. S. Bach, "Goldberg" Variations, Variation 5, (c) Brahms, Symphony No. 4, first movement.

tained even in the face of a contravening musical surface (e.g., syncopations, off-beat accents).[2] To be clear, metrical attending structure does not merely consist of temporal encoding and then a feed-forward replication of event structure. Metrical behaviors are essentially dynamic and time continuous.

Metric attending involves several layers of attentional activity, and often times the metric pattern will be either "richer" or "poorer" than the durational surface. Figure 2a is a familiar melody, "Frère Jacques," made up of an isochronous series of tones at a moderate tempo. Each note articulates the pulse or tactus, and the melodic patterning makes the organization of each four-beat measure quite clear. In this case, there is a one-to-one correspondence between the pattern of our expectations and the unfolding of events on the musical surface. Figure 2b, from Bach's "Goldberg" Variations, is comprised of a stream of rapid notes, only some of which articulate the beat or tactus. When the piece begins, listeners do not know if the rapid rhythmic activity will continue and thus reward their continued attention at the smallest/most rapid levels of motion. In this case, initially

2. I have consciously avoided using the term "entrainment" here, although many do in their descriptions of meter (Brower, 1993; Gjerdingen, 1993; Jones, 1987, 1990, 1992; Jones & Boltz, 1989). Entrainment, strictly speaking, involves a phase locking of one oscillating system to another. In the discussion that follows, it is proposed that metrical attending is at times more (or less) than a phase locking of the listener's attentional rhythms with temporal invariants in the musical surface and hence is not entrainment per se. Nonetheless, I do think that meter is related to (and is perhaps a complex form of) entrainment behavior.

one might only generate expectancies relative to the quarter- or eighth-note levels, but not for each and every 16th note. As the piece continues, and the ongoing presence of the running 16ths seems secure, it then becomes possible and useful to expect a continuous level of 16th-note subdivision. To put it another way, the consistency of the rapid articulations justifies heightened expectations at that (very short) metric level. As levels of meter emerge in Figure 2b, they are first provisional (as indicated by the parentheses) and then stabilize.

Figure 2c, the opening melody from Brahms' Fourth Symphony, presents the opposite problem. When taken by itself, apart from the accompaniment in the lower strings, this melody does not articulate the tactus level. The ¢ or "cut time" signature indicates a performance tempo where the half note carries the beat, and this piece is typically performed at a tempo of half note about 90 beats/minute (660-ms IOI per half note). On the one hand, the quarter notes, which serve as anacruses to the following half notes, are too brief to carry the tactus, and thus articulate a subtactus level. On the other hand, they are also intermittent, and their lack of continuity stymies the establishment of a quarter-note level of the meter (i.e., a level of beat subdivisions). As indicated in this analysis, the listener does not generate a continuous pattern of expectation at the quarter-note level, but only at higher levels. Therefore, in this melody, one cannot generate the tactus from the bottom up. In hearing this melody (again, apart from any accompaniment), it seems likely that most listeners will interpolate the "missing" beats that occur on the second half note of each measure; these are indicated by the dots below the staff. To set up a stable meter in this instance, listeners will have to generate periodicities than are not phenomenally present in the music.[3]

A single level of periodic attending gives the listener a very limited degree of temporal expectancy: if this activity is heard as a beat or tactus, then we expect something should happen on or about the next beat. In musical terms, such a "meter" would consist of a single beat, which one would count: "1, 1, 1, 1 . . . " Although such a form of synchronization (as in many tapping studies) or expectation may be familiar to psychologists, it does not make much sense to musicians or music theorists. This is because musical meters do more than this; they also *differentiate* among our

3. Interpolation has received some recent attention in psychological research and modeling (Jones et al., 1995; Franěk, Mates, & Nártová, 2000; Rousseau & Rousseau, 1996; Yee, Holleran, & Jones, 1994). Desain and Honing (1994) have included beat interpolation component as part of their more general "beat extraction" model. Nonetheless, beat extraction or induction is not an exact analogue to beat generation by human listeners. While we base our sense of beat on periodicities that are present in the music, as noted earlier, meter is more than just an extraction of temporal invariants: it also involves the creation of invariants. Given the innate rhythmicity of human action and attention, there is more to metric attending than the optimal quantization of a rhythmic surface.

expectancies. Consider a simple duple measure: 1, 2, 1, 2, 1, 2 . . . If we are at the beginning of such a measure, although we expect something on the next beat (2), our expectations are even greater that a musically significant event will occur on the following downbeat (1). Thus meter involves a coordination of two or more attentional periodicities— two rates of attending activity. We may further stipulate that one of those periodicities must be heard as the beat or tactus—at least one must fall within the "range of usable tempos" noted below, so that, for example, a pattern consisting of a two-beat "measure" where each beat had a 2.5-s IOI would not qualify as a metrical pattern.

Although two levels of meter are necessary, three (or more) are preferable, as this provides an attending framework that allows the listener to track rapid, moderate, and relatively slow event onsets. Interestingly, the basic metric archetypes in Western music employ this threefold division of time. The nomenclature for Western meters employs a pair of descriptors, duple or triple (2 or 3 beats per measure) and simple or compound (binary or ternary subdivision of the beat). Table 1 gives the most common time signature of each metric type, though the choice of the integer beat unit is arbitrary. For example, 2/2, 2/4, and 2/8 are all time signatures for "simple duple" meters. There is often some confusion regarding compound meters. A measure of 6/8 does not articulate six beats. Because normal note orthography does not allow for a triplet division of a quarter note, in compound meters, the beat is represented by a dotted note (in 6/8, a dotted quarter), and the subdivision level is represented by an 8th. A true six-beat measure is indicated by "6/4" or "18/8" (these may also be notated as two measures of 3/4 or two measures of 9/8, respectively). Common time, or 4/4, is often represented by the ¢ and it may be regarded as a pairing of two measures of 2/4 (though it is also possible to have a 4/4 measure that is simply an ordering of 4 beats). Finally, although simple duple, simple triple, and compound duple meters are all relatively common, compound triple meter is quite rare.

Perceptual and Cognitive Constraints on Rhythm, Meter, and Tempo

As Fraisse remarked,

> In his pioneering study, Bolton (1894) worked on the problem of the limits of the frequencies at which subjective rhythmization could appear. He gave as the lower limit an interval of 115 msec, and as the upper limit, 1580 msec. These limits should command our attention, since they are approximately those of the durations on which all of our perceptions of rhythm are based. (Fraisse, 1982, p. 156)

TABLE 1
Basic Metric Types

Descriptors	Simple	Compound
Duple	2/4, 2/8, 2/2, 4/4	6/8 (6/4)
Triple	3/4, 3/2	9/8 (9/4)

As is well known to readers of this journal, Bolton's findings have been supplemented by much subsequent research, but his findings have held up remarkably well. His lower limit (≅100 ms) for subjective rhythmization also appears to be a limit for durational discrimination. Monohan and Hirsh (1990) found "listeners can discriminate an interval between two brief sounds with great precision . . . for intervals down to about 100 ms" (pp. 223–224). They also noted that "timings less than 100 ms obey a different set of rules" (p. 217). One hundred milliseconds is also the minimum time to allow for the cortical processing of musical elements (Roederer, 1995), and it correlates with Lehiste's (1970) observation that the fastest possible vocal articulation of rapidly repeated syllables is about 120 ms (pp. 6–7). Friberg and Sundström (1999), in a study of jazz drumming, found that "the absolute duration of the short note was found to be constant at 100 ms for medium to fast tempi, thus indicating a limit on tone duration." Given this pattern of results from both the perceptual and performance side of rhythmic behavior, it seems reasonable to propose that the lower limit for elements in a metrical pattern is around 100 ms.

This 100-ms limit is shorter than the lower end of the range in which one may perceive a beat or pulse. Table 2 is based on Westergaard's (1975, p. 274) chart of the range of "useful tempos" from his tonal theory textbook. Westergaard's intuitive judgments of "too fast" (as well as "too slow") correspond to empirical tests for beat perception. Warren (1993) investigated subjects' ability to perceive repeated melodic patterns and found that

TABLE 2
Range of Useful Tempos

Beats/minute	Interonset Interval (ms)	Tempo Comment
30	2000	Too slow to be useful
42	1414	Very slow
60	1000	Moderately slow
80	700	Moderate
120	500	Moderately fast
168	350	Very fast
240	250	Too fast to be useful

After Westergaard (1975), p. 274.

they were able to do so within a range of 200–2000 ms for successive note onsets. Similar results were found in studies in which subjects were asked to tap to metronomic ticks (Duke, 1989) or to musical excerpts (Drake, Penel, & Bigand, 2000). Parncutt (1994), drawing on earlier research and his own experimental studies in which subjects tapped to a variety of metric patterns at different tempos, found a range of "maximal pulse salience" that extends from 200 to 1800 ms, with a pronounced peak anchored between 600 and 700 ms (i.e., 80–100 bpm). Although we can feel a beat at any tempo within the extended range (200–1800 ms), beats are most strongly felt at a moderate tempo, around 80–90 beats/minute (600–700 ms). Fraisse (1982) notes that above 1800 ms, subjective rhythmization becomes impossible and successive sounds are not perceptually linked (p. 156). The lack of perceptual linkage prevents us from hearing such sounds in terms of a coordinated motion or movement, and Fraisse has emphasized the importance of the connection between hearing rhythm and perceiving movement (see Clarke, 1999).

Thus there is a difference at the lower end of the metric spectrum as to the shortest IOIs that can be heard as beats (≅200 ms) versus the shortest IOIs that can be accurately produced and discriminated (≅100 ms). This makes some musical sense, as the 100-ms intervals may correspond to the smallest possible IOI for a beat subdivision, which by definition will be, in the minimal case of binary subdivision (i.e., simple duple or simple triple meter), 1/2 of the beat IOI (the implications of this observation are discussed at greater length in the following section). Likewise, the upper limit for the total timespan of a metrical unit is longer than that for beat succession and subjective rhythmization (such as that reported by Bolton) because a measure is a higher order pattern of events. Woodrow (1932), in a series of extended synchronization and tapping tests, found that

> There thus appears to be one duration, at around 1.5–2 seconds, at which the reproduction of empty intervals, synchronization, and the experiencing of rhythm all begin to become difficult and another duration around 3.4 sec (2–4 sec) which represents the vanishing point of the capacity for synchronization, and (if taken as the duration of a single foot) for experiencing rhythm. The ability to reproduce empty time intervals, it is true, does not disappear at 3.4 sec. It is, however, barely possible that there may be some connection between a duration of twice this length, namely 6.8 sec, which is near the upper limit of rhythmical measure, and the fact that in this neighborhood the increase in the relative variability in the reproduction of temporal intervals . . . appears to reach its maximum. (pp. 377–378)

The constraint on the scope of larger temporal patterns is correlated with our sense of the psychological present. Although the idea of the psychological present (or "specious" present) goes back to James (1890/1950), Michon has more recently defined it as "the time interval in which sensory information, internal processing, and concurrent behavior appear to be integrated within the same span of attention" (cited in Royal, 1995, p. 33; see also Dowling & Harwood, 1986, Gabrielsson, 1993). This span of attention may vary, but as Royal (1995) notes:

> Some agreement is evident among writers concerning estimates of the typical time span of this integration. Fraisse (1984) states that it seems to have an upper limit of about 5 seconds, but is more typically nearer 2–3 seconds in time span. Michon (1978) agrees with the average value of 2–3 seconds, but places the upper limit at 7–8 seconds. Pöppel (1972) has similarly estimated the maximum duration of the psychological present to lie between 4 and 7 seconds. (p. 33)

The variation reported by these researchers is probably related to the perceptual context, and it seems reasonable to suppose that attending to a highly patterned rhythmic stimulus may afford the listener the opportunity to hear a metrical unit over a 5- to 8-second interval, whereas other, less integrated patterns will permit only a more limited metric scope. And if 2 seconds seems to be the limit for hearing successive events as temporally connected outside of a metric hierarchy, then it makes sense that the absolute value for a measure might be from about 4–6 seconds (i.e., twice or three times the length of the "slowest possible beat"). Therefore, it seems reasonable to require that rhythmic patterns be comprised of elements that are at least 100 ms in length, and that the total length of a metrical pattern cannot exceed 5-6 s. Remembering the caveats noted earlier, these values define a temporal "envelope" for musical rhythm and meter.

Durations and duration sequences that occur in the metric envelope are thus "accessible to the senses," to use Heusler's phrase (quoted in Sachs, 1953, p. 15). But we do not perform or perceive all durations and sequences within this range in the same way. Repp (1995) has observed:

> Although it may seem that rhythm should scale proportionally and remain perceptually invariant across changes in global tempo—and certainly the relative note values of simple rhythms can be reproduced and recognized across changes in tempo—several studies have suggested that subjective rhythmic organization changes with tempo (Handel, 1992; Handel & Lawson, 1983; Monohan & Hirsh, 1990; Parncutt, 1994), so that rhythms may not be executed in exactly the same way at different tempi, and listeners can find it difficult to match or recognize proportionally-scaled rhythmic patterns when the tempo is changed substantially (Handel, 1993; Sorkin & Montgomery, 1991). (p. 40)[4]

4. See also Gentner (1987) regarding the "nonscalability" of rhythmic behaviors in nonmusical contexts.

Similarly, Collyer and Church (1998) summarize:

> The temporal spectrum encompasses many orders of magnitude of time, a range so large that different mechanisms in different sub-ranges would likely be needed to achieve adequate sensitivity in all of them. . . . the range from 175ms to 1000ms is divided into at least two sub-ranges." (p. 85)

There would seem to be three significant thresholds within the metric envelope, one around 200–250 ms, another around 600–700 ms, and a third around 1.5–2 s. The last threshold is, of course, related to the limits on subjective rhythmization and beat perception noted earlier.

The special significance of the 600–700 ms period has long been known, and this value appears in many studies. Early psychophysical researchers sought to determine an "indifference interval," a tempo that sounded neither too slow nor too fast. Wundt (1911) found an average indifference interval of 600 ms, as did Fraisse (1963). Another approach investigated "spontaneous tempo" or "natural pace," which often is gathered by simply asking subjects to tap (fingers, hand, or foot) at a "comfortable" rate. Although there is a great deal of intersubject variation in these "personal tempos," a mean value across subjects also tends toward 600 ms. For example, Semjen, Vorberg, and Schulze (1998) obtained preferred tempo rates from 428.8 ms to 725 ms, with a mean around 565.3 ms (see also Fraisse, 1982). And as noted above, Parncutt found a peak in pulse salience in the 600–700 ms range.[5]

The 250-ms threshold crops up in many studies of durational discrimination (see McAuley & Kidd, 1998 for a research summary). For example, Friberg and Sundberg (1995) studied the perturbation of an element in a six-tone sequence (rather than simply a pair of durations) over a wide range of IOIs (from 100 to 1000 ms), and they reported that "the absolute JND [just noticeable difference] was found to be approximately constant at 6 ms for tone interonset intervals shorter than 240 ms and then a relative JND constant at 2.5% of the tone interonsets above 240 ms" (p. 2524). Apart from durational and tempo discrimination, other studies point to a threshold around 250 ms. Michon (1964) argued that a shift between holistic versus analytic processing occurs around 250 ms, Massaro (1970) reported 250 ms as a threshold for auditory backward masking, and

5. Drake, Jones, and Baruch (2000) have investigated developmental aspects of natural pace and preferred tempo. They found that children's preferred tempos tend to be faster (and more variable) and then seem to slow down (and become less variable) as they mature. If our sense of meter and tempo is kinematic or kinesthetic in nature, as some have suggested (Friberg & Sundberg, 1999; Palmer, 1996; Todd, 1995), and because smaller bodies will tend to have shorter eigenvalues for their fundamental movement periods (for leg motions in walking, arm swings, etc.), the results of Drake et al. make good sense.

Crowder (1993) argued that 250 ms is the limit of a short-term auditory memory that may play a causal role in some of these other observations. Fraisse (1982) argued that there are two durational categories, "*short times* of 200 to 300 ms and *long times* of 450 to 900 ms" (p. 167). Large (2000), in his study of metric categorization found that "there is some evidence that perceptual categorization operates differently depending on the absolute time intervals involved" and this difference was manifest around 250 ms.

Interactions Between Beats, Beat Subdivisions, and Tempo

There is an important relationship between beats and beat subdivisions. Recall that whereas the fastest/shortest IOIs for subjective rhythmization and meter are around 100 ms, the fastest/shortest IOIs for a beat or tactus are around 200–250 ms. As noted earlier, the roughly 2:1 relationship here is suggestive: does one or both of these limits affect or perhaps generate the other? To put it another way, does hearing a beat require, at least latently, hearing a subdivision of the beat? Subdivisions give rise to both quantitative and qualitative aspects of the beat: simple subdivision often has a stiff, march-like affect, whereas compound subdivision may have a lilt or shuffle.

One may first approach the qualitative differences in subdivision in terms of even versus uneven partitionings of the beat span (or, from a bottom-up perspective, whether or not the beat itself is composed of categorically even versus uneven subbeat units). When we think of "even" subdivision, most obviously this involves subdivisions that are ≅1/2 of the beat IOI (in other words, simple subdivision). An "uneven" subdivision involves categorically different durations, a combination of a distinct long (L) and a short (S). Yet this sets up a relationship between the duration of the S versus that of the L, and typically these are represented by ≅2:1 ratio (though other ratios are of course possible). Although these distinct durational units can be represented by any ratio (or by any set of discrete time values), uneven subdivisions give rise to a different metrical context, in that the L and S involve different "quantities" of subdivision. A subdivision that is ≅1/3 of the beat IOI is thus able to mark Ls and Ss in terms of their composition (i.e., 2 versus 1 subdivision units). Hence triplet subdivision undergirds "uneven" division of the beat into an L-S or S-L pattern. Note that this sense of "uneven" is at odds with a mathematical or psychophysical point of view, as both simple and compound subdivisions give rise to nominally isochronous (i.e. "even") metrical levels.

Table 3 illustrates how changes in the beat rate will constrain the organization of sub-beat levels. The first column gives the beat or tactus in several different ranges, each referring to average IOIs in milliseconds (N.B., these

TABLE 3
Simple and Compound Subdivision Interonset Intervals Relative to Tempo

	Subdivision	
Beat Rate	**Simple**	**Compound**
< 200 ms	None	None
200–250 ms	100–125 ms	None (66–83 ms)
250–300 ms	125–150 ms	85–95 ms (?)
300–500 ms	< 250 ms	< 250 ms
500–750 ms	> 250 ms	< 250 ms
750+ ms	> 250 ms	> 250 ms

are expressed as average IOIs, but one presumes that a degree of expressive variation may be present on each level). The second and third columns show the concomitant ranges for the IOIs involved in binary and ternary subdivision of the tactus relative to the 100-ms and 250-ms thresholds discussed earlier.[6] The first row is for a (hypothetical) beat that is less than 200 ms in duration; this row is included simply for completeness. In the next row, corresponding to very fast beat IOIs (200–250 ms) only simple (binary) subdivision is possible. As the beat slows down and approaches 300 ms, compound subdivision becomes possible. And it is here that the 250-ms threshold may come into play. As we have noted, the values for these various thresholds are only approximate and are dependent on task and context. As the IOI for the fastest level of events approaches 80–90 ms (i.e., nears the 100-ms threshold), the IOI for a ternary beat passes the 250-ms threshold. Thus if ≅200 ms is related to a "floor" for simple subdivision, then ≅250 ms may be related to a floor for compound subdivision. As the beat level IOI moves into the 300–500 ms range, both simple and compound subdivision fall below 250 ms. Interestingly, as one moves into the 500–700 ms range—that is, near the center of the range of maximal pulse salience and closest to the indifference interval—simple and compound subdivisions fall on opposite sides of the 250-ms threshold. At these tem-

6. One could also construct a similar table to describe the relationship between beats and subdivisions in more complex metrical contexts, that is, where one has uneven (nonisochronous) beats (as is the case in various African and Balkan musics, for example). Although a complete account of beat-subdivision relationships in complex meters is beyond the scope of this article (but see London, 1995), it is worth noting that the uneven beats typically involve a 3:2 relationship, so that, for example, a Short-Short-Long pattern would have a 2-2-3 relationship. Moreover, these musics very often have the subdivisions sounding in some part of the musical texture (thus making the 2s and 3s palpably clear) and are performed at relatively rapid tempi, so that both the Long and Short beats tend to fall within a 300–500 ms range.

pos, there may be a categorical difference between units of simple versus compound subdivision. Or to put it another way, near the center of the range of maximal pulse salience binary beats and measures are made up of the "same" kind of time (e.g., 300-ms subdivisions of 600-ms beats), whereas ternary beats and measures are not. This might explain in part the subjective or affective differences between simple and compound subdivision, and the use of different subdivisions to express different emotional qualities. Moreover, the differences in the temporal composition of various kinds of subdivision may also explain perceived tempo or gestural differences among passages with the same beat-level IOI.

Recall Fraisse's distinction between "*short times* of 200 to 300 ms and *long times* of 450 to 900 ms" cited above. In light of the data presented in Table 3, we may construe his distinction (which may retain a good deal of subjective validity) not simply in terms of the absolute value of short versus long times, but rather as a hierarchic manifestation of various metrical relationships. First, one may note that that long times comprise two or more short times. A somewhat more hierarchically involved way of explaining the distinction is to note that these "times" are taken as beat-level units, then short times will be made up of subdivisions less than 250 ms in length, whereas long times involve subdivisions greater than 250 ms. As such, long times also hold the potential for greater depth of subdivision, that is, that the 250 ms (or greater) level of attention may itself be divided into yet-shorter intervals, whereas short intervals have already reached the temporal floor for metrical subdivisions.

Systematic Aspects of Meter and Tempo

Having explored some of the interactions between beat and subbeat periodicities, we may now take a somewhat broader view of the hierarchic relationships amongst the periodicities in more extended metric family or system. Figure 3 is a graph of metrically related periodicities, with the central beat arbitrarily serving as the origin for the various branches of the graph.[7] Each node (apart from the central tactus) is produced by multiplicative operations on an adjacent node. As one moves out from the central

7. It is presumed that this node represents the tactus, but as Meyer and Palmer (2001) have shown, performers (and one presumes listeners) can in some contexts choose which level of activity to construe as the tactus. One could therefore generalize Figure 3 by adding additional dimensions to account for a greater number of branching possibilities at each node. As a model of a particular "tactus state" (i.e., a representation of a performer's or listener's sense of the tactus at a particular point in the course of a playing or a hearing), the graph in Figure 3 is adequate, and as a simplified representation of metrical possibilities, it allows us to see how various thresholds interact with metrical structure at different tempos.

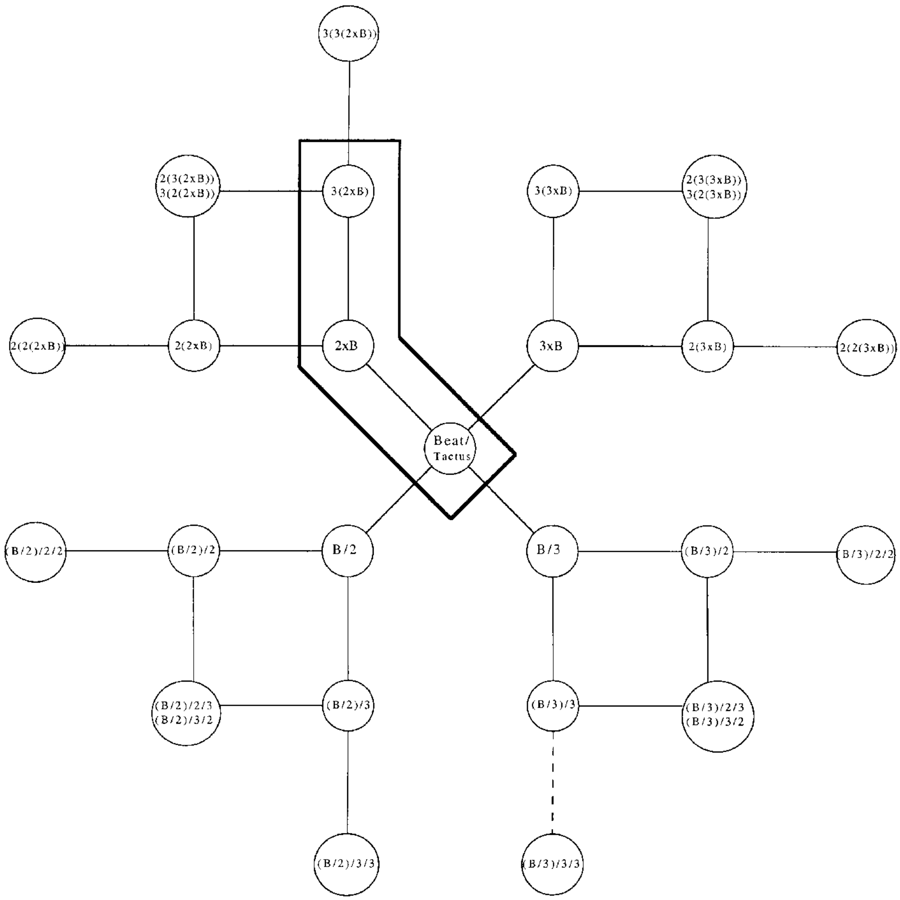

Fig. 3. Multiplicative periodicities relative to a central tactus.

tactus, the horizontal branches of the graph carry binary operations, whereas the vertical branches carry ternary operations. Each node is created by recursive operations, such that nodes that are "farther out" from the central beat are hierarchic composites of "closer" nodes. For example, if one starts with a central beat, one may create a measure of two beats, then a larger unit of 3 two-beat measures (readers are reminded that this is not a three-measure phrase, but a metrical pattern involving three layers of organization). One should therefore construe the (3(2 x B)) node on the graph as entailing three nested levels of metric structure; these are boxed in Figure 3. Likewise, any subordinate level below the beat involves nested periodicities up to the beat level itself. In this graph, a number of nodes overlap, for example 2(3(2 x B)) and 3(2(2 x B)). Although the absolute value of these two operations is the same (i.e., both result in a periodicity that is 12 times the duration of the central beat interval), their hierarchical arrangements differ. The reason for this overlap will be made clear in the discussion of Figure 5.

Figure 4 is a relabeling of Figure 3 using time signatures found in Western music. For convenience, the central beat is represented by a quarter note (a value that lies in the middle of the range of durational orthography, which is probably why 2/4, 3/4, and 4/4 signatures are so common). It can be seen that binary and ternary operations give rise to almost all of the meters found in Western classical music. Nodes above the central beat represent familiar measures; nodes below the central beat represent various layers of subdivision. As in Figure 3, higher level meters are to be under-

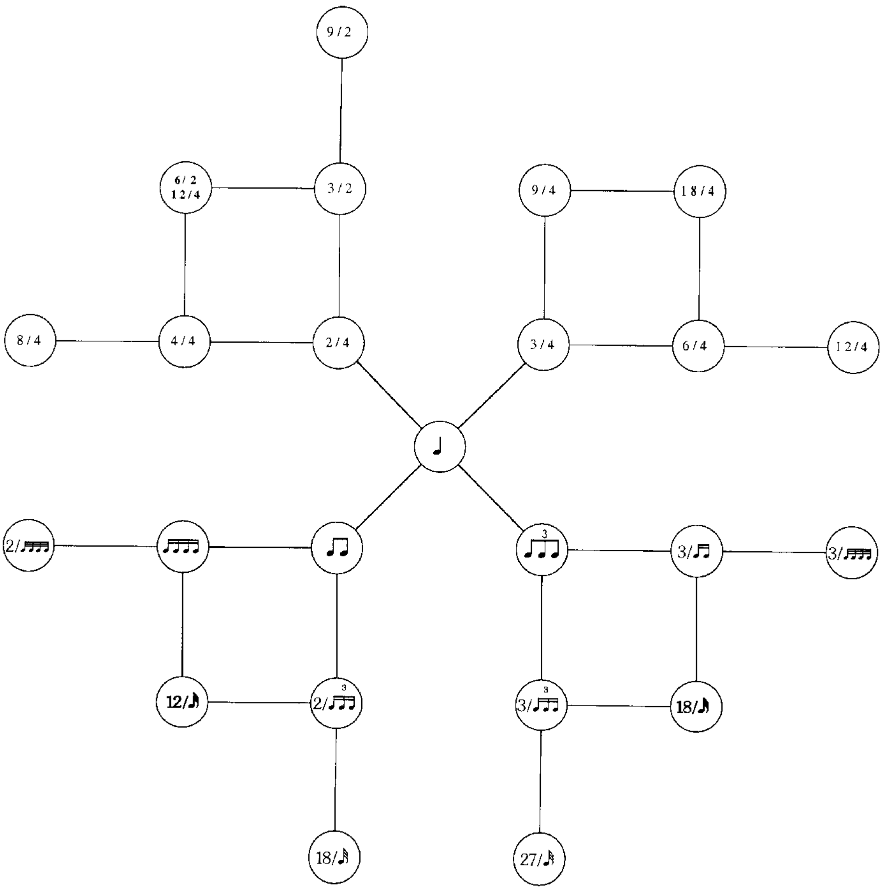

Fig. 4. Figure 3 relabeled using time signatures commonly found in Western musical notation.

stood as hierarchic composites, so that, for example, a measure 9/4 is comprised of three units of 3/4, and those 3/4 units are comprised of 3 beat-level units (and of course, those beats may be further subdivided).

Figure 4 is a generalized form of metric relationships, apart from any particular tempo value for the tactus. Once a tempo for the central beat is chosen, however, one may consider the absolute values of the resulting periodicities. Figure 5 shows the result where the central beat rate is set at 92 beats/minute (a 650-ms IOI). As can be seen, the longest periodicity is 12 s, whereas the shortest is 24 ms; these as well as many other periodicities lie outside the metric envelope. Figure 5 allows us to see how the choice of tempo serves as a constraint on metrical possibilities and on interlevel relationships relative to various perceptual and cognitive thresholds. It is for

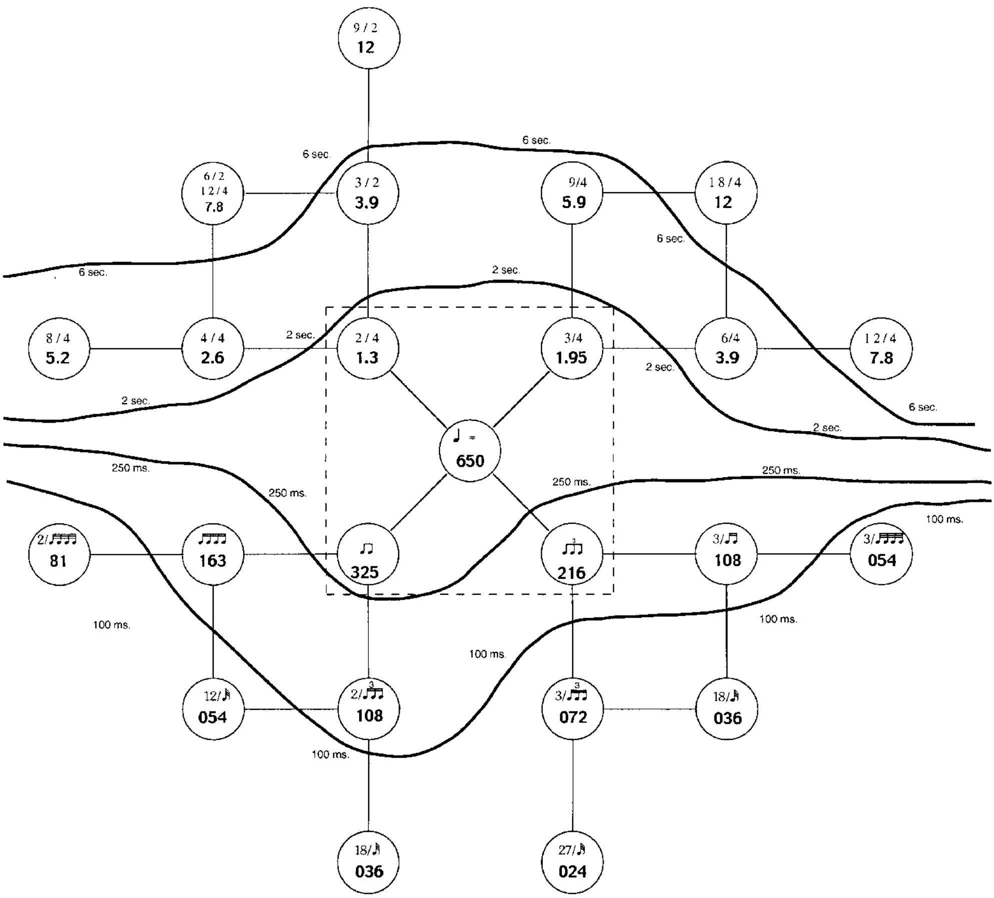

Fig. 5. Interonset intervals for various metrical levels when the tactus = 650 ms (levels above the tactus are given in seconds; levels below are given in milliseconds).

this reason that nodes with identical periodicities have been allowed to overlap [e.g., 2(3(2 x B)) and 3(2(2 x B))]. Although we wish to acknowledge that there are two distinct hierarchic arrangements at this node, in both cases the resultant IOI will be the same, and hence subject to the same tempo-related constraints. Thus if one prunes those nodes that lie outside the limiting thresholds, the graph that remains indicates what metrical relationships are possible at any given tempo.

As one might also expect, in Figure 5, the number of possible configurations is skewed toward 2:1 ratios. At this tempo, full measures of both duple or triple time fall under the 2-s threshold, suggesting that downbeats in either meter will seem strongly connected and inviting higher levels of metric structure. As noted earlier, in this tempo range, the duple subdivision is longer than 250 ms, whereas triplet subdivisions are shorter, which may give rise to categorical differences between simple versus compound subdivision.

Several metric "family trees" at various tempos are shown in Figure 6. Figure 6a shows the possible relationships when the tactus is very slow. There are a limited number of patterns above the beat (indeed, only 4/4, 2/4, or 3/4), but rich possibilities for subdivision, including periodicities in both the 100–250 ms range and the 250–600 ms range. Figure 6b is a "pruned" version of Figure 5; note how there are about the same number of metric levels both above and below the central beat. In Figure 6c, one finds the mirror image of Figure 6a—although many layers of organization are possible above the beat, there are only limited possibilities for subdivision below the beat (and note that any subdivision, if present, will have an IOI that is < 250 ms). Thus there are floor effects (which limit the possibilities for subdivision at faster tempos) and ceiling effects (which limit the possibilities for higher levels of meter at slower tempos). The graphic representations in Figure 6 make the relationship between tempo (that is, IOI of the central tactus) and the extent and shape of the "metric design space" immediately apparent.

Figure 6c also illustrates another tempo effect. If we have a strong proclivity toward periodicities in the center of the range of maximal pulse salience (i.e., around 600–700 ms), we would tend to prefer meters that include (or have the potential to include) such periodicities over those that do not. At some tempos, however, periodicities in this range are not possible. When the IOI for the central beat is at 430 ms (MM = 140), given the hierarchical relationships between subtactus and supratactus levels, there are *no* periodicities in even a 550–750 ms range. This suggests a perceptual basis regarding the performer's choice of tempo: one may gravitate toward those tempos that allow for an attentional resonance in the range of maximal salience and avoid those tempos that do not.

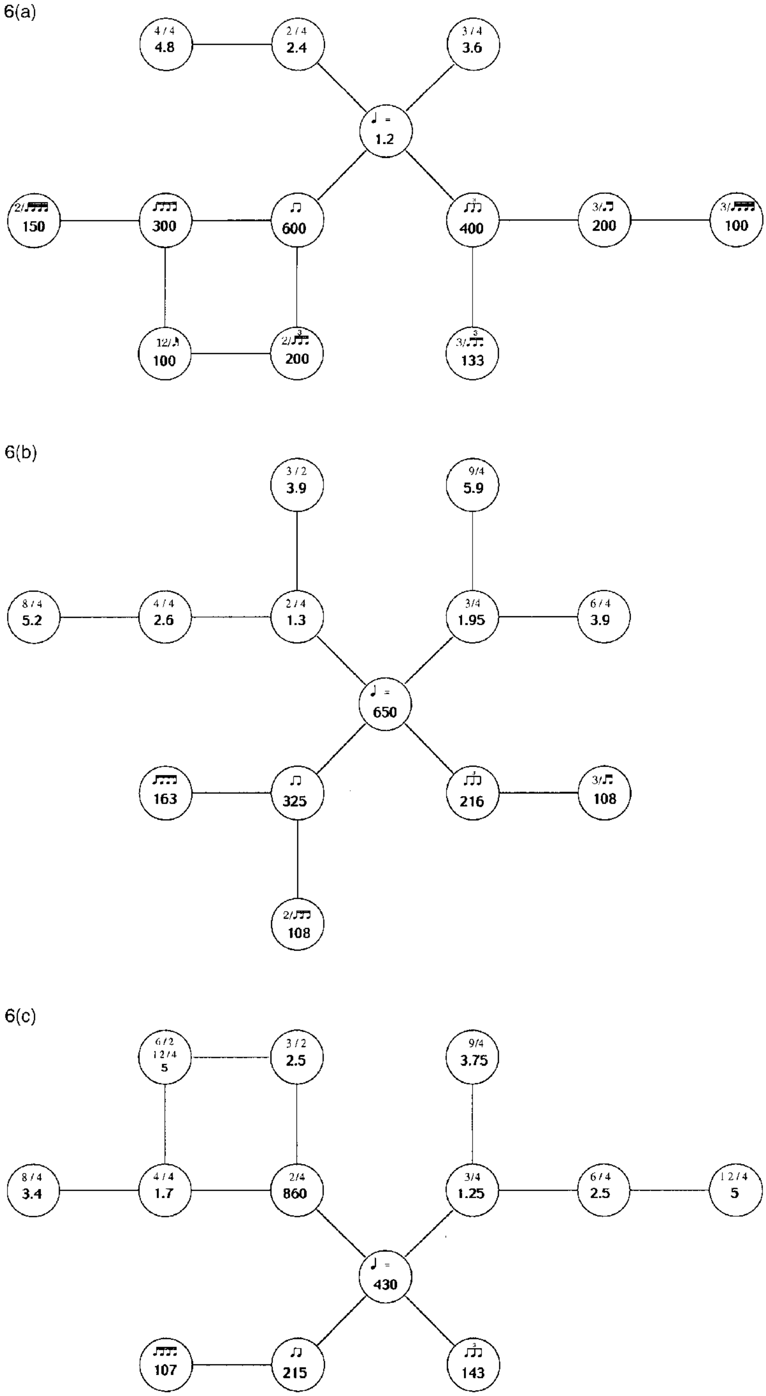

Fig. 6. Tree diagrams of metrical relationships at various tempos.

Concluding Discussion

A summary of various studies on temporal perception and performance suggests that meter can only occur with respect to periodicities in a range from about 100 ms to about 6 s. In addition, we may grasp a sense of beat or tempo in a subrange of 200 ms to about 2 s, though we have a preference for periodicities around 500–700 ms. There seems to be another important temporal threshold around 200–250 ms. Given these ranges and thresholds, most meters are a composite of different "kinds of time."

The tactus level of meter is of cardinal importance, as a metrical pattern requires a tactus coordinated with at least one other level of organization (usually a superior level, but not always). There is a strong interaction between tempo—the rate of the tactus—and the formal organization of a metric hierarchy. As the tempo changes, there may be changes in the perception of the perceived beat and there may be changes in the affective quality of the beat as well. The choice of a particular tempo also limits the scope of metric possibilities, for as the tactus moves across its range in the metric envelope, other levels are subject to various floor and ceiling affects. Changes in tempo affect both the number and the kinds of metric patterns that are possible.

Beats and subdivisions have a special relationship. The relationship between the absolute value for the lower limit on metrical attending (100 ms) and the lower limit on beat perception (200 ms) suggests that even when not phenomenally present in the musical surface, beat subdivisions may be latent in our perception of the beat. The organization of subtactus level(s) affects the perceived quality of motion of the tactus (and higher) levels. For example, the difference between an expressively uneven binary beat subdivision and a Long-Short figure with a sense of a rhythmic "lilt" lies, in the latter case, in the emergence of a compound metrical underpinning to the beat.

The subjective or qualitative differences between meters (and the same meter in different tempos) proposed here may make more sense to musicians than to psychologists—these differences are not necessarily ones that will be manifest in terms of empirical measures (reaction times, discrimination thresholds, etc.). They are more likely to be found via listener introspection, or perhaps in the kinds of affective and motional descriptions listeners will tend to give various rhythmic figures at different tempos.

The examination of how metrical hierarchies may interact with perceptual and cognitive thresholds raises a chicken-and-egg problem: are the thresholds found in psychological research artifacts of hierarchical attending/motor control strategies? Or are some aspects of temporal behavior and judgment (e.g., long-term tempo drift, categorical perceptions of rhythm and meter, attractor tempos, time estimation biases) the product of hierarchical attending/complex motor behavior interacting with certain absolute temporal thresholds? What the observations just given suggest, however, is that in order to unravel whatever interrelationship(s) there may be between various temporal thresholds and complex attentional strategies, researchers in music perception and cognition will need to develop experiments and experimental stimuli that explicitly take the hierarchical relationships among metrical levels into account.

References

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

Brower, C. (1993). Memory and the perception of rhythm. *Music Theory Spectrum, 15*(1): 19–35.

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

Collyer, C., & Church, R. M. (1998). Interresponse intervals in continuation tapping. In D. A. Rosenbaum & C. E. Collyer (Eds.),*Timing of behavior: Neural, psychological, and computational* (pp. 63–87). Cambridge, MA: MIT Press.

Crowder, R. G. (1993). Auditory memory. In S. McAdams & E. Bigand (Eds.), *Thinking in sound: The cognitive psychology of human audition* (pp. 113–145). Oxford: Oxford University Press.

Desain, P., & Honing, H. (1994). Does expressive timing in music perfomance scale proportionally with tempo? *Psychological Research, 56,* 285–292.

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

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

Drake, C., Penel, A., & Bigand, E. (2000). Tapping in time with mechanically and expressively performed music. *Music Perception, 18,* 1–23.

Duke, R. A. (1989). Musicians' perception of beat in monotonic stimuli. *Journal of Research in Music Education, 37,* 61–71.

Fraisse, P. (1963). *Psychology of time.* New York: Harper.

Fraisse, P. (1982). Rhythm and Tempo. In D. Deutsch (Ed.), *The psychology of music* (pp. 149–180). New York: Academic Press.

Franěk, M., Mates, J., & Nártová, M. (2000). Tempo change: Timing of simple ratios. In P. Desain & W. L. Windsor (Eds.), *Rhythm perception and production* (pp. 143–156). Lisse: Swets and Zeitlinger.

Friberg, A., & Sundberg, J. (1995). Time discrimination in a monotonic, isochronous sequence. *Journal of the Acoustical Society of America, 98,* 2524–2531.

Friberg, A., & Sundström, A. (1999). Jazz drummers' swing ration and its relation to the soloist. Presented at the bi-annual meeting of the Society for Music Perception and Cognition, Evanston, IL.

Gabrielsson, A. (1993). The complexities of rhythm. In T. J. Tighe & W. J. Dowling, (Eds.), *Psychology and music: The understanding of melody and rhythm* (pp. 93–120). Hillsdale, NJ: Lawrence Erlbaum.

Gjerdingen, R. O. (1993). "Smooth" rhythms as probes of entrainment. *Music Perception, 10,* 503–508.

Gentner, D. R. (1987). Timing of skilled motor performance: Tests of the proportional duration model. *Psychological Review, 94,* 255–276.

Handel, S. (1992). The differentiation of rhythmic structure. *Perception and Psychophysics, 52,* 497–507.

Handel, S. (1993). The effect of tempo and tone duration on rhythm discrimination. *Perception & Psychophysics, 54,* 370–382.

Handel, S., & Lawson, G. R. (1983). The contextual nature of rhythmic interpretation. *Perception & Psychophysics, 34,* 103–120.

Hirsh, I. J., Monohan, C. B., Grant, K. W., & Singh, P. G. (1990). Studies in auditory timing: 1. Simple patterns. *Perception and Psychophysics, 47,* 215–226.

James, W. (1950). *The principles of psychology.* New York: Dover Reprint. (Original work published 1890)

Jones, M. R. (1987). Dynamic pattern structure in music: Recent theory and research. *Perception and Psychophysics, 41,* 621–634.

Jones, M. R. (1990). Musical events and models of musical time. In R. A. Block (Ed.), *Cognitive models of psychological time* (pp. 207–240). Hillsdale, NJ: Lawrence Erlbaum.

Jones, M. R. (1992). Attending to musical events. In M. R. Jones. & S. Holleran (Eds.), *Cognitive bases of musical communication* (pp. 91–110). Washington, DC: American Psychological Association.

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

Jones, M. R., Jagacinski, R. J., Yee, W., Floyd, R. L., & Klapp, S. T. (1995). Test of attentional flexibility in listening to polyrhythmic patterns. *Journal of Experimental Psychology: Human Perception and Performance, 21,* 293–307.

Large, E. W. (2000, August). Rhythm categorization in context. In *Proceedings of the 6th International Conference on Music Perception and Cognition* [CD-ROM]. Keele, Staffordshire, U.K.: Keele University, Department of Psychology.

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

Lehiste, I. (1970). *Suprasegmentals.* Cambridge, MA: MIT Press.

London, J. (1995). Some examples of complex meters and their implications for models of metric perception. *Music Perception, 13,* 59–78.

Massaro, D. W. (1970). Retroactive interference in short-term recognition memory for pitch. *Journal of Experimental Psychology, 83*(1, Pt. 1), 32–39.

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

Meyer, R. K., & Palmer, C. (2001). Rate and tactus effects in music performance. Manuscript submitted for publication.

Michon, J. A. (1964). Studies on subjective duration. I. Differential sensitivity in the perception of repeated temporal intervals. *Acta Psychologica, 22,* 441–450.

Monohan, C. B., & Hirsh, I. J. (1990). Studies in auditory timing: 2. Rhythm patterns. *Perception and Psychophysics, 47,* 227–242.

Palmer, C. (1996). Anatomy of a performance: Sources of musical expression. *Music Perception, 13,* 433–453.

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

Repp, B. (1995). Quantitative effects of global tempo on expressive timing in music performance: Some perceptual evidence. *Music Perception, 13,* 39–57.

Roederer, J. G. (1995). *The physics and psychophysics of music: An introduction.* New York: Springer Verlag.

Rousseau, L., & Rousseau, R. (1996). Stop-reaction time and the internal clock. *Perception and Psychophysics, 58,* 434–448.

Royal, M. S. (1995). *The perception of rhythm and tempo modulation in music.* Unpublished doctoral dissertation, University of Western Ontario, London.

Sachs, C. (1953). *Rhythm and tempo.* New York: Norton.

Semjen, A., Vorberg, D., & Schulze, H. H. (1998). Getting synchronized with the metronome: Comparisons between phase and period correction. *Psychological Research, 61,* 44–55.

Sorkin, R. D., & Montgomery, D. A. (1991). Effect of time compression and expansion on the discrimination of tonal patterns. *Journal of the Acoustical Society of America, 90*(2, Pt 1), 846–857.

Todd, N. P. M. (1995). The kinematics of musical expression. *Journal of the Acoustical Society of America, 97,* 1940–1950.

van Noorden, L. (1975). *Temporal coherence in the perception of tone sequences.* Unpublished doctoral dissertation, Technische Hogeschool Eindhoven, The Netherlands.

van Noorden, L., & Moelants, D. (1999). Resonance in the perception of musical pulse. *Journal of New Music Research, 28,* 43–66.

Warren, R. M. (1993). Perception of acoustic sequences: Global integration versus temporal resolution. In S. McAdams & E. Bigand (Eds.), *Thinking in sound* (pp. 37–68). Oxford: Oxford University Press.

Wertheimer, M. (1912). Experimentelle Studien über das Sehen von Bewegung. *Zeitschrift für Psychologie, 61,* 161–265.

Westergaard, P. (1975). *An introduction to tonal theory.* New York: W.W. Norton.

Woodrow, H. (1932). The effect of rate of sequence upon the accuracy of synchronization. *Journal of Experimental Psychology, 15,* 357–379.

Wundt, W. (1911). *Grundzüge der physiologischen Psychologie.* Leipzig: Wilhelm Engelmann.

Yee, W., Holleran, S., & Jones, M. R. (1994). Sensitivity to event timing in regular and irregular sequences: Influences of musical skill. *Perception and Psychophysics, 56,* 461–471.