Tempo and Rhythm

Chapter *in* Music Perception · August 2010

DOI: 10.1007/978-1-4419-6114-3_6

J. Devin Mcauley
Michigan State University
114 PUBLICATIONS **2,586** CITATIONS

Chapter 6
Tempo and Rhythm

J. Devin McAuley

6.1 Introduction

It is a remarkable feat that listeners develop stable representations for auditory events, given the varied, and often ambiguous, temporal patterning of acoustic energy received by the ears. The focus of this chapter is on empirical and theoretical approaches to tempo and rhythm, two aspects of the temporal patterning of sound that are fundamental to musical communication.

The chapter is organized into five sections. Section 6.2 introduces basic concepts in research on tempo and rhythm and previews the topic areas that are covered in the chapter. Section 6.3 provides a general overview of two contrasting theoretical approaches that have broadly influenced research on tempo and rhythm. Sections 6.4 and 6.5 provide a selective review of research on tempo and rhythm, respectively. The chapter concludes with a summary of key points and a short discussion of promising avenues for future research.

6.2 Basic Concepts

There are two aims of this section. First, the terms tempo, rhythm, grouping, beat, and meter are defined in turn, with an emphasis on how these terms are used in the field of music perception and cognition. Second, a broad overview of the to-be-covered topic areas in research on musical tempo and rhythm is provided.

J.D. McAuley (✉)
Department of Psychology, Michigan State University, East Lansing, MI 48823, USA
e-mail: dmcauley@msu.edu

M.R. Jones et al. (eds.), *Music Perception*, Springer Handbook of Auditory Research 36,
DOI 10.1007/978-1-4419-6114-3_6,

6.2.1 Tempo

In a general sense, tempo simply means the rate (or pace) of events in the environment. A fast (or slow) tempo is a fast (or slow) event rate. In a musical sense, tempo communicates the pace of a piece of music (i.e., how fast or slow it is) and is typically associated with the rate of periodic events (beats) that listeners perceive to occur at regular (equal) temporal intervals. In a notated musical score, the intended tempo is given in terms of beats per minute (e.g., 120 bpm). The beat-per-minute convention is used in research on musical tempo, but tempo is also expressed in the literature as the time interval between successive beats (the beat period). The latter convention has been valuable because it permits a more direct comparison with the broader time perception literature.

Four areas of tempo research are discussed in the chapter: the concept of preferred tempi and limits on the range of time intervals that convey tempo information (Sect. 6.4.1), absolute memory for tempo (Sect. 6.4.2), tempo discrimination (Sect. 6.4.3), and tempo production (Sect. 6.4.4).

6.2.2 Rhythm

In music, the term *rhythm* has been used in at least two ways. It can refer either to the sound pattern or to the perception of that pattern. With respect to the sound pattern, rhythm is the serial pattern of durations marked by a series of events; in the case of music, the rhythm of a melody is the serial pattern of durations marked by sounds (notes) and silences (rests). Musical notation specifies this serial duration pattern using relative, rather than absolute, units to represent the duration or notes and rests. With respect to perception, rhythm refers to the perceived temporal organization of the physical sound pattern (i.e., the series of notes and rests). Also commonly associated with the perception of rhythm is a feeling of movement in time (Fraisse 1963; Lerdahl and Jackendoff 1983). Three fundamental characteristics of the perceived temporal organization of music are linked to the concepts of grouping, beat, and meter. Questions regarding grouping concern the *figural coding* of rhythms, while those concerning beat and meter concern the *metric coding* of rhythms (Bamberger 1980; Smith et al. 1994).

Work on rhythm is discussed in three sections: perception of grouping (Sect. 6.5.1), perception of beat and meter (Sect. 6.5.2), and models of rhythm (Sect. 6.5.3). The remainder of this section focuses on definitions for the terms grouping, beat, and meter and how these terms have been used in the literature.

6.2.2.1 Grouping

Grouping refers to how a series of notes are perceived to be clustered or grouped together. Research on principles of grouping and their role in the figural coding of rhythms has a long history, sharing similarities with work on the development of Gestalt principles of perceptual organization (Wundt 1874; Bolton 1894; Woodrow 1909; Wallin 1911). The section on the perception of grouping (Sect. 6.5.1) considers how sound characteristics, such as duration, frequency, and intensity, influence the way listeners perceive the grouping of notes in musical patterns as well as the tendency for listeners to impose grouping structure in the absence of any acoustic cues in the signal.

6.2.2.2 Beat and Meter

A basic response to music is to clap, tap, or move the body in time in a periodic fashion with a perceived pulse or beat (Parncutt 1994; Snyder and Krumhansl 2001; Temperley 2001; Large and Palmer 2002). Beats, by definition, occur at periodic intervals; in a musical score, the beat is notated. However, what is meant by a *perceived beat* in the preceding example is a series of approximately periodic time points in music that stand out in some way to the listener (i.e., they are accented). Beats (either notated or perceived) often coincide with onsets of musical notes (sounded events), but they do not have to. That is, beats can occur on silent musical elements. Ideally, the perceived beat coincides with the beat that a performer intends and/or is notated in the musical score, but whether this is true in practice depends on a variety of factors. Most music evokes a sense of beat, but there are examples of musical styles, such as Gregorian chant, wherein a periodic beat is not readily discernible (Crocker 2000).

Meter refers to the temporal organization of beats on multiple time scales. These time scales form a *metric hierarchy*, such that the beats at each level of the hierarchy periodically coincide. Beats at lower levels of the hierarchy are at a faster tempo than beats at higher levels of the hierarchy. In a musical score, the meter of the music is typically marked by a time signature that specifies two levels of the hierarchy: the primary beat (also called tactus) level and one higher level, called the *measure*. Each measure in music is subdivided into an equal number of beats, where each beat has a prescribed (notated) duration (also referred to as the beat period). Figure 6.1 provides two examples. In the left half of the figure, a 2/4 time signature specifies a measure that is equally subdivided into two equal-duration quarter notes where each quarter note is assigned a beat; this means that in a 2/4 time signature, there is a 2:1 ratio between the time span of the measure and the primary beat period; this is also called a duple meter. In the right half of the figure, a 3/4 time signature, in contrast, the measure is subdivided into three equal-duration quarter notes. In this case, there is a 3:1 ratio between the time span of the measure and the primary beat period; this is also called a triple meter. Duple and triple meters are common in Western music. *Meter perception* refers to hearing beats on multiple time scales with some beats heard as more accented (stronger) than others based on the metric hierarchy implied by the temporal structure of the rhythm. Figure 6.1b illustrates metric hierarchies for the two notated musical examples in Fig. 6.1a. Ideally, there is a correspondence between the perceived meter and the meter that a performer intends and/or is notated in the musical score. If the meter

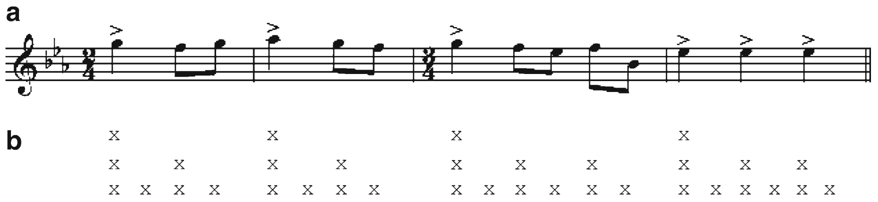

Fig. 6.1 (**a**) Musical notation for two rhythms, one with a 2/4 (duple) meter and the other with a 3/4 (triple) meter. Notated elements are quarter notes (♩) and eighth notes (♪). Eighth notes are half the duration of quarter notes. For both 2/4 and 3/4 meters, the quarter note is assigned to be the primary beat. Tempo is indicated by specifying the number of primary beats (quarter notes) per minute (e.g., ♩ = 120). (**b**) Corresponding metric hierarchies for the two rhythm examples with duple and triple meters; each row of x's marks periodic beats at one level within the hierarchy. The number of stacked x's shows the number of overlapping levels in the hierarchy at that time point. Time points with more x's are perceived to be more strongly accented than time points with fewer x's

of a piece of music is successfully communicated to a listener, then the listener hears a periodic pattern of stronger (S) and weaker (W) accents that roughly corresponds to the notated metrical structure (Cooper and Meyer 1960; Lerdahl and Jackendoff 1983). For example, for a duple meter, characteristic of marches, listeners would be expected to hear a binary SWSWSW … accent pattern, whereas for a triple meter, characteristic of waltzes, listeners would be expected to hear a ternary SWWSWWSWW … accent pattern. However, sometimes perceived meter does not precisely match the notated meter. For example, a duple meter may not be heard with a duple accent pattern, but rather with a quadruple accent pattern.

Some meters are called simple meters because they involve simple integer ratios between the measure and beat levels (e.g., 2:1, 3:1 or 4:1 for duple, triple, and quadruple meters, respectively). Complex meters, in contrast, involve subdivisions of the measure into more than four beats; sometimes these subdivisions can be unequal. Balkan folk music, with a 7/8 time signature, is one example of a complex meter (Snyder et al. 2006). Sometimes a complex meter is difficult to grasp and listeners may settle on a simpler interpretation.

Research on perception of beat and meter and the metric coding of rhythms has considered both bottom-up contributions of acoustic cues and top-down contribution of listener knowledge. The section on the perception of beat and meter (Sect. 6.5.2) focuses on contributions of different types of accents to perception of metrical structure, the role of tempo, and the role of listener knowledge. The next section turns to a general theoretical overview to provide a framework for the remainder of the chapter.

6.3 Theoretical Overview

A central theoretical issue in research on tempo and rhythm, and more broadly the field of timing, concerns the nature of the internal "clock" used to measure time. One prevalent view is that the mind's clock has many of the characteristics of a stopwatch or hourglass (Gibbon et al. 1984; Ivry and Hazeltine 1995; Mattel and Meck 2000). A contrasting view is that the mind's clock resembles in many ways a self-sustaining oscillatory process (Large and Jones 1999; McAuley and Jones 2003). Hourglass and oscillator conceptions of the clock have been linked to two general theoretical approaches, commonly referred to in the literature as interval theories and entrainment theories, respectively.

6.3.1 Interval Theories

Interval theories are based on an information-processing framework. Formal models developed within this framework typically posit distinct clock, memory, and decision stages of temporal processing (Gibbon et al. 1984; Meck 2003). Interval models have been applied to both perceptual and motor aspects of responding to musical events (Keele et al. 1989; Ivry and Hazeltine 1995; Meck 2003). Perhaps the most influential and tested of these models is scalar expectancy theory (Gibbon 1977; Gibbon et al. 1984; Church 2003). Complementary to scalar expectancy theory (SET) is the Wing and Kristofferson (W&K) model (Wing and Kristofferson 1973), which has been widely applied to rhythmic tapping behavior in an effort to distinguish clock and motor sources of performance variability.

In SET and related interval models, the clock stage involves a pacemaker, which emits over time a continuous stream of pulses that flow into an accumulator via an attention-controlled switch. At the start of a to-be-estimated (target) time interval, the switch closes, allowing pulses to flow into an accumulator; at the end of the to-be-estimated time interval, the switch opens, stopping the flow of pulses into the accumulator. The number of pulses accumulated during the target time interval provides a representation of duration. With each time interval, the accumulator is cleared with the closing of the switch; thus, the switch acts like an arbitrary reset signal, similar to how a stop watch or hourglass is reset. At the memory stage, each estimate of duration is stored in a long-term reference memory; over time, a distribution of duration codes develops. At the decision stage, the current accumulator account is compared with a temporal criterion (a stored duration code) sampled from reference memory to permit a temporal judgment, such as “shorter” or “longer.”

The following example illustrates how the general interval approach to timing works. Imagine approaching a stop light in a car for the first time. As the stop light turns red, the switch closes and pulses begin to collect in the accumulator. Similarly, when the stop light turns green the switch opens and the accumulation process stops. The number (count) of the pulses over the temporal extent of the red light provides a representation of the duration of the stop light (a duration code) that can be stored in memory. With many visits to the same stop light, a memory distribution of duration codes develops. Thus, each new visit to the light refines temporal expectations about when the stop light will turn green. Expectations come into play at the decision stage. At the decision stage, the current accumulator count is compared in a continuous fashion with a sampled count from memory. The sampled count functions as a temporal criterion. When the current count reaches the temporal criterion, then, in essence, “time is up” and the currently timed interval is expected to end. If the objective time interval ends before the count reaches the temporal criterion, then the interval is shorter than expected; conversely, if the count reaches the temporal criterion before the objective time interval ends, then the interval is longer than expected.

Interval theories have generally provided an elegant explanation for a number of timing phenomena (Meck 1996, 2005). With respect to musical tempo and rhythm, interval theories have been applied primarily to tempo discrimination (Sect. 6.4.3). However, interval theories have also been also extended to address motor aspects of responding to musical events (Sect. 6.4.4) and to a lesser extent beat and meter perception (Sect. 6.5.3).

6.3.2 Entrainment Theories

Entrainment theories derive from a dynamical systems perspective, rather than an information-processing framework. In broad terms, entrainment approaches propose that the tempo and rhythm of everyday events engage people on a moment-to-moment basis through attentional synchrony (Jones 1976; Jones and Boltz 1989; Large and Jones 1999; McAuley and Jones 2003). This approach builds on the observation that entrainment is a widespread phenomenon in nature, with entrainment of human circadian rhythms (e.g., sleep–wake cycle) with various environmental rhythms as one familiar biological example (Moore-Ede et al. 1982). In the case of circadian rhythms, the *driving* rhythm is an environmental rhythm, such as the daily light–dark cycle and the *driven* rhythm is the self-sustaining oscillation generated by pacemaker neurons in the surprachiasmatic nucleus.

Entrainment theories on a millisecond-to-second time scale are similarly based on the concept of self-sustaining entrainable oscillation (Jones and Boltz 1989; McAuley 1995; McAuley and Kidd 1998; Large and Jones 1999; Eck 2002; McAuley and Jones 2003). Applied to music, the driving rhythm is the external musical rhythm and the driven rhythm is a self-sustaining neural oscillation. Unlike with circadian rhythms, the peaks in oscillator amplitude in models of musical entrainment are assumed to represent periodic changes in gross neural activity, rather than the response of individual pacemaker cells, with the period of the oscillation providing a potential referent for judgments about relative timing; for the express purposes of chapter goals, the amplitude peaks is referred to as “beats.” Note that these “beats” of the entrainment model are subjective beats that may or may not correspond precisely with objective beats in the signal. That is, musical events (notes) may be either temporally aligned (in phase) or misaligned (out of phase) with the periodic timing of peaks in oscillation amplitude (subjective beats). Rather than distinguishing between clock and memory stages of temporal processing, the entrainment approach assumes that individuals make judgments about the timing of a sequence of musical events by detecting the synchrony/asynchrony of successive stimuli with these periodic beats. Thus, entrainment models forgo an explicit memory comparison between two coded durations in favor of dynamic information afforded by synchrony versus asynchrony between the internal driven rhythm and the external driving rhythm.

Musical events that arrive unexpectedly "early" (i.e., before an amplitude peak/beat) provide evidence that the local tempo of a sequence of musical events is accelerating, while musical events that arrive unexpectedly "late" (i.e., after a beat) imply that the local tempo of a sequence of musical events is "decelerating." McAuley and Jones (2003; see also Jones and Boltz 1989) refer to the discrepancy between expected and actual onset times as *temporal contrast*. Critically, two types of adaptive processes are assumed to operate on temporal contrast to facilitate entrainment: phase correction and period correction. The role of phase correction is to align amplitude peaks (expected time points) with musical event onsets. The role of period correction is to adjust the period of the internal oscillator so that it eventually matches salient time intervals marked by the sequence of musical events. The range of rates (tempi) that afford stable entrainment (synchronization) is referred to as an *entrainment region*. Dynamic attending theory (DAT) is a generalization of entrainment theory, whereby the internal driven rhythm is conceptualized as an attentional rhythm (Jones 1976; Large and Jones 1999).

Empirical support for entrainment theories and more broadly DAT comes from a range of sources. First, studies of overt motor tracking of event sequences indicate less variability and greater accuracy in responding to rhythmically simple events than to complex events (Jones and Pfordresher 1997; Large et al. 2002; Large and Palmer 2002; Pfordresher 2003). Second, in perceptual monitoring tasks, simple rhythms enhance performance with rhythmically expected targets, suggesting that attentional synchrony has a facilitating influence on detection of pitch, timbre, or time changes (Jones et al. 1982, 2002; Klein and Jones 1996; Jones and Yee 1997; Barnes and Jones 2000; McAuley and Jones 2003). Third, rhythms facilitate detection of temporal order of target pitches that are embedded in longer sequences (Jones et al. 1981). Fourth, in recognition memory tasks people mistakenly identify decoy melodies as target melodies when the decoys occur in a target's rhythm (Jones and Ralston 1991). Finally, the recall of pitch sequences is better when various accents (e.g., pitch skips, contour changes) are timed regularly rather than irregularly (Boltz and Jones 1986).

In sum, there are a number of key differences between entrainment and interval theories. First, from an interval perspective, duration is explicitly represented as a duration code that is stored in memory, whereas from an entrainment perspective, duration is implicitly represented by the oscillator period. Second, assumptions about responses to stimulus onsets differ. Interval models assume arbitrary reset of the pacemaker-accumulator clock with each stimulus onset, whereas entrainment models assume more gradual correction of oscillator phase and period. Third, with interval models, successive time estimates are independent, whereas in entrainment models these are dependent. Finally, with respect to duration/timing judgments, interval models involve explicit comparison of two stored duration codes whereas entrainment models involve a phase-based temporal contrast metric.

Having distinguished two major theoretical perspectives on tempo and rhythm and covered basic terminology, the next section turns to a synthesis of past and current approaches to the study of musical tempo and the relationship of this work more broadly to research on time perception.

6.4 Tempo

What is the function of the tempo of a piece of music? For one thing, tempo communicates emotion, with fast music tending to be perceived as “happy” and slow music tending to be perceived as “sad” by both children and adults (Dalla Bella et al. 2001; Gagnon and Peretz 2003; see Schellenberg, Chap. 5). More generally, tempo helps listeners track musical events as they unfold in time and enables predictions about when future events are likely to occur. This latter function of tempo appeals beyond the study of music because of the importance of prediction to general cognitive processes.

6.4.1 Tempo Limits and the Concept of Preferred Tempo

The range of tempi over which beats are perceived is limited (Fraisse 1963, 1982). If music is performed too quickly, successive sounds become indistinguishable. Conversely, if music is performed too slowly, rhythmic organization tends to fall apart, leaving only a series of isolated sounds. Between the two extremes, music and other sound patterns have perceivable rhythm. A conservative estimate for the upper (fast) tempo limit is around 100 ms between sounds (Friberg and Sundström 2002; London 2004), while an estimate for the lower (slow) tempo limit is around 2.5 s between sounds; reported values for fast and slow tempo limits vary considerably in the literature, however (Fraisse 1982; Clarke 1999; McAuley et al. 2006). Fraisse (1963) referred to the range of time intervals between 0.1 and 2.5 s as the “psychological present”; see also James (1890). Pöppel and colleagues link the slow tempo limit to the temporal capacity of working memory (Szelag et al. 1996; Pöppel 1997).

Other work has considered motor tempo limits on people’s ability to clap, tap, or generally move in synchrony with music. Notably, motor tempo limits closely parallel those observed for perception. The upper (fast) tempo limit for 1:1 synchronization (e.g., one tap per beat) tends to be partly constrained by the rate at which an individual can move, but is also likely driven by the increased proportional variability that is found at fast rates. For hand tapping, the upper (fast) limit is approximately 150–200 ms between taps (Fraisse 1982; Repp 2003; McAuley et al. 2006) and the lower (slow) tempo limit is about 2 s between taps. For time intervals longer than 2 s or so between taps, participants tend to have difficult predicting when the musical event (e.g., tone) will occur and simply reacting to, rather than anticipating, tone onsets. Research on tempo limits associated with synchronizing movements with music is generally consistent with entrainment theories, which have proposed that synchronization should be most accurate within a limited entrainment region, with less stable synchronization performance and increased variability outside that region. At least one developmental study has suggested that the entrainment region is narrower for children and older adults compared with young adults (McAuley et al. 2006).

Within the tempo limits that define perceivable rhythms and afford synchronization, individuals demonstrate clear tempo preferences. The concept of a preferred tempo has been widely studied (Stern 1900; Wallin 1911; Fraisse 1963, 1982; Jones 1976; McAuley et al. 2006) and names given for preferred tempo in the literature vary considerably; these include “mental tempo,” “personal tempo,” “psychic tempo,” and “internal tempo” (Stern 1900; Rimoldi 1951; Mishima 1956; Fraisse 1963, 1982; Boltz 1994; Vanneste et al. 2001). The various names for preferred tempo reflect different assumptions about what preferred tempo means. Does preferred tempo simply index a preferred rate of spontaneous motor activity, or a preferred rate of listening, or does it measure a broader cognitive tempo preference? From an entrainment perspective, the concept of a preferred tempo is linked to the intrinsic period of the driven oscillator (McAuley et al. 2006). From an interval perspective, preferred tempo is sometimes associated with clock speed (i.e., rate of pulse accumulation; Vanneste et al. 2001). Independent of theoretical orientation or assumptions about the underlying meaning of preferred tempo, assessments of preferred tempo have generally emphasized either spontaneous motor measures or perceptual measures. Work on motor and perceptual measures of preferred tempo is reviewed in turn next.

6.4.1.1 Spontaneous Motor Tempo

Stern (1900) was one of the first researchers to suggest that the tempo of spontaneous motor activity provides some insight about the pace of mental activity. To measure this, Stern asked individuals to tap their hands on a table at a rate they considered just right (not too fast or too slow). This assessment of the tempo of spontaneous motor activity has become one of the most widely used measures of preferred tempo. Not all researchers agree, however, that spontaneous motor tempo reflects the pace of mental activity. Nonetheless, a common assumption of much work in this area is that the preferred tempo of spontaneous rhythmic motor activities, such as walking or clapping, does provide some insight about the pace of an internal “mental” clock involved in the perception of time (Boltz 1994; Vanneste et al. 2001).

The most representative value of spontaneous motor tempo (SMT) reported in the literature for adults is around 600 ms (Fraisse 1982). Across studies a representative range extends from approximately 300 ms to 800 ms (Frischeisen-Köhler 1933;

Mishima 1956; Smoll and Schutz 1978; Fraisse 1982; McAuley et al. 2006). One notable exception is Collyer et al. (1994), who report a bimodal distribution of spontaneous motor tempi with modes around 272 ms and 450 ms.

Although 600 ms is a representative value of SMT, there are also large individual differences. Observed values in individual assessments of SMT vary widely; SMT can be as short as 200 ms or as long as 1,600 ms (Drake et al. 2000; McAuley et al. 2006). Despite large individual differences, measures of spontaneous motor tempo tend to be reliable. Spontaneous motor tempo is very stable within a given production, with the sequence of produced intervals varying on average by ≈5%. This is similar to the degree of variability observed in assessments of duration discrimination. When multiple measures of SMT are taken, correlations across these measures range from 0.75 to 0.95 (Harrel 1937; Rimoldi 1951; McAuley et al. 2006).

6.4.1.2 Preferred Perceptual Tempo

Whereas spontaneous motor tempo refers to the natural or preferred rate of rhythmic motor activity (e.g., tapping), preferred perceptual tempo refers to the rate of a series of sounds or lights that is judged to be neither too fast, nor too slow, but appears to be “just right” (Fraisse 1982; McAuley et al. 2006). Early work on preferred perceptual tempo (PPT) concerned identifying an *indifference interval* that individuals perceived as neither too short nor too long (Vierordt 1868; Woodrow 1951). Like SMT, the most commonly reported value for this interval is around 600 ms (Frischeisen-Köhler 1933; Mishima 1956; McAuley et al. 2006), but a wide range of values have also been reported over the years (Wallin 1911; Woodrow 1951; Fraisse 1982; McAuley et al. 2006). Thus, notably, SMT and PPT have comparable frequencies. However, evidence for the correlation between the two is mixed (Fraisse 1982). Some of the strongest support for a correlation between the two is reported by McAuley et al. (2006), who found a large, positive, correlation between SMT and PPT, i.e., near 0.75. Such correlations support the view that motor and perceptual tempo preferences have a common psychological basis.

6.4.1.3 Factors Affecting Preferred Tempo

Developmental studies of preferred tempo reveal that both preferred perceptual tempo and spontaneous motor tempo slow with increased age (Drake et al. 2000; Vanneste et al. 2001; McAuley et al. 2006). In the most extensive of these studies, McAuley et al. (2006) examined PPT and SMT for participants between the ages of 4 and 95 years and found that the preference for slower sequences (PPT) increased systematically with age, thus paralleling a similar age-related motor slowing in SMT. Moreover, PPT was also found to highly correlate with SMT across all age ranges. For children between the ages of 4 and 7, preferred tempo was typically between 300 and 400 ms; for adults, preferred tempo was around 600 ms, while for older adults preferred tempo was close to 700 ms. Drake et al. (2000) and Vanneste et al. (2001) reported similar age-related trends. In sum, there is emerging support for the age-related slowing of perceptual and motor measures of preferred tempo across the lifespan. Interpreted within an interval model framework, these data support the view that the internal clock slows with increased age. From an entrainment model perspective, these data suggest an age-related lengthening of intrinsic oscillator period.

Fraisse (1982) reported that differences in SMT between two identical twins are no different than two productions of SMT by the same subject, but that differences in SMT between two fraternal twins are as large as those found between two individuals selected at random. This suggests that there may be a genetic basis for preferred tempo, but there is limited evidence to support this view. For example, it appears that elements of preferred tempo are learned. Drake et al. (2000) showed that preferred tempi of children with musical training tend to shift to adult levels sooner that children without musical training. Thus, some of the observed developmental changes in preferred tempo may be a consequence of experience, rather than maturation, with more musical experience speeding the rate of developmental change. A learning perspective is consistent with other work showing that even on the time scale of an experiment, participants develop a general sense of the average pace of the events they experience, and this average pace can serve as a referent for participants judgments about the tempo (or duration) of events in their environment (Jones and McAuley 2005; Miller and McAuley 2005).

Aside from effects of age, individual differences in preferred tempo do not appear to be associated with other intrinsic factors, such as gender, handedness, or body size. There is also little evidence to link preferred tempo to specific physiological variables, such as heart rate. One study suggestive of a link between preferred tempo and physiological variables is the work of Boltz (1994), who showed that exposure to annoying sounds, such as a car horn, hypothesized to increase general arousal levels, tended to speed up preferred tempo, while exposure to relaxing music, hypothesized to decrease general arousal levels, tended to slow down preferred tempo. Boltz further showed that changes in preferred tempo predictably influenced individuals' judgments about learned event durations. This latter finding supports a link between preferred tempo and the speed of an internal clock if a theorist is interval model inclined, or a link to intrinsic oscillator period if one is more entrainment oriented.

6.4.2 Absolute Memory for Tempo

Music typically can be performed at a range of tempi and still retain its identity (Andrews et al. 1998). For example, "Happy Birthday" can be sung fast or slow and still be recognized by listeners as "Happy Birthday." This flexibility in recognition occurs because the identity of a piece of a melody does not, typically, depend on the absolute durations of individual notes, but rather on the pattern of relative durations; indeed, this aspect of music perception is captured by musical notation, which specifies only in relative terms how long each note should be sung or played. The fact that tune identity typically depends on the relative, rather than absolute, duration of notes, raises the interesting question of whether individuals possess memories for the absolute (e.g., millisecond) temporal features of music even though it is the relative patterning of durations that matters most for recognition.

Researchers interested in this question have tended to focus on whether people possess absolute memory for tempo (Farnsworth et al. 1934; Halpern 1988; Bergeson and Trehub 2002; Levitin and Cook 1996). Bergeson and Trehub (2002) compared the tempi of mothers' singing to infants to the tempi of their spoken utterances and found that for singing, repeated productions of the same song demonstrated remarkable stability in tempo, as well as pitch and rhythm. With respect to tempo, differences across song productions varied by only around 3%. For speech, in contrast, repeated productions of the same utterance showed much less stability in pitch and tempo, although speech rhythm tended to be preserved. Tempo differences for repeated productions of the same utterance were on the order of 20%.

Other work suggests that whether listeners retain the absolute temporal features of music may depend on the degree to which they have directly experienced tempo variations in a piece of music. Music with little variation in the experienced tempo, such as commercial recordings of pop songs heard on the radio, are said to have a tempo standard; that is, repeated exposures to the song are always at about the same tempo. Levitin and Cook (1996) examined absolute memory for the tempo by having adult listeners without musical training sing pop songs from memory using a cued list of songs typically heard on the radio. Consistent with the view that individuals encode the absolute temporal features of music, song productions of 72% of the participants were within 8% of the actual tempo of the commercial recordings frequently heard on the radio. Notably, this degree of variability was no different that the typical range of tempo discrimination thresholds reported in the literature. Moreover, the correlation between produced and actual tempo was 0.95. In contrast, vocal productions of a set of familiar folk songs without a tempo standard revealed substantially more variability. Moreover, because people were shown to be capable of producing song tempi at very fast and very slow rates, this ruled out the possibility that the precise tempo memory observed for pop songs was simply due to production (performance) constraints.

Halpern (1988) considered how closely the remembered (imagined) tempo of familiar music matched the perceived tempo using a set of familiar songs that notably did not have a tempo standard. For the perception version of this task, participants listened to a song and adjusted a metronome until they arrived at their preferred tempo for the song. For the imagery version, participants imagined the tune they heard and adjusted clicks of a metronome until they matched beats of the imagined song. Providing support for the view that tempo is represented in auditory imagery, the correlation between perceived and imagined song tempi was around 0.6. Moreover, there was a tendency for the imagined tempo of a song to be faster than the perceived tempo when the perceived tempo was slow, whereas the reverse was true when the perceived tempo was fast. This suggests that the imagined tempo of a song tends to gravitate to a mean rate. Notably, the average time interval between beats for perceived and imagined tempi was around 600 ms, which is consistent with work on preferred tempo. Taken together, representation of absolute tempo information in auditory imagery appears to include both a local learned component that reflects the tempo of the particular piece of music and a more global temporal context component. It is not clear whether the more global component simply reflects the average tempo experienced by listeners or an individual's intrinsic preferred tempo, which conceivably may have a genetic basis.

In sum, there is converging evidence that musicians and nonmusicians alike develop fairly precise memories for the absolute tempo of music, as long as the music has a tempo standard (i.e., it is almost always performed at about the same tempo). This is not true for music without a tempo standard or for the tempo of spoken utterances. Systematic deviations in memory for the absolute tempo of songs with a tempo standard and for speech are reminiscent of some of the work on preferred tempo. Having discussed tempo preferences and listeners' ability to develop long-term memory for tempi, the next section addresses listeners' ability to detect changes in tempo.

6.4.3 Tempo Discrimination

Another area of tempo research addresses listeners' ability to detect changes in tempo. Two questions have guided this research. The first concerns whether tempo discrimination thresholds obey Weber's law; the second question concerns the impact of multiple sequence intervals on thresholds.

6.4.3.1 Weber's Law and the Multiple-Interval Advantage

Studies of tempo discrimination have involved both auditory and visual sequences (Michon 1964; Schulze 1978, 1989; Drake and Botte 1993; ten Hoopen et al. 1994; Ivry and Hazeltine 1995; Vos et al. 1997; McAuley and Kidd 1998; Grondin 2001a). A central debate in this literature concerns whether discrimination thresholds obey Weber's law. Weber's law is assessed by either perceptual measures or motor measures. Motor assessments of Weber's law is discussed in Sect. 6.4.4. With respect to perceptual measures, if Weber's law holds for a range of tempi (time intervals), then the just-noticeable difference in tempo (ΔT) between two sequences should be a constant proportion (or percentage) of a base time interval, T, where the value of T is typically given by a fixed referent time interval marked by successive tone onsets in one of the two sequences. A formal way of expressing Weber's law is that $\Delta T/T$ should be equal to a constant, k. The value of k is called a Weber fraction. A Weber fraction of 0.1 indicates that a listener can detect a 50-ms difference given a T of 500 ms, but would need at least a 100-ms difference for a T of 1,000 ms. Empirical studies have examined the simple version of Weber's law described in the preceding text and generalized variants; see Grondin (2001b) for a comprehensive review.

Many of the studies addressing the issue of Weber's law in a musical context have focused on tempo judgments about isochronous (equal interval) rhythms. Figure 6.2 illustrates a typical tempo discrimination task, in which listeners hear two monotone sequences, a standard sequence followed by a comparison sequence, and they must judge the tempo of the comparison relative to the standard. In general, tempo discrimination is poorer when the two sequences are irregularly timed than when they are regularly timed sequences (Drake and Botte 1993).

Michon (1964) measured listeners' ability to detect tempo differences for interonset intervals (IOIs) between 67 and 2,700 ms using relatively long sequences and very well practiced subjects. Counter to Weber's law, he found a minimum relative just-noticeable difference (JND) that was as small as 1.0% for a 100-ms IOI with a secondary minimum region around 2.0% for IOIs between 300 and 1,000 ms. Both minimums were much lower than the approximate 6% relative JNDs reported for isolated intervals. This latter finding suggested an advantage afforded by multiple sequence intervals.

The issue of single versus multiple intervals was considered in detail by Drake and Botte (1993). They examined listeners' ability to detect tempo differences in isochronous sequence for IOIs between 100 and 1,500 ms for 1, 2, 4, and 6 interval sequences. For single-interval sequences, tempo relative JNDs were approximately 6%, which is a similar value to that found in duration discrimination studies (Woodrow 1951; Creelman 1962; Small and Campbell 1962; Abel 1972; Getty 1975; Allan 1979; Drake and Botte 1993). For multiple-interval sequences, thresholds improved, on average, to 3%. Best performance was found for 6-interval sequences for a 400-ms tempo with the reported threshold slightly below 2%. Similar to Michon (1964), relative JNDs for tempi between 100 and 1,500 ms were not constant, as predicted by Weber's law, but rather were a U-shaped function. This finding is consistent with the view taken by some that tempo discrimination is best (i.e., thresholds lowest) at a listener's preferred tempo. Relative JNDs were approximately constant for IOIs between 300 and 800 ms, increasing for both shorter and longer values.

Overall, Drake and Botte observed more improvement with increasing number of intervals at fast tempi than at slow tempi, which is also in line with the work of Michon (1964). Results consistent with these findings were also reported by McAuley and Kidd (1998) for time intervals between 100 and 1,000 ms for 1- and 3-interval sequences; increasing the number of sequence intervals reduced thresholds, especially at the faster tempi.

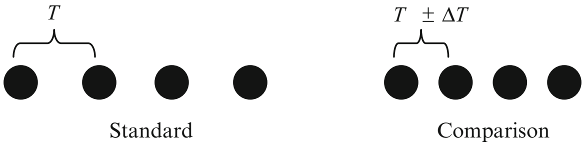

Fig. 6.2 Illustration of a tempo-discrimination task. Two monotone sequences, a four-tone standard sequence with a fixed time interval, T, between tone onsets, followed by a four-tone comparison sequence with a fixed time interval $T \pm \Delta T$. The value of ΔT typically varies from trial to trial. The listeners' task is to judge the tempo of the comparison sequence relative to the standard

In sum, researchers considering the question of Weber's law for tempo discrimination and effects of sequence length (number of equal intervals) on thresholds have found that Weber's law holds only within a limited range of tempi. Notably, this optimal tempo zone encompasses commonly reported values for preferred tempo. Moreover, as the number of equal intervals increases, listeners' ability to detect changes in tempo improves. Across studies JNDs for single-interval sequences are typically on the order of 6%, while JNDs for multiple-interval sequences are sometimes less than about 2% (Michon 1964; Drake and Botte 1993; Friberg and Sundberg 1995).

6.4.3.2 Locus of the Multiple-Interval Advantage

Within an interval-model framework, reduced tempo sensitivity (increased threshold) is typically attributed to increased variability of clock, memory, or decision processes. From this perspective, Drake and Botte proposed to explain improvements in tempo discrimination thresholds associated with the number of sequence intervals using a multiple-look model whereby each equal time interval in an isochronous standard sequence provides an independent but variable estimate of sequence tempo. They hypothesized, as have others, that listening to the standard sequence leads to a series of independently sampled estimates of the tempo of the standard sequence, which are averaged to form an aggregate memory trace (Keele et al. 1989; Schulze 1989; Drake and Botte 1993; Ivry and Hazeltine 1995). As the number of independent "looks" increases, the average sampling error between the estimated and actual standard tempo decreases, leading to lower discrimination thresholds. Drake and Botte predicted that the JND in tempo, taken as the standard deviation of the sampling distribution, should decrease inversely to the square root of the number of standard sequence intervals, as shown here:

$$\mathrm{JND}n = \frac{\mathrm{JND1}}{\sqrt{n}} \tag{6.1}$$

In this model, JND_1 is the observed JND for a single-interval standard sequence and JND_n is the predicted JND for an n-interval standard sequence.

Studies have reported data consistent with the multiple-look model for both auditory and visual sequences for tasks involving time-interval perception as well as production (ten Hoopen and Akerboom 1983; Ivry and Hazeltine 1995; Rousseau and Rousseau 1996; McAuley and Kidd 1998; McAuley and Jones 2003). There are notable exceptions, however. Some studies have reported mixed results (Schulze 1989; Hirsh et al. 1990; Grondin 2001a), whereas others have found no multiple-interval advantage (ten Hoopen et al. 1994; Pashler 2001). One factor preventing a clear interpretation of some of this research is that the numbers of standard and comparison intervals have sometimes covaried, making the locus of the multiple-interval advantage unclear (Drake and Botte 1993; Grondin 2001a). That is, does the multiple-interval advantage occur because of multiple intervals in the first (standard) sequence, the second (comparison) sequence, or both?

Miller and McAuley (2005) investigated the locus of the multiple-interval advantage by independently varying the number of standard and comparison intervals in standard-comparison pairs of isochronous tone sequences. They found, somewhat surprisingly, that with a *fixed* standard tempo on each trial, the multiple-interval advantage occurs because of multiple intervals in the comparison sequence, rather than multiple intervals in the standard sequence. However, when tempo of a standard sequence varies from trial to trial, both the standard sequence and the comparison sequence contribute to the multiple-interval advantage; see also Grondin and McAuley (2009).

To account for distinct contributions of the number of standard and comparison intervals to tempo discrimination thresholds, Miller and McAuley (2005) proposed a generalized multiple-look (GML) model, which extended the original multiple-look model of Drake and Botte (1993) to measure average sampling error for two independent samples corresponding to the number of sampled intervals from the standard and comparison sequences, respectively. As with the Drake and Botte multiple-look model, thresholds in the GML model are predicted to be inversely related to the number of equal sequence intervals. However, unlike the Drake and Botte model, a weight parameter, w, in the GML model permits the threshold contribution of the standard and comparison sequences to vary. One general observation about multiple-look models and related interval approaches is that although they provide a descriptive account of tempo threshold data, a common weakness is that they often do not make explicit predictions about other dependent measures such as points of subjective equality (see Jones and McAuley 2005).

The entrainment approach offers an alternative account of tempo discrimination data. According to this perspective, an effect of the number of standard intervals on tempo thresholds is due to period correction processes, whereas an effect of the number of comparison intervals on tempo thresholds is due to listeners' reliance on temporal contrast information (i.e., judgments about how aligned successive tone onsets in the comparison sequence are with internally generated beats). Thus, increasing the number of standard intervals should produce improvements in tempo sensitivity in situations that require substantial period correction (e.g., when the standard sequence tempo differs from the average tempo) but not in situations that require little or no period correction (e.g., when the standard sequence tempo is at the average tempo). Conversely, increasing the number of comparison intervals is likely to lead to improvements in tempo sensitivity in situations that require little or no period correction (i.e., when there is a close match between the period of an induced internal oscillation and the standard interval). It is precisely in the latter situations where relative phase discrepancies of tone onsets in the comparison sequence (temporal contrasts) provide the most reliable information about the relative tempo of the comparison sequence ("faster" or "slower"). Empirical data from time judgment tasks have supported these entrainment model predictions (Large and Jones 1999; Barnes and Jones 2000; McAuley and Jones 2003; Miller and McAuley 2005; Jones and McAuley 2005).

6.4.4 Produced Tempo

Some of the same issues addressed in research on tempo discrimination discussed in the preceding section emerge in research on performance aspects of music. As with tempo discrimination, a key question concerns whether the tempi that people produce in music performance or in simply tapping along to music obey Weber's law. This question has been most often addressed in the context of a synchronize-continue tapping task, first introduced by Stevens (1886). In its simplest form, synchronize-continue tapping requires individuals to generate a series of finger taps in synchrony with a metronome set to a particular tempo (or target time interval, *T*), and then to continue tapping at the same rate in the absence of the pacing stimulus. Variants of this paradigm have asked individuals to tap in synchrony with the beat of a musical excerpt, at subdivisions or multiples of the beat (i.e., at different levels of the metric hierarchy), and have compared tapping to simple and complex meters (Patel et al. 2005; Repp 2005; Snyder et al. 2006). Performance across all versions of the paradigm is typically evaluated by the phase alignment of taps with sounds during the synchronization phase, as well as the mean and variance of both the produced intervals for both the synchronization and continuation phases of the task. For the purposes of comparison to work on perceived tempo, the emphasis of this section is on the simple version of the paradigm involving isochronous sequences and assessments of the mean and variability of produced time intervals (i.e., the produced tempo).

A widely cited modeling approach to synchronize-continue tapping performance is the interval model of Wing and Kristofferson (termed the W&K model; Wing and Kristofferson 1973). In the W&K model, synchronize-continue tapping entails a series of responses (taps) to each in a series of multiple intervals. Each tap is triggered by an internal clock which, in synchronizing to an isochronous sequence, reflects an encoding of *T*. That is, stimulus tempo, expressed as a time interval *T*, determines the number of pacemaker pulses corresponding to the stored interval code (*C*) and is used to meter out each time interval between successive taps. Further, in continuation tapping, where people must sustain the induced sequence rate in the absence of tones, this model predicts the *n*th produced tapped interval: I_n. Specifically, this interval is given by an additive combination of the nth interval code (C_n) and peripheral (motor) delays associated with taps that intiate (D_{n-1}) and terminate (D_n) this interval:

$$I_n = C_n + D_n - D_{n-1} \tag{6.2}$$

According to this model, the variability of produced intervals (*I*) derives from two sources: clock variance and motor variance. Moreover, with an assumption of independence between clock and motor components, it is possible to decompose the total tapping variance (σ_I^2) into separate estimates of clock (σ_C^2) and motor (σ_D^2) sources of variability (Wing 1980; Wing and Kristofferson 1973).

As with interval models of tempo discrimination described in the preceding section, the W&K model makes predictions about accuracy and variability of performance (here motor performance) as a function of tempo, *T*. With regard to accuracy, a central prediction of the W&K model holds that the averaged produced interval in continuation tapping will approximate T; no systematic over- or underestimations of T are predicted. This is because clock values (C_n) should always center on the target interval, T. Although this prediction generally holds for young adults, children and older adults frequently fail to maintain a constant rate of tapping (Ivry and Keele 1989; Williams et al. 1992; Greene and Williams 1993; McAuley et al. 2006), suggesting rather periodic drift to a preferred tempo. In these cases, period drift is typically treated as a nuisance variable and eliminated from the time series using a linear or nonlinear detrending procedure (Ogden and Collier 1999). In contrast, for some entrainment models detrending eliminates an important component of behavior. This is because these models assume that an intrinsic oscillator period, which gradually adjusts during synchronization tapping, will exhibit a diagnostic trend of drifting back to an intrinsic preferred rate in continuation tapping.

A second prediction of the W&K model concerns tapping variability. This revives issues surrounding tempo and Weber's law in tempo discrimination. For the W&K model the decomposition of total tapping variance into clock and motor components is crucial because it leads to the prediction that clock variance increases linearly with the target interval, T, whereas motor variance remains constant. This is an important prediction, not only because it is the mainstay of the W&K model, but also because it violates Weber's law. According to Weber's law, the standard deviation of produced intervals should be linearly related to tempo: that is, as T increases, the standard deviation of taps, which reflect a motor JND, should increase proportionally. However, because the W&K model assumes clock variance, σ_C^2, increases linearly with T, it predicts that observed tapping variance will increase much more with slower tempi than predicted by Weber's law. Although some applications of the W&K model to continuation tapping of young adults have been moderately successful (Wing 1980) in explaining tapping variability, their generality remains in doubt. Such contrasting predictions about tempo and Weber's law remain the source of much lively debate in the literature.

Using the W&K approach, researchers have explored the possibility that children and older adults show increased clock variance, increased motor variance, or both, relative to young adults. From this work, it is clear that tapping of young children is more variable than tapping of young adults for the reported tempi (Ivry and Keele 1989; Greene and Williams 1993; Geuze and Kalverboer 1994). Less clear is the extent of age differences between older and younger adults (Duchek et al. 1994; Vanneste et al. 2001; Krampe et al. 2005). Many of these studies have attributed observed age-related differences in tapping variability to clock rather than motor sources. One limitation of some of the research using the W&K model is that despite theoretical predictions about the relationship between tempo and variability, the W&K model often has been evaluated at only a single tempo. When a broad range of tempi are considered, the W&K model does not generally fare too well. Rather, consistent with the Drake and Botte (1993) research on tempo discrimination and the multiple-interval advantage discussed in the preceding section, Weber's law appears to provide a better account of the data, but only for a limited range of tempi. That is, for a restricted range of target intervals, T, the standard deviation, not the variance, of produced intervals increased linearly with T, resulting in a constant Weber fraction. Outside this range at both faster and slow tempi, the Weber fraction tends to increase.

Having discussed theoretical approaches to musical tempo and key empirical findings, the next section turns to a synthesis of past and current approaches to the study of musical rhythm.

6.5 Rhythm

Empirical and theoretical approaches to rhythm have often distinguished the perception of grouping and the figural coding of rhythms from the perception of beat and meter and the metric coding of rhythms (Bamberger 1980; Povel and Essens 1985; Ross and Houtsma 1994; Hébert and Cuddy 2002). Perception of grouping involves the segmentation of a rhythm into clusters of elements (e.g., a group of two, following by a group of three). A figural code based on principles of grouping preserves the number of tones in each successive group (figure) and the number of figures in each repetition of the rhythm. Perception of beat and meter involves the perception of temporal regularity on multiple hierarchically related time scales. Metric coding minimally involves the representation of at least two level of the metric hierarchy and preserves the relative durations of elements of the rhythm. All rhythms do not necessarily afford metric coding (i.e., fit within a metric framework; see Povel and Essens 1985). Moreover, rhythm discrimination can rely on differences in grouping structure (figural coding), metrical structure (metric coding), or both (Handel 1998).

A selective review of rhythm research is presented in three sections. Section 6.5.1 describes work on perception of grouping. Section 6.5.2 describes work on perception of beat and meter. Section 6.5.3 considers several influential models of rhythm.

6.5.1 Perception of Grouping

A range of acoustic cues, including frequency, duration, and amplitude (intensity) have the potential, depending on their patterning, to convey to the listener a sense of inherent sequence organization or structure (Fraisse 1956; Handel 1989). The focus of this section is on the perception of grouping, namely that some sequence elements belong together (i.e., they are grouped) whereas others do not; moreover, within a group, some elements are accented, while others are not. There is substantial evidence that grouping affords a figural coding of rhythms that is distinct from the coding of metrical structure, and which plays an important role in listener ability to discriminate and remember rhythms (Handel 1998).

Figure 6.3 illustrates how various acoustic cues influence perceived grouping of elements of a rhythm. First, if every second or third element in a sequence is accented by increasing its intensity, then the elements tend to be perceived to be grouped into twos or threes, respectively, with the element of increased intensity beginning the group (Fig. 6.3a); moreover, the time interval between groups tends to be incorrectly perceived as longer than the time intervals separating elements within the group (Bolton 1894; Woodrow 1909).

Second, if the duration of every second or third element in an otherwise isochronous sequence is lengthened, then elements of the sequence are often perceived in

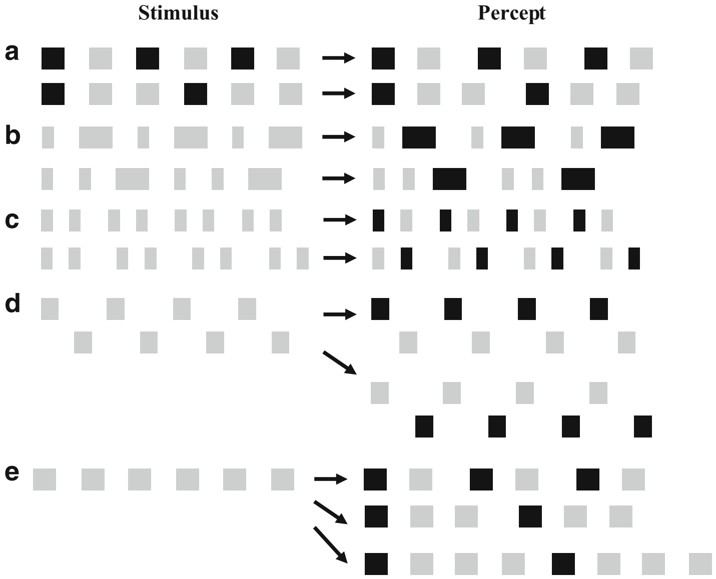

Fig. 6.3 Examples of principles of grouping. The left half of the figure shows a series of sound sequence examples (wider squares indicate lengthening of a musical element, while darker shading indicates an increase in sound intensity). The right half of the figure illustrates typical percepts for each example (darker shading indicates perceived accentuation of the musical element). (**a**) The intensity of every second or third element in an isochronous sequence is increased, leading to grouping by twos or threes, respectively. The more intense (louder) tone begins each group with the perceived time interval between groups longer than the perceived time intervals separating elements within a group. (**b**) The duration of every second or third element in an isochronous sequence is lengthened, leading to grouping by twos or threes, respectively. The longer element ends each group with the perceived time interval between groups longer than the perceived time intervals separating elements within a group. (**c**) Every other time interval between tone onsets is lengthened, leading to binary grouping. For differences between adjacent time intervals that are relatively small, the first element of each group is perceived as accented, whereas for differences that are relatively large, the second tone of each group is perceived as accented. (**d**) Alternating patterns of high and low tones lead to binary grouping with either the high or low tone perceived as accented and beginning the group. (**e**) Common subjective rhythms for isochronous sequences of identical elements

groups of two or three, with lengthened elements perceived as accented (Woodrow 1909); here as end accents of each group (Fig. 6.3b). As with intensity accentuation, the time intervals in between groups tend to be subjectively longer than the time intervals separating elements within a group (Handel 1989; Woodrow 1951).

Third, grouping of elements of a rhythm is also affected through variation in the time intervals between element onsets (i.e., holding constant element durations). Povel and Okkerman (1981) showed that when every second or third inter-onset interval is lengthened, then individuals tend to hear the groups of two or threes, respectively. In this case, when the difference between two successive intervals is relatively small, the first element of each group tends to be perceived as accented, whereas when this difference is relatively large, the final element of each group tends to be perceived as accented (Fig. 6.3c); see also the classic work of Garner (1974).

Fourth, frequency (pitch) cues also affect grouping (Fig. 6.3d). For pitch cues, there is a tendency to perceive the rhythmic organization of sequences according to repeated pitch patterning (Woodrow 1911; Steedman 1977). For example, when individuals listen to an isochronous sequence of tones of equal amplitude and duration that alternate between a fixed high and fixed low frequency (e.g., HLHLHLHL), they tend to hear a binary grouping of elements with accents on either the high or low tone, with the accented element beginning the group (Woodrow 1909, 1911).

Finally, some of the earliest work by Bolton (1894) on principles of grouping and the figural coding of rhythms found that even for isochronous sequences of identical sounds, listeners tend to perceive the elements of the sequences to be grouped in twos, threes, or fours, with the first element of each group judged to be more accented than the others (Fig. 6.3e). As the tempo of a rhythm increases, there is a tendency for the number of elements perceived to form a group to increase, suggesting that there may be an intrinsic preferred total duration for each group (Harrel 1937).

Much work has assumed that principles of grouping are universal. However, several studies have shown that performers use their knowledge of musical structure to emphasize the grouping of elements by increasing the intensity or duration of the final element in a group (Drake 1993; Drake and Palmer 1993; Repp et al. 2002). Moreover, at least one recent study has shown that perception of grouping can be dependent on cultural and linguistic experience of the listener (Iversen et al. 2008).

In sum, patterns of change in acoustic dimensions, such as frequency, duration, and intensity produce subjective accents on musical elements and influence how those elements are perceived to be grouped together in time, permitting a figural coding of rhythm. Grouping can also be imposed by the listener or emphasized by a performer, and is in addition impacted by the linguistic experience of the listener. In the absence of explicit acoustic cues, listeners tend to hear groups of twos, threes, and fours; preferred group size, however, is influenced by tempo. Section 6.5.2 turns to a discussion of central issues in research on beat and meter, including different types of accents, effects of tempo, and contributions of listener knowledge and experience.

6.5.2 Perception of Beat and Meter

Beat and meter have been studied using a variety of different empirical methods. This work has shown that metric coding of rhythms confers a number of advantages to listeners. Critically, metric coding provides information about the relative timing of elements of a rhythm and a basis for generating expectations about "when" in time future events will occur. Rhythms that can be described by a metric hierarchy (i.e., they afford a metric coding) have been shown to be more easily discriminated than rhythms that do not fit within a metric framework (Bharucha and Pryor 1986). Listeners also have more trouble discriminating pitches that occur at metrically weak locations than those that occur at metrically strong locations (Jones et al. 1982). Moreover, memory for the temporal position of a probe tone is better when probe tones are placed at strong metrical positions than at weak metrical positions in an implied metric hierarchy (Palmer and Krumhansl 1990). Finally, beat and meter facilitate judgments about time and pitch changes (Jones et al. 1982, 2002; Jones and Yee 1997; Barnes and Jones 2000; McAuley and Jones 2003) and enhance judgments of event durations (Boltz 1991, 1998) and melodic phrase completeness (Boltz 1989).

6.5.2.1 Contribution of Different Types of Accents

An important theoretical issue in work on beat and meter concerns the contribution of different types of accents and their timing to the communication of metrical structure to a listener (Cooper and Meyer 1960; Benjamin 1984). Two kinds of accents that have received substantial consideration in the literature are temporal accents and melodic accents. Temporal accents include pause accents and duration accents. Pause (or rhythmic) accents are produced by an empty time interval (marked by the onsets of two successive tones) that is relatively long compared to preceding inter-onset intervals (Povel and Okkerman 1981; Narmour 1996; Jones 1987; Jones and Pfordresher 1997). Duration accents are accents that occur on tones with a relatively long duration compared to the duration of preceding tones (Woodrow 1951; Handel 1989).

There are several varieties of melodic accents (Jones 1993; Hannon et al. 2004). Interval accents are created on a tone when the tone is much higher or lower in pitch that the surrounding events (Lerdahl and Jackendoff 1983; Huron and Royal 1996). With an interval accent, it is the element after the pitch jump (from high-to-low or low-to-high) that is accented (Jones 1981, 1993). Contour accents occur on tones at the point of change in a musical contour (e.g., the middle tone in three-tone melody than ascends and then descends in pitch); these points have been called contour pivot points or turnaround points (Thomassen 1982). It is not surprising that interval and contour accents frequently overlap because turnaround points often involve large pitch leaps. A third instance of melodic accent is a tonal accent (Smith and Cuddy 1989; Dawe et al. 1993). Tonal accents arise from a shift in tonal stability (e.g., from a leading tone to the tonic) within a particular musical context.

One issue that has received increased attention recently concerns the relative weight that listeners place on melodic and temporal accents in inferring metrical structure (Ellis and Jones 2009). Although there is consistent support that temporal accents are very important for metric coding (Povel and Essens 1985; Large and Jones 1999), data on the importance of melodic accents for perceiving metrical structure are mixed (Huron and Royal 1996). On the one hand, a number of studies have shown relatively little contribution of melodic accents to meter. Snyder and Krumhansl (2001) showed that when participants were asked to tap to pitch-varied and monotone versions of ragtime piano music, there was very little difference in tapping performance for the two conditions. Woodrow (1911) showed that increases in loudness and duration, but not changes in pitch influence the perceived beginning of a group of elements. Metrical stability ratings for events in melodies interrupted at various points show larger effects for temporal accents than for pitch accents (Bigand 1997). Finally, in expressive musical performance, pitch accents tend to be less consistent than temporal accents and highly context dependent (Drake and Palmer 1993).

On the other hand, there is also empirical evidence that both melodic accents and temporal accents contribute significantly to the perception of metrical structure; see Ellis and Jones (2009) for a comprehensive review. Much of this work highlights the importance of the temporal alignment of the two types of accents. This view is most comprehensively expressed by entrainment-based approaches, such as Dynamic Attending Theory (DAT; Jones 1976; Jones and Boltz 1989; Large and Jones 1999). In DAT, the periodic timing of temporal accents and melodic accents is assumed to drive entrainment and contribute to the emergence of what Jones and colleagues refer to as joint accent structure. A key prediction of this theory is that coincident (concordant) melodic and temporal accents (corresponding to a simple joint accent structure) should lead more efficient entrainment and a stronger perception of metrical structure than conflicting (discordant) accent timing (corresponding to a complex joint accent structure). Empirical support for a joint accent structure (JAS) hypothesis has been found in a number of studies using a variety tasks. In general, melodies with a concordant JAS have been shown to produce perceptual and performance advantages over melodies with a discordant JAS (Deutsch 1980; Boltz and Jones 1986; Dowling et al. 1987; Monahan et al. 1987; Boltz 1989, 1991; Drake et al. 1991; Jones et al. 1993; Jones and Pfordresher 1997; Pfordresher 2003).

Overall, conflicting results concerning the importance of melodic accents to meter perception raise the obvious question of why findings across studies have been so inconsistent. One possible reason is a failure of most studies to control for accent salience (Huron and Royal 1996; Snyder and Krumhansl 2001; Temperley and Bartlette 2002; Toiviainen and Snyder 2003). Thus, in cases where melodic accents are likely to be less salient than temporal accents, it is perhaps not surprising that the data favor temporal accents with melodic accents contributing weakly or not at all to perceived meter. This suggests that the relative contribution of melodic and temporal accents can be carefully assessed only when melodic and temporal accents are equated for their salience. In one of the only studies to explicitly equate melodic and temporal salience, Ellis and Jones (2009) provide conclusive evidence that melodic accents do contribute to meter perception and that metrical clarity ratings are greater for melodies with a concordant JAS than for melodies with a discordant JAS. Moreover, metrical clarity ratings are found to increase and reaction times decrease as the number of temporally aligned (coincident) accents increases.

6.5.2.2 Role of Tempo in Perception of Metrical Structure

As with many aspects of rhythm perception, tempo plays an important supporting role. This is particularly true for the perception of beat and meter, where the tempo of a rhythm can influence how listeners organize the elements of the rhythm both in terms of figural coding and in terms of metric coding (London 2004). The focus of this section is on the role of tempo in metric coding. One common empirical method for studying effects of tempo on perceived beat and meter requires listeners to synchronize taps with various rhythms presented at different tempi. A key dependent measure in studies using this method is the level of the implied metric hierarchy that listeners decide to tap.

In an influential set of studies on this topic, Handel and colleagues (Oshinsky and Handel 1978; Handel and Oshinsky 1981; Handel and Lawson 1983; Handel 1984) had listeners tap in synchrony with what they perceived to be the most natural placement of accents for a variety of constructed polyrhythms. Polyrhythms pit one isochronous rhythm against another, and so naturally create several possible rhythm interpretations of the emerging pattern. For example, for a 3 × 4 polyrhythm, Oshinsky and Handel (1978) observed three different tapping responses depending on tempo: (1) Listeners subdivided the pattern into three equal time intervals, tapping in synchrony with the three-element sequence; (2) they subdivided the pattern into four equal time intervals, tapping in synchrony with the four-element sequence; or (3) they tapped once every 12 elements when the two sequences coincided. Overall, listeners were more likely to subdivide into three equal time intervals at fast tempi than at slow tempi. Handel and colleagues have reported similar effects for other polyrhythms.

Related research has shown that for isochronous rhythms, individuals tend to tap beats every two, three, or four elements in a tempo-dependent fashion (Duke 1989); this finding is reminiscent of work by Bolton (1894) on grouping, but differs in that listeners are not asked to explicitly group elements, but rather to tap what they perceived to be the beat of the rhythm. Still other work has shown that for both simple and complex meters characteristic of Western music, there is a tendency to tap out higher time levels in an implied metric hierarchy at fast tempos than at slow tempos (Duke 1989; Parncutt 1994; London 2002, 2004). In many cases at very fast tempos, there is a tendency for individuals to tap once with each repetition of the rhythm (the highest level in the hierarchy; McAuley and Semple 1999).

In sum, for variety of rhythms, the most salient time level (i.e., the perceived beat) for listeners tends to be at a higher level in the metric hierarchy at fast tempi than at slow tempi. One interpretation of the reported effects of tempo on perceived beat and meter is that there is an interaction between preferred tempo and metrical structure. As the tempo of a rhythm changes, the relative time level that is nearest to an absolute preferred time interval also changes; in turn, this affects the relative salience of each level of the hierarchy. Having considered the role of different type of accents and the contributions of tempo to beat and meter, the discussion turns to the role of knowledge and experience in the next section.

6.5.2.3 Role of Knowledge and Experience in Perception of Metrical Structure

Elements of beat and meter perception are present relatively early in infancy (Hannon and Johnson 2005). Nonetheless, perception of beat and meter is also subject to learning and effects of enculturation (Hannon and Trehub 2005); see Trainor and Corrigall, Chap. 4, for a developmental perspective on these issues. In adulthood, questions about the role of knowledge in the perception of metrical structure have centered on the type of knowledge that musically trained and untrained individuals bring to bear when listening to music. That is, to what extent does meter have a psychological basis and do both trained and untrained listeners apply their knowledge of meter in everyday musical listening situations?

In a seminal study on this topic, Palmer and Krumhansl (1990) provided several sources of evidence that support mental representations for metrical structure and its use in perception. First, analysis of the frequency distribution of notes in a corpus of musical excerpts revealed statistical regularities in the temporal distribution of musical events that permitted accurate identification of the meter for at least the selected subset of meters examined in the study. Second and more important, in the absence of any acoustical cues to meter, listener judgments about the goodness-of-fit of the timing of probe tones placed at different metrical positions paralleled the pattern of statistical regularities observed in the corpus analyses. Third, effects of musical training were also evident with musicians showing finer-grained representations of metrical structure than did nonmusicians. Finally, listeners' memory for the temporal placement of a probe tone was influenced by its position in the implied metrical hierarchy with better memory for probes at strong metrical positions than at weak metrical positions. Support for the mental representation of musical meter is also evident in expressive musical performance. For example, pianists use their knowledge of the noted meter to accent events at strong metrical positions by making them longer, louder, or more legato (Shaffer 1981; Sloboda 1983; Drake and Palmer 1993).

Having now identified key theoretical issues and empirical findings in research on musical rhythm, including work on grouping, beat, and meter, a final section on rhythm provides a more detailed consideration of several influential modeling approaches.

6.5.3 Models of Rhythm

An important distinction made in the development of various models of rhythm is between the processing of rhythms that permit a metric coding and those that only afford a figural coding (Essens and Povel 1985; Povel and Essens 1985; Hébert and Cuddy 2002). Povel and Essens (1985) refer to the former metrical temporal patterns and to the later as nonmetrical temporal patterns. Much of the work on modeling rhythm has focused on the induction of a beat and the perception of metrical structure (i.e., metric coding). These models range from algorithmic rule-based approaches (e.g., Longuet-Higgens and Lee 1982; Povel and Essens 1985; Desain and Honing 1999) to dynamical systems accounts (e.g., Large and Kolen 1994; Large and Jones 1999; Eck 2002).

In general, rule-based models of rhythm share similarities with interval theories of tempo. Specifically, as with tempo, rule-based models of rhythm tend to derive from an information processing perspective with the common assumption that people perceive, remember, and reproduce rhythms by structuring their mental representation according to an internal clock. As with tempo models, this internal clock is assumed to involve a pacemaker-accumulator mechanism that ticks out regular intervals that are aligned with particular stimulus onsets that correspond to induced beats.

One particularly influential rule-based approach is that of Povel and Essens (1985). These authors proposed what amounts to a three-stage clock model. First, subjective accents were assigned to the rhythm according to a set of empirically derived preference rules (Povel and Okkerman 1981). The emphasis of these preference rules was on temporal accents. Accents were assumed to occur on (a) temporally isolated tones, (b) the second in a group of two tones, and (c) the first and last tone in a run or three or more elements (Povel and Okkerman 1981; Povel and Essens 1985). Second, all possible clock intervals were generated in an algorithmic fashion, allowing some restrictions on what would be considered viable responses. Finally, in a "matching" stage, the amount of counter (negative) evidence was calculated for each potential clock and the clock with the least counter-evidence was determined to be the most likely induced beat. Negative evidence consisted of clock pulses (beats) falling on unaccented elements or silent elements of a sequence, with silent elements contributing more negative evidence than unaccented elements. In support of their model, Povel and Essens found that rhythms with less negative evidence were reproduced more accurately and judged to be simpler than rhythms with more negative evidence (Povel and Essens 1985; Essens and Povel 1985).

McAuley and Semple (1999) generalized the Povel and Essens (1985) clock model approach by in addition considering models that determined the best clock (beat) according to a positive evidence heuristic, as well as a hybrid model that combined in a weighted fashion both positive and negative (counter-evidence) evidence heuristics. The use of positive evidence as a heuristic has been similarly examined by Parncutt (1994), who also incorporated into his model the concept of pulse salience in an attempt to address effects of tempo. Because the three models differ in the nature of the evidence used to determine the strength of the best-fitting clock (beat), they each have different biases. Clock models involving negative evidence heuristics tend to favor slow clocks because slow clocks afford less opportunity to accumulate negative evidence. In contrast, clock models involving positive evidence heuristics tend to favor fast clocks because fast clocks afford more opportunity to accumulate positive evidence. Hybrid models that combine positive and negative evidence in a weighted fashion tend to offer a balance between these two extremes and favor clocks that are neither too fast nor too slow.

In a comparison of these three models (positive evidence, negative evidence, hybrid), McAuley and Semple (1999) found that models based solely on a negative-evidence heuristic best predicted the perceived beat of nonmusicians, whereas models incorporating a positive-evidence heuristic best predicted the perceived beat of musicians. In addition, tempo tended to affect the type of evidence that best predicted the perceived beat. A negative-evidence heuristic worked best for fast tempi because it favored, like listeners, beats at higher levels of the metric hierarchy. In contrast, a positive-evidence heuristic worked best for slow tempi because it favored, like listeners, beats at lower levels of the metric hierarchy. The tempo shift favoring negative-evidence heuristics at fast tempi and positive-evidence heuristics a slow tempi reflects listeners tendency to prefer to tap beats at a level within a metric hierarchy that it is at an absolute tempo that is neither too fast nor too slow, but rather falls at an intermediate rate. This intermediate rate is in the range of commonly reported preferred tempi.

Three general weaknesses of the different varieties of the internal clock approach to rhythm are that these models: (1) describe, but fail to explain why various factors such as tempo and musical experience alter the perception of a beat; (2) do not operate in real time, but rather consider all possible clocks in an algorithmic fashion before settling on a solution; and (3) rely primarily on temporal accents to infer beat and meter. In addition to the important role that different types of accents play, another somewhat overlooked factor is the role of repetition. Both the repetition of melodic fragments and rhythmic patterns contribute to the perception of metrical structure (Steedman 1977; Temperley 2001). See research by Steedman (1977) and Temperley and Bartlette (2002) for examples of models incorporating a principle of repetition (or parallelism); these models have been used to recover the scored meter of a piece of music with modest success.

A variety of models have been proposed that do operate in real time, many arising from an entrainment perspective. Most entail hypotheses about self-sustaining oscillations that respond to both the timing and intensity of event onsets and facilitate the extraction of a musical pulse, which is then used for real-time beat tracking (Large and Kolen 1994; McAuley 1995; Toiviainen 1998; Eck 2002; Large and Palmer 2002). The perception of meter from an entrainment perspective requires a multiple-oscillator model in which each oscillator corresponds to each metrical level. One proposal along these lines is for the relative salience of each metrical level to be modeled by the resonance region of an oscillator that is centered on preferred tempo (van Noorden and Moelants 1999).

Overall, entrainment-based approaches to beat and meter offer advantages over internal clock approaches in that they operate in real time and have the potential for providing a more comprehensive explanation of how factors such as tempo influence beat perception and meter (Large and Kolen 1994; McAuley 1995; Large and Jones 1999). As with interval models, entrainment approaches are similarly limited in their reliance primarily on temporal accents to infer beat and meter.

6.6 Summary

Tempo and rhythm are two fundamental and related aspects of musical communication. Tempo refers to pace of a piece of music (i.e., how fast or slow it is) and is typically measured in beats per minute or by the beat-onset interval. The concept of rhythm has been less straightforward to define and can be viewed either as a serial pattern of durations or the perceived temporal organization of that pattern. In the second sense of the term, rhythm as a perception has several elements including grouping, beat, and meter.

Theoretical approaches to tempo and rhythm can, for the most part, be classified as either interval theories or entrainment (beat-based) theories. Key differences between the two approaches concern (1) the nature of the clock's responses to stimulus onsets (arbitrary reset via a switch in interval theories vs. more gradual phase and period correction processes in entrainment theories), (2) representation of duration (duration code stored in memory in interval theories vs. oscillator period in entrainment theories), (3) successive time estimates (independent vs. dependent), and (4) the nature of duration judgments (explicit comparison of two stored duration codes in interval theories vs. temporal contrast metric in entrainment theories).

Perceptual and motor preferences for particular tempi emerge at a young age. In adulthood, representative values for preferred perceptual tempo and spontaneous motor tempo are around 600 ms, although correlations between the two are modest. Across the life span, there is converging support for the view that preferred tempo slows with increased age. There is also support for the view that listeners develop absolute memory for tempi of music that is experienced at the same tempo (i.e., the music has a tempo standard). There is little support, however, for differences in preferred tempo associated with gender, handedness, of for clear associations with specific physiological measures. Experience does appear to play a role as musicians tend to have slower preferred tempi than nonmusicians, especially in childhood.

Studies of tempo discrimination and tempo production reveal violations in Weber's law consistent with the concept of a preferred tempo. With respect to perception, JNDs for tempo tend to be a minimum in a range of tempi (optimal tempo region) that is centered on 600 ms, with thresholds for single interval sequences around 6% and those for multiple-interval isochronous sequences around 2% in ideal listening conditions. Weber fractions for variability in tempo productions tend to mirror threshold results observed for perception. There is also converging evidence both musicians and nonmusicians develop fairly precise memories for the absolute tempo of music, as long as the music has a tempo standard

Research on rhythm distinguishes the perception of grouping (figural coding) from the perception of beat and meter (metric coding). Some of the earliest work on rhythm showed that for isochronous sequences of identical sounds, listeners tend to perceive the elements of the sequences to be grouped in twos, threes, or fours, with the first element of each group judged to be more accented than the others. Acoustic factors affecting perceived grouping (i.e., figural coding of a rhythm) include intensity, duration, and frequency patterning. The topic of beat and meter has been studied using a variety of different methods, with the general finding that musical events that occur at strong metrical positions are better perceived, remember, and reproduced than musical events that occur at weak metrical positions. Both stimulus-based accents of different types and musical knowledge contribute to perceived beat and meter (i.e., metric coding of a rhythm). There is consistent support that the timing of temporal accents is important for the perception of metrical structure, but the role of melodic accents in metric coding is less clear cut. However, many studies revealing weak or negligible effects of melodic accents on perceived beat and/or meter have failed to control for accent salience.

Overall, interval and entrainment models of tempo and rhythm have met with mixed success. Interval models of tempo, such as the multiple-look and generalized multiple-look models have been quite successful in describing discrimination thresholds for tasks requiring the discrimination of isolated time intervals or isochronous sequences. However, these approaches have been less successful when applied to more complex rhythms. Moreover, because the pacemaker-accumulator conception of an internal clock passively records time as the number of ticks, interval models, in general, are agnostic about the concept of a preferred tempo. Entrainment models offer an alternative approach, which addresses both findings on preferred tempo and tempo discrimination.

One outstanding challenge of many models of rhythm is to address the finding that the perceived grouping, beat and meter (and hence figural and metric coding) vary with tempo, as well as vary from listener to listener. Another issue that has been difficult to address in models is the nature of interactions between pitch and time cues in perception of grouping, beat, and meter. Most models of rhythm have focused on the relative timing of temporal accents; relatively few, in contrast, address the role of melodic accents. Dynamic attending theory developed within an entrainment framework is one approach that shows promise for capturing the rich interactions between melodic and temporal accents in the perception of rhythm.

Finally, a generally understudied area of research is individual differences. With respect to rhythm, there are numerous anecdotal reports that suggest that there are large individual differences in the ability to perceive a beat; that is, some people appear to have much more difficulty perceiving a beat than others. The nature of these individual differences in adulthood is only beginning to be addressed (Iversen and Patel 2008; Grahn and McAuley 2009). The importance of this line of investigation is highlighted by recent evidence supporting a link between rhythmic ability and language processing (Alcock et al. 2000; Thompson and Goswami 2008) and evidence of rhythm perception deficits in neurological disorders, such as Parkinson's disease (Grahn and Brett 2009).

Acknowledgements Many thanks to Mari Riess Jones, Laura Dilley, Molly Henry, and Nathan Miller for detailed comments on an earlier version of this chapter. Thanks also to Laura Dilley for help with figures. The author received support for this work from the National Science Foundation (BCS 0818271).

References

Abel SM (1972) Discrimination of temporal gaps. J Acoust Soc Am 52:519–524.

Alcock KJ, Passingham RE, Watkins KE, Vargha-Khadem F (2000) Pitch and timing abilities in inherited speech and language impairment. Brain Lang 75:34–46.

Allan LG (1979) The perception of time. Percept Psychophys 26:340–354.

Andrews MW, Dowling WJ, Bartlett JC, Halpern AR (1998) Identification of speeded and slowed familiar melodies by younger, middle-aged, and older musicians and nonmusicians. Psychol Aging 13:462–471.

Bamberger J (1980) Cognitive structuring in the apprehension and description of simple rhythms. Arch Psychol 48:171–199.

Barnes R, Jones MR (2000) Expectancy, attention, and time. Cogn Psychol 41:254–311.

Benjamin WE (1984) A theory of musical meter. Music Percept 1:355–413.

Bergeson T, Trehub S (2002) Absolute pitch and tempo in mothers' songs to infants. Psychol Sci 13:72–75.

Bharucha J, Pryor JH (1986) Disrupting the isochrony underlying rhythm: an asymmetry in discrimination. Percept Psychophys 40:137–141.

Bigand E (1997) Perceiving musical stability: the effect of tonal structure, rhythm, and musical expertise. J Exp Psychol: Hum Percept Perform 23:808–822.

Bolton TL (1894) Rhythm. Am J Psychol 6:145–238.

Boltz M (1989) Rhythm and "good endings": effects of temporal structure on tonality judgments. Percept Psychophys 46:9–17.

Boltz M (1991) Some structural determinants of melody recall. Mem Cogn 19:239–251.

Boltz M (1994) Changes in internal tempo and effects on the learning and remembering of event durations. J Exp Psychol: Learning Mem Cogn 20:1154–1171.

Boltz M (1998) The processing of temporal and nontemporal information in the remembering of event durations and musical structure. J Exp Psychol: Hum Percept Perform 24:1087–1104.

Boltz M, Jones MR (1986) Does rule recursion make melodies easier to reproduce? If not, what does? Cogn Psychol 18:389–431.

Church RM (2003) A concise introduction to scalar timing theory. In Meck WH (ed), Functional and Neural Mechanisms of Interval Timing. Boca Raton, FL: CRC Press.

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

Collyer CE, Broadbent HA, Church, RM (1994) Preferred rates of repetitive tapping and categorical time production. Percept Psychophys 55:443–453.

Cooper GW, Meyer LB (1960) The Rhythmic Structure of Music. Chicago, IL: University of Chicago Press.

Creelman CD (1962) Human discrimination of auditory duration. J Acoust Soc Am 34:582–593.

Crocker RL (2000) An Introduction to Gregorian Chant. New Haven, CT: Yale University Press.

Dalla Bella S, Peretz I, Rousseau L, Gosselin N (2001) A developmental study of the affective value of tempo and mode in music. Cognition 80:1–10.

Dawe LA, Platt JR, Racine RJ (1993) Harmonic accents in inference of metrical structure and perception of rhythm patterns. Percept Psychophys 54:794–807.

Desain P, Honing H (1999) Computational models of beat induction: the rule-based approach. J New Music Res 28:29–42.

Deutsch D (1980) The processing of structured and unstructured tonal sequences. Percept Psychophys 28:381–389.

Dowling WJ, Lung KM, Herrbold S (1987) Aiming attention in pitch and time in the perception of interleaved melodies. Percept Psychophys 41:642–656.

Drake C (1993) Perceptual and performed accents in musical sequences. Bull Psychonomic Soc 31:107–110.

Drake C, Botte MC (1993) Tempo sensitivity in auditory sequences: evidence for a multiple-look model. Percept Psychophys 54:277–286.

Drake C, Palmer C (1993) Accent structure in music performance. Music Percept 10:343–378.

Drake C, Dowling WJ, Palmer C (1991) Accent structures in the reproduction of simple tunes by children and adult pianists. Music Percept 8:315–334.

Drake C, Jones MR, Baruch C (2000) The development of rhythmic attending in auditory sequences: attunement, referent period, focal attending. Cognition 77:251–288.

Duchek, JM, Balota, DA, Ferraro FR (1994) Component analysis of a rhythmic finger tapping task in individuals with senile dementia of the Alzheimer type and in individuals with Parkinson's disease. Neuropsychology 8:218–226.

Duke RA (1989) Musicians' perception of beat in monotonic stimuli. J Res Music Ed 37:61–71.

Eck D (2002) Finding downbeats with a relaxation oscillator. Psychol Res 66:18–25.

Ellis RJ, Jones MR (2009) The role of accent salience and joint accent structure in meter perception. J Exp Psychol: Hum Percept Perform 35:264–280.

Essens PJ, Povel DJ (1985) Metrical and nonmetrical representations of temporal patterns. Percept Psychophys 37:1–7.

Farnsworth P, Block H, Waterman W (1934) Absolute tempo. J Gen Psychol 10:230–233.

Fraisse P (1956) Les structures rythmiques. Louvain, Belgium: Publications Universitaires de Louvain.

Fraisse P (1963) The Psychology of Time. New York: Harper and Row.

Fraisse P (1982) Rhythm and tempo. In Deutsch D (ed), The Psychology of Music. Orlando, FL: Academic, pp. 149–180.

Friberg A, Sundberg, J (1995) Time discrimination in a monotonic, isochronous sequence. J Acoust Soc Am 98:2524–2531.

Friberg A, Sundström (2002) Swing ratios and ensemble timing in jazz performance: evidence for a common rhythmic pattern. Music Percept 19:333–349.

Frischeisen-Köhler, I (1933) The personal tempo and its inheritance. Character Pers 1:301–313.

Gagnon L, Peretz I (2003) Mode and tempo relative contributions to "happy-sad" judgements in equitone melodies Cogn Emotion 17:25–40.

Garner WR (1974) The Processing of Information and Structure. Oxford: Lawrence Erlbaum.

Getty DJ (1975) Discrimination of short temporal intervals: a comparison of two models. Percept Psychophys 18:1– 8.

Geuze RH, Kalverboer, AF (1994) Tapping a rhythm: a problem of timing for children who are clumsy and dyslexic? Adapt Phys Act Q 11:203–213.

Gibbon J (1977) Scalar expectancy theory and Weber's law in animal timing. Psychol Rev 84:279–325.

Gibbon J, Church RM, Meck WH (1984) Scalar timing in memory. Ann NY Acad Sci 423:52–77.

Grahn JA, Brett M (2009) Impairment of beat-based rhythm discrimination in Parkinson's disease. Cortex 45:54–61.

Grahn JA, McAuley JD (2009) Neural bases of individual differences in beat perception. NeuroImage 47:1894–1903.

Greene LS, Williams, HG (1993) Age-related differences in timing control of repetitive movement: application of the Wing-Kristofferson model. Res Q Exerc Sport 64:32–38.

Grondin S (2001a) Discriminating time intervals presented in sequences marked by visual signals. Percept Psychophys 63:1214–1228.

Grondin S (2001b) From physical time to the first and second moments of psychological time. Psychol Bull 47:22–44.

Grondin S, McAuley JD (2009) Duration discrimination in crossmodal sequences. Perception 38:1542–1559.

Halpern AR (1988) Perceived and imagined tempos of familiar songs. Music Percept 6: 193–202.

Handel S (1984) Using polyrhythms to study rhythm. Music Percept 1:465–484.

Handel S (1989) Listening: An Introduction to the Perception of Auditory Events. Cambridge, MA: MIT Press.

Handel S (1998) The interplay between metric and figural rhythmic organization. J Exp Psychol: Hum Percept Perform 24:1546–1561.

Handel S, Lawson GR (1983) The contextual nature of rhythmic interpretation. Percept Psychophys 34:103–120.

Handel S, Oshinsky, JS (1981) The meter of syncopated auditory polyrhythms. Percept Psychophys 30:1–9.

Hannon EE, Johnson SP (2005) Infants use meter to categorize rhythms and melodies: implications for musical structure learning. Cogn Psychol 50:354–377.

Hannon EE, Trehub SE (2005) Metrical categories in infancy and adulthood. Psychol Sci 16:48–55.

Hannon EE, Snyder JS, Eerola T, Krumhansl CL (2004) The role of melodic and temporal cues in perceiving musical meter. J Exp Psychol: Hum Percept Perform 30:956–974.

Harrel TW (1937) Factors influencing preference and memory for auditory rhythm. J Gen Psychol 17:63–104.

Hébert S, Cuddy LL (2002) Detection of metric structure in auditory figural patterns. Percept Psychophys 64:909–918.

Hirsh IJ, Monahan CB, Grant KW, Singh PG (1990) Studies in auditory timing: 1. Simple patterns. Percept Psychophys 47:215–226.

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

Iversen JR, Patel AD (2008) The beat alignment test (BAT): surveying beat processing abilities in the general population. In Adachi et al. (eds), Proceedings of the 10th International Conference on Music Perception and Cognition, August 25–29. Adelaide: Causal Productions.

Iversen JR, Patel AD, Ohgushi K (2008) Perception of rhythmic grouping depends on auditory experience. J Acoust Soc Am 124:2263–2271.

Ivry RB, Hazeltine RE (1995) Perception and production of temporal intervals across a range of durations: evidence for a common timing mechanism. J Exp Psychol: Hum Percept Perform 21:3–18.

Ivry RB, Keele SW (1989) Timing functions of the cerebellum. J Cogn Neurosci 6:136–152.

James W (1890) The Principles of Psychology, Vol. I. New York: Holt.

Jones MR (1976) Time, our lost dimension: toward a new theory of perception, attention, and memory. Psychol Rev 83:323–355.

Jones MR (1981) A tutorial on some issues and methods in serial pattern research. Percept Psychophys 30:492–504.

Jones MR (1987) Dynamic pattern structure in music: recent theory and research. Percept Psychophys 41:621–634.

Jones MR (1993) Dynamics of musical patterns: how do melody and rhythm fit together? In Tighe TJ, Dowling WJ (eds), Psychology and Music: The Understanding of Melody and Rhythm. Hillsdale, NJ: Lawrence Erlbaum, pp. 67–92.

Jones MR, Boltz M (1989) Dynamic attending and responses to time. Psychol Rev 96:459–491.

Jones MR, McAuley JD (2005) Time judgments in global temporal contexts. Percept Psychophys 67:398–417.

Jones MR, Pfordresher PQ (1997) Tracking melodic events using joint accent structure. Can J Exp Psychol 51:271–291.

Jones MR, Ralston JT (1991) Some influences of accent structure on melody recognition. Mem Cogn 19:8–20.

Jones MR, Yee W (1997) Sensitivity to time change: the role of context and skill. J Exp Psychol: Hum Percept Perform 23:693–709.

Jones MR, Kidd GR, Wetzel R (1981) Evidence for rhythmic attention. J Exp Psychol: Hum Percept Perform 7:1059–1073.

Jones MR, Boltz M, Kidd GR (1982) Controlled attending as a function of melodic and temporal context. Percept Psychophys 32:211–218.

Jones MR, Boltz M, Klein JM (1993) Expected endings and judged duration. Mem Cogn 21:646–665.

Jones MR, Moynihan H, Mackenzie N, Puente J (2002) Temporal aspects of stimulus-driven attending in dynamic arrays. Psychol Sci 13:313–319.

Keele SW, Nicoletti R, Ivry RI, Pokorny RA (1989) Mechanisms of perceptual timing: beat-based or interval-based judgments? Psychol Res 50:221–256.

Klein JM, Jones MR (1996) Effects of attentional set and rhythmic complexity on attending. Percept Psychophys 58:34–46.

Krampe R, Mayr U, Kliegl R (2005) Timing, sequencing, and executive control in repetitive movement production. J Exp Psychol: Hum Percept Perform 31:379–397.

Large EW, Jones MR (1999) The dynamics of attending: how people track time-varying events. Psychol Rev 107:119–159.

Large EW, Kolen JF 1994. Resonance and the perception of musical meter. Connection Sci 6:177–208.

Large EW, Palmer C (2002) Perceiving temporal regularity in music. Cogn Sci 26:1–37.

Large EW, Fink P, Kelso JAS (2002) Tracking simple and complex sequences. Psychol Res 66:3–17.

Lerdahl F, Jackendoff R (1983) A Generative Theory of Tonal Music. Cambridge, MA: MIT Press.

Levitin DJ, Cook PR (1996) Memory for musical tempo: additional evidence that auditory memory is absolute. Percept Psychophys 58:927–935.

London J (2002) Cognitive constraints on metric systems: some observations and hypotheses. Music Percept 19:529–550.

London J (2004) Hearing in Time: Psychological Aspects of Musical Meter. New York: Oxford University Press.

Longuet-Higgens C, Lee C (1982) The perception of musical rhythms. Perception 11:115–128.

Mattel MS, Meck WH (2000) Neuropsychological mechanisms of interval timing behavior. BioEssays 22:94–103.

McAuley JD (1995) Perception of time as phase: toward an adaptive-oscillator model of rhythmic pattern processing. Unpublished doctoral dissertation, Indiana University, Bloomington.

McAuley JD, Jones MR (2003) Modeling effects of rhythmic context on perceived duration: a comparison of interval and entrainment approaches to short-interval timing. J Exp Psychol: Hum Percept Perform 29:1102–1125.

McAuley JD, Kidd GR (1998) Effect of deviations from temporal expectations on tempo discrimination of isochronous tone sequences. J Exp Psychol: Hum Percept Perform 24:1786–1800.

McAuley JD, Semple P (1999) The effect of tempo and musical experience on perceived beat. Aust J Psychol 51:176–187.

McAuley JD, Jones MR, Holub S, Johnston HM, Miller NS (2006) The time of our lives: lifespan development of timing and event tracking. J Exp Psychol: Gen 135:348–367.

Meck WH (1996) Neuropharmacology of timing and time perception. Cogn Brain Res 3:227–242.

Meck WH (2003) Functional and Neural Mechanisms of Interval Timing. Boca Raton, FL: CRC Press.

Meck WH (2005) Neuropsychology of timing and time perception. Brain Cogn 58:1–8.

Michon JA (1964) Studies on subjective duration: I. Differential sensitivity in the perception of repeated temporal intervals. Acta Psychol 22:441–450.

Miller NS, McAuley JD (2005) Tempo sensitivity in isochronous tone sequences: the multiple-look model revisited. Percept Psychophys 67:1150–1160.

Mishima J (1956) On the factors of mental tempo. Japanese Psychol Res 4:27–38.

Monahan CB, Kendall RA, Carterette EC (1987) The effect of melodic and temporal contour on recognition memory for pitch change. Percept Psychophys 41:576–600.

Moore-Ede MC, Sulzman FM, Fuller CA (1982) The Clocks that Time Us: Physiology of the Circadian Timing System. Cambridge, MA: Harvard University Press.

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

Ogden RT, Collier GL (1999) On detecting and modeling deterministic drift in long run sequences of tapping data. Comm Stat Theory Methods 28:977–987.

Oshinsky JS, Handel S (1978) Syncopated auditory polyrhythms: discontinuous reversals in meter interpretation. J Acoust Soc Am 63:936–939.

Palmer C, Krumhansl CL (1990) Mental representations of musical meter. J Exp Psychol: Hum Percept Perform 16:728–741.

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

Pashler H (2001) Perception and production of brief durations: beat-based versus interval-based timing. J Exp Psychol: Hum Percept Perform 27:485–493.

Patel AD, Iversen JR, Chen Y, Repp BH (2005) The influence of metricality and modality on synchronization with a beat. Exp Brain Res 163:226–238.

Pfordresher, PQ (2003) The role of melodic and rhythmic accents in musical structure. Music Percept 20:431–464.

Pöppel E (1997) A hierarchical model of temporal perception. Trends Cogn Sci 1:56–61.

Povel DJ, Essens PJ (1985) Perception of temporal patterns. Music Percept 2:411–440.

Povel DJ, Okkerman H (1981) Accents in equitone sequences. Percept Psychophys 30:565–572.

Repp BH (2003) Rate limits in sensorimotor synchronization with auditory and visual sequences: the synchronization threshold and the benefits and costs of interval subdivision. J Mot Behav 35:355–370.

Repp BH (2005) Rate limits of on-beat and off-beat tapping with simple auditory rhythms: 2. The roles of different kinds of accent. Music Percept 23:165–187.

Repp BH, Windsor WL, Desain P (2002) Effects of tempo on the timing of simple musical rhythms. Music Percept 19:565–593.

Rimoldi HJA (1951) Personal tempo. J Abnorm Soc Psychol 46:280–303.

Ross J, Houtsma AJM (1994) Discrimination of auditory temporal patterns. Percept Psychophys 56:19–26.

Rousseau L, Rousseau R (1996) Stop-reaction time and the internal clock. Percept Psychophys 58:434–448.

Schulze HH (1978) The detectability of local and global displacements in regular rhythmic patterns. Psychol Res 40:173–181.

Schulze HH (1989) The perception of temporal deviations in isochronic patterns. Percept Psychophys 45:291–296.

Shaffer L (1981) Performances of Chopin, Bach and Bartok: studies in motor programming. Cogn Psychol 13:327–376.

Sloboda JA (1983) The communication of musical metre in piano performance. Q J Exp Psychol A 35:377–396.

Small AM, Campbell RA (1962) Temporal differential sensitivity for auditory stimuli. Am J Psychol 75:401–410.

Smith KC, Cuddy LL (1989) Effects of metric and harmonic rhythm on the detection of pitch alterations in melodic sequences. J Exp Psychol: Hum Percept Perform 15:457–471.

Smith KC, Cuddy LL, Upitis R (1994) Figural and metric understanding of rhythm. Psychol Music 22:117–135.

Smoll FL, Schutz RW (1978) Relationships among measures of preferred tempo and motor rhythm. Percept Mot Skills 46:883–94.

Snyder JS, Krumhansl CL (2001) Tapping to ragtime: cues to pulse finding. Music Percept 18:455–489.

Snyder JS, Hannon EE, Large EW, Christiansen MH (2006) Synchronization and continuation tapping to complex meters. Music Percept 24:135–146.

Steedman MJ (1977) The perception of musical rhythm and metre. Perception 6:555–569.

Stern W (1900) Das psychisch Tempo. In Uber psychologie der individuellen differenzen. Leipzig: Barth.

Stevens LT (1886) On the time sense. Mind 11:393–404.

Szelag E, von Steinbüchel N, Reiser M, Gilles de Langen E, Pöppel E (1996) Temporal constraints in processing of nonverbal rhythmic patterns. Acta Neurobiol Exp 56:215–225.

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 Percept 20:117–149.

ten Hoopen G, Akerboom S (1983) The subjective tempo difference between interaural and monaural sequences as a function of sequence length. Percept Psychophys 34:465–469.

ten Hoopen G, Boelaarts L, Gruisen A, Apon I, Donders K, Mul N, Akerboom S (1994) The detection of anisochrony in monaural and interaural sound sequences. Percept Psychophys 56:110–120.

Thomassen JM (1982) Melodic accent: experiments and a tentative model J Acoust. Soc Am 71:1596–1605.

Thompson JM, Goswami U (2008) Rhythmic processing in children with developmental dyslexia: auditory and motor rhythms link to reading and spelling. J Physiol Paris 102:120–129.

Toiviainen P (1998) An interactive MIDI accompanist. Comput Music J 22:63–75.

Toiviainen P, Snyder JS (2003) Tapping to Bach: resonance-based modeling of pulse. Music Percept 21:43–80.

van Noorden L, Moelants D (1999) Resonance in the perception of musical pulse. J Music Res 28:43–66.

Vanneste S, Pouthas V, Wearden J (2001) Temporal control of rhythmic performance: a comparison between young and old adults. Exp Aging Res 27:83–102.

Vierordt K (1868) Der Zeitsinn nuch Versuchen. Tubingen: H. Laupp.

Vos PG, van Assen M, Franek M (1997) Perceived tempo change is dependent on base tempo and direction of change: evidence for a generalized version of Schulze's (1978) internal beat model. Psychol Res 59:240–247.

Wallin J (1911) Experimental studies of rhythm and time. Psychol Rev 18:100–133.

Williams HG, Woollacott MH, Ivry RB (1992) Timing and motor control in clumsy children. J Mot Behav 24:165–172.

Wing A (1980) The long and short of timing in response sequences. In Stelmach GE, Requin J (eds), Tutorials in Motor Behavior. Amsterdam: North-Holland.

Wing AM, Kristofferson AB (1973) The timing of inter-response intervals. Percept Psychophys 13:455–460.

Woodrow H (1909) A quantitative study of rhythm. Arch Psychol 14:1–66.

Woodrow H (1911) The role of pitch in rhythm. Psychol Rev 18:54–77.

Woodrow H (1951) Time perception. In Stevens SS (ed), Handbook of Experimental Psychology. New York: John Wiley, pp. 1224–1236.

Wundt W (1874) Grudzüge der physiologischen psychologie. Leipzig: Engelmann.