

Three Things Linguists Need to Know About Rhythm and Time in Music

JUSTIN LONDON
Carleton College, USA

ABSTRACT: This paper, directed at researchers in linguistics, introduces three key aspects of musical rhythm and time for their consideration: (1) the distinction between groups of durations (i.e., acoustical events in the world) and our endogenous sense of beats and beat cycles, that is, musical meter; (2) the active nature of rhythmic perception and cognition, which involves both innate and enculturated responses to music, and (3) that musical rhythms involve temporal processes on different time scales (from 100ms to 5-7 seconds), though they are integrated into a coherent perceptual framework. In addition, the relationships between musical rhythm and sensorimotor entrainment, as well as some important differences between musical and linguistic rhythms, are discussed.

Submitted 2012 January 6; accepted 2012 July 13.

KEYWORDS: *meter, grouping, entrainment, time-scales, attention*

THREE things a linguist needs to know about rhythm and time in music are:

1. There is an important distinction between *meter*—our endogenous sense of rhythmic organization that arises in the perception of periodic stimuli—and *grouping*, which consists of phenomenal patterns of duration out in the world;
2. Our perception of musical rhythm is *active*, not passive, and involves both bottom-up and top-down processing of the musical sounds, usually with concomitant motor behavior;
3. Different time scales for musical rhythm (i.e., very short, short, moderate, and long durations/intervals) play by different perceptual and cognitive rules.

This paper will discuss each of these in turn, and then conclude with a few cautionary remarks about some important ways in which rhythm in language and music differ.

METER VERSUS GROUPING

Grouping (sometimes referred to as rhythmic grouping, or simply rhythm (Lerdahl & Jackendoff, 1983; London, 2001) involves the organization of durations, or more precisely, the intervals from sound onset to sound onset ("IOI" or "inter-onset interval"). Group boundaries can be marked by changes in loudness, timbre (spectral properties), pitch, and/or duration. *Meter* involves our endogenous sense of a cyclic pattern of beats, though meter typically involves additional levels of temporal structure. Beats are perceptual abstractions, characterized by peaks of attentional energy (Jones & Boltz, 1989; Large & Jones, 1999; London, 2012).

Meter is a product of our dynamic interaction with patterns of sound events in the world; as such, it is best understood a form of *entrainment*, our ability to synchronize a periodic aspect of our attention and/or behavior with rhythms present in the environment. To put it another way, meter, which is phenomenally manifest to us as our perception of a pattern of accentually differentiated beats, is a musically-specific form of entrainment. Entrainment is bound up with the linkages between our auditory and sensorimotor systems. Indeed, Repp (2007a) hypothesized that metrical interpretation rests upon covert sensorimotor action, and Repp's hypothesis is supported by research in neuroscience that has uncovered strong associations between the auditory and motor regions of the brain (Chen, Penhune, & Pastore, 2008; Grahn & Brett, 2007; Zatorre, Chen & Penhune, 2007). Entrainment also enhances our temporal sensitivities and judgments; we are better at noticing differences in duration, for example, in the context of a continuous series of beats than when we hear only an isolated tone or interval (Drake & Botte, 1993; Friberg & Sundberg, 1995; Repp, 2001). As metric entrainment involves the behavior of a dynamic system, often described in terms of coupled oscillators (Large & Kohlen, 1994; Large & Palmer, 2002), while one can talk about metrical "structure" one should remember that meter is not a musical structure, but as Gjerdingen (1989) aptly put it, "a mode of attending."

Both grouping and meter create and/or can be described in terms of temporal patterns, and indeed, in some cases the two may be perfectly congruent. There are important differences the two, however, summarized in Table 1:

Meter	Grouping
Endogenous Beats	Phenomenal Durations/IOIs
Beats organized into cyclical and stable measures; pattern repeats exactly and consistently	Durations/IOIs organized into groups which may or may repeat
Measures are continuous; no gaps within a measure	Groups are continuous; no gaps within a group
Measures are contiguous; no gaps between measures	Groups need not be contiguous; often gaps between groups
Measures involve at least two beats to establish a repeating cycle	A group may consist of a single duration or onset.
Metrical Accent stems from the listener's sense of the beat cycle and where it repeats	Grouping/rhythmic Accent stems from the physical properties of the sounds
Metrical Accent always at the head of a metrical cycle	Grouping Accent may occur anywhere within the group
Meter is strongly predictive	Grouping is weakly predictive

Table 1. Comparison of essential characteristics of meter versus grouping

Most of the distinctions listed in Table 1 stem from the cyclical/modular aspects of metrical entrainment (a constantly repeating pattern of sensorimotor attention and/or action), versus the more continually varying durational surface of the music. While in some forms of music, such as electronic dance music, grouping patterns repeat continuously and exactly, in most musical contexts they do not. Meter, by contrast, is the stable pattern of attention and action that both predicts when most events (and the most salient events) will occur. In short, meter involves *when* events will happen, while grouping involves *what* events will happen.

One last and important point regarding meter versus rhythm: meter and grouping are not always in phase, though their phase relationship is stable. That is, group and measure boundaries may or may not align; if they do not, one has a figure that “begins with an upbeat” or anacrusis, as illustrated by Figure 1 (click [here](#) for audio clip #1).



Fig. 1. Examples of anacrusis (grouping and meter out-of-phase) versus thetic (downbeat-oriented; grouping and meter in phase) melodies.

The top staff in Figure 1 shows a melody which starts with two “pick up” notes before the quarter note/crochet on the downbeat of the first full measure. The top melody is characterized by a repeating short-short-long rhythmic figure. The melody on the bottom staff, while similar, starts squarely on the downbeat with two long notes (and then a short-short-long figure). The rhythmic initial rhythmic figures, and their corresponding repetitions throughout each phrase are boxed in red in each melody. The red boxes also show how the phase relationship between measures and groups remains stable throughout the phrase; for historical context on this stability see Mirka (2010).

RHYTHMIC PERCEPTION IS AN ACTIVE PROCESS

Rhythmic perception, like other aspects of auditory and visual perception, is not passive, but involves the active engagement of the listener with the unfolding music. In large part this is due to the nature of entrainment itself, and as a result our entrainment can give supplementation to the acoustical signal. One of the oldest and best known forms of this is “subjective rhythmization” (James, 1890; Bolton, 1894; Fraisse, 1963), or more precisely, “subjective metricization” (London, 2012). When confronted with an acoustically identical series of ticks or tones, we subjectively hear them in groups of twos or threes, groups which seem to be accentually differentiated, even though phenomenally they are not (Brochard, Abecasis, et al., 2003). Likewise our metric entrainment involves not just a sense of the

beat period, but also an organization of those beats into a repeating cycle of 2, 3, or 4 beats, though 2-beat cycles are most common, other cycles, including 5 or more beats, and more complex cycles involving uneven beats, are also possible.

Once we are entrained to a regular cycle of beats, we are able to interpolate missing beats (and may even "hear" a phenomenal articulation) where there is a silence. Figure 2 is the well-known "stop time" figure found in blues, jazz, and rock (click [here](#) for audio clip #2):



Fig. 2. The "stop time" bass/accompaniment figure. Arrows indicate the location of missing beats filled in by the listener's endogenous metric entrainment.

The initial five note pattern gives a strong sense of the underlying 8th note layer and the dotted-quarter layer which articulates the primary pulse or tactus—that is, the level at which you reflexively tap your toes. These periodicities continue in our minds even when the music does not. The audio clip which accompanies Figure 2 is an edited version of the accompaniment from Muddy Water's "Hootchie-Cootchie Man;" it was created by taking the initial five note figure and looping it through the appropriate interval (this was done to eliminate other rhythmic cues and noise from the original recording). The red arrows in the figure indicate locations where most listeners will endogenously generate the "missing" beats. Our sense of these missing beats is an instance of what have been called "loud rests" (Cooper & Meyer, 1960; London, 1993; Windsor, 1993). Little if any musical enculturation is involved in our entrainment which generates a sense of beat in "Hootchie-Cootchie Man," though familiarity with the stop-time figure will allow you to do so almost instantly. Bottom-up entrainment and our familiarity with the pattern give a strong sense of the underlying 8th note layer and the dotted-quarter layer which articulates the primary pulse or tactus—that is, the level at which you reflexively tap your toe.

In other contexts our musical enculturation plays a greater role in our rhythmic perception and cognition. The "standard pattern" in much West African Music (Agawu, 2006) provides an excellent case in point. This rhythmic pattern is asymmetrical; it can be thought of as series of durations with the proportions (2+2+1+2+2+2+1); Figure 3 gives the pattern in musical notation:

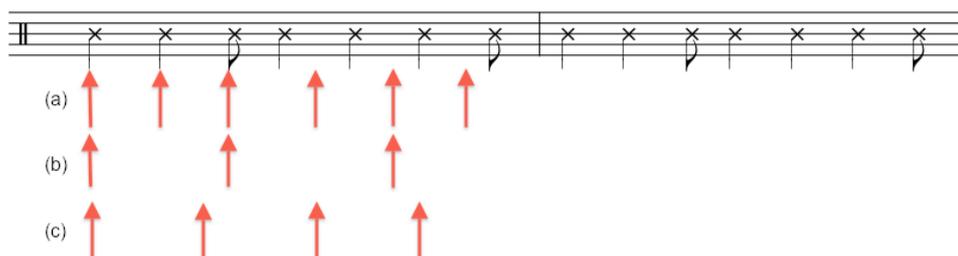


Fig. 3. A rhythmic pattern (or "bell pattern") common in West African music. Arrows indicate possible beat locations (6, 3, or 4 beats per measure, respectively), dependent upon the listener's metric construal of the pattern

This pattern is typically played by a bell (gankogui), but may be performed by other instruments; in the audio examples it is played on a hollow wood drum. First listen to the pattern in isolation (click [here](#) for audio clip #3). It is clear is that any sense of beat we might gather from the pattern moves at slower pace (i.e., slower than the 8th-note layer that the pattern strongly projects). What then determines our sense of beat? Here are three possibilities (there are others). The first, marked with the (a) arrows in Figure 3, is a fairly rapid series of six pulses over the course of the pattern, a six beat measure (click [here](#) for audio clip #4). The second, corresponding to the (b) layer of arrows, is a slower three beat measure (click [here](#) for audio clip #5); note that this also entails the 6 beat layer. Level (c) is a highly syncopated alignment of four beats with the seven-stroke pattern (click [here](#) for audio clip #6). This

pattern may not be so easy to hear for North American and European listeners, nor is it one that is readily produced by beat-finding algorithms used in computer music applications. However, the four beat meter is the normal metrical interpretation for West African music (Agawu, 2006); for an enculturated West African listener this sense of beat comes readily, especially when associated with the dances one performs with this music. Thus metric entrainment relies both on low-level, relatively automatic and modular processes of finding and sustaining periodicities present in the music (so-called "stimulus-driven attending") as well as the application of previously acquired templates for sensorimotor perception and action.

The discussion of Figure 3 also illustrates how two (or more) people listening to the same music can hear it in the context of different meters. Indeed, it is possible the same listener to hear the pattern differently on different listening occasions, akin to the figure-ground switching of dubbits or necker cubes in vision. One should notice here not only that there are different ways of construing the standard pattern, but more importantly, how different the pattern itself becomes in different metrical contexts. On the other hand, if two (or more) listeners do hear the same meter on a common listening occasion, as at a concert or other public event, their parallel entrainments will greatly enhance their joint attention and joint action (Knoblich & Sebanz, 2006). This in turn leads to a strong sense of our collective, shared experience, and hence the ubiquity of music in social, religious, political, and sports settings.

DIFFERENT TIME SCALES PLAY BY DIFFERENT RULES

Finally, our perception of duration, numerosity, sense of continuity, and ability to synchronize is different for very short (less than ≈ 400 ms), moderate (≈ 400 to ≈ 1500 ms) and longer (≈ 1500 to ≈ 5000 ms) intervals. The first two categories correspond to the distinction between "short" and "long" times proposed by Fraisse (1963, 1982). At the same time, music typically involves concurrent levels of activity that involve periodicities in different ranges. Here is a familiar example from the Beatles (click [here](#) for audio clip #7). In "Eleanor Rigby" you can hear a slow moving cello, faster moving violins, and Paul McCartney's voice moving at an intermediate (and irregular) rate. We track these different rates of motion within a single meter, and thus one of the things meter does is integrate periodicities on different time-scales.

Our ability to hear rhythmically is limited to events and event patterns within a range of 100ms to about 5-6 seconds (London 2012). If a melody is performed too fast its notes become just a blur, hence 100ms is known as the "trill threshold" (Miller & Heise, 1950). Conversely, when a melody is played too slow (i.e., with more than 1.5-2.0 seconds between notes), it is longer heard as temporally connected. Rhythm perception crucially depends on having at least one periodicity in the 300-1500ms range. Within this range we are strongly attracted to periodicities with an IOI around 600ms, the zone of "maximal pulse salience" (Parncutt, 1994; van Noorden & Moelents, 1999). Other periodicities, both faster and slower, are often present, but they are heard in relation to the anchoring range, the range in which we can hear a beat. Events with IOIs in the range of 100-300ms are too fast to be heard as beats, though we can still make durational discriminations (Hirsch, Monohan, et al. 1990) and determine numerosity, i.e., quadruplet versus quintuplet (Repp, 2007b). We are also able to synchronize with these very short events (Repp, 2006a, 2006b), though we cannot do so quite as fast if the synchronization pattern is uneven (Repp, London & Keller, 2005). Similarly, events whose IOIs are greater than 1200-1500ms lose their temporal coherence, and we no longer hear them as beats in a metrical cycle, but as isolated acoustic events (Woodrow, 1932).

Of course, we do encounter both faster and slower events in music—but these are extra-metrical/extra-rhythmic. Musicians can produce notes up to about 20 per second (the limit for "extreme drumming" competitions); beyond 10 per second (i.e., 100ms IOI) the result is a perceptual blur—in other words, a trill or a melodic sweep. Likewise, we can track slower periodicities if they are integrated into a metrical cycle which involves more rapid levels of motion. This is precisely what occurs when we listen to the cello line in "Eleanor Rigby". The intervening beats provide temporal continuity and give timing information that allows one to entrain to the slower periodicities.

CONCLUSION

There are, of course, many more things to be said about rhythm perception and cognition that are relevant to music-language relationships, ranging from the relationship between a musician's native language and their compositions and performances (e.g., Patel & Daniele (2003), which compares durational variability in music and language by composers of various nationalities), to neuro-physiological correlations between musical and linguistic processing (Patel, 2007), to relations between rhythmic accent and linguistic accent (Palmer & Kelly (1992), on the nuclear stress rule), and to

linguistic versus musical syntax (Swain, 1997). But let me conclude with a few cautionary observations on some important ways in which musical and linguistic rhythms are not alike.

Most obviously, music involves isochrony, while language does not. To be precise, while neither language nor music has an isochronous rhythmic surface, unlike language music affords patterns of sensorimotor entrainment, metric cycles that do involve isochronous periodicities. Thus when texts are set to music the rhythmic result is a "flattening out" of the normal speech timings. As noted above, music involves events in broad temporal range, a range that is a combination of several different time-scales. Language also involves a broad range of timings, but in spoken language important structural information occurs in very short time intervals (20-50ms range), especially with respect to differentiation of consonants. While too short for musical rhythm, timings in this range are important for the differentiation of musical timbres.

Finally, musical rhythm is continuous, as once a pulse gets going it often remains constant for at least several minutes (hence the notion of playing "without missing a beat"); by contrast, spoken language proceeds by fits and starts. Limited by the extent of a breath group and the temporal vagaries created by multi-speaker conversations, spoken language will naturally give little affordance for entrainment. In language one is constantly adapting the changing discursive context, as we try to figure out what we are going to say and as well as how our interlocutors are going to proceed. Something similar occurs at beginning of piece of music (or when one is flipping through the radio dial), when we must discover the basic beat period, meter, tempo, and so forth. But then the music keeps on going . . . whereas an utterance does not.

REFERENCES

- Agawu, K. (2006). Structural analysis or cultural analysis? Competing perspectives on the "standard pattern" of West African rhythm. *Journal of the American Musicological Society*, Vol. 59, No. 1, pp. 1-46.
- Bolton, T.L. (1894). Rhythm. *The American Journal of Psychology*, Vol. 6, No. 2, pp. 145-238.
- Brochard, R., Abecasis, D. et al. (2003). The "ticktock" of our internal clock: Direct brain evidence of subjective accents in isochronous sequences. *Psychological Science*, Vol. 14, No. 4, pp. 362-366.
- Chen, J.L., Penhune, V.B., & Zatorre, R.J. (2008). Listening to musical rhythms recruits motor regions of the brain. *Cerebral Cortex*, Vol. 18, No. 12, pp. 2844-2854.
- Drake, C., & Botte, M.-C. (1993). Tempo sensitivity in auditory sequences: Evidence for a multiple-look model. *Perception and Psychophysics*, Vol. 54, No. 3, pp. 277-286.
- Fraisse, P. (1963). *Psychology of Time*. New York, Harper.
- Fraisse, P. (1982). Rhythm and Tempo. In D. Deutsch (Ed.), *The Psychology of Music, 1st Edition*. New York, Academic Press, pp. 149-180.
- Friberg, A., & Sundberg, J. (1995). Time discrimination in a monotonic, isochronous sequence. *Journal of the Acoustical Society of America*, Vol. 98, No. 5, pp. 2524-2531.
- Gjerdingen, R.O. (1989). Meter as a mode of attending: A network simulation of attentional rhythmicity in music. *Integral*, Vol. 3, pp. 67-91.
- Grahn, J.A., & Brett, M. (2007). Rhythm and beat perception in motor areas of the brain. *Journal of Cognitive Neuroscience*, Vol. 19, No. 5, pp. 893-906.
- Hirsh, I.J., Monohan, C.B., et al. (1990). Studies in auditory timing: 1. Simple patterns. *Perception and Psychophysics*, Vol. 47, No. 3, pp. 215-226.
- James, W. (1890, 1950). *The Principles of Psychology*. New York: Dover Reprint.
- Jones, M.R., & Boltz, M. (1989). Dynamic attending and responses to time. *Psychological Review*, Vol. 96, No. 3, pp. 459-91.
- Knoblich, G., & Sebanz, N. (2006). The social nature of perception and action. *Current Directions in Psychological Science*, Vol. 15, No. 3, pp. 99-104

- Large, E.W., & Palmer, C. (2002). Perceiving temporal regularity in music. *Cognitive Science*, Vol. 26, pp. 1-37.
- Large, E.W., & Kohlen, J.F. (1994). Resonance and the perception of musical meter. *Connection Science*, Vol. 6, pp. 177-208.
- Lerdahl, F., & Jackendoff, R. (1983). *A Generative Theory of Tonal Music*. Cambridge: MIT Press.
- London, J.M. (1993). Loud rests and other strange metric phenomena (or, meter as heard). *Music Theory Online*, Vol. 0, no. 2 <http://www.mtosmt.org/issues/mto.93.0.2/mto.93.0.2.london.art>
- London, J.M. (2001). Rhythm. In J. Tyrrell & S. Sadie (Eds.), *The New Grove Dictionary of Music and Musicians*. Oxford: Oxford University Press, Vol. 21, pp. 277-309.
- London, J.M. (2012). *Hearing in time: Psychological Aspects of Musical Meter*, 2nd edition. New York: Oxford University Press.
- Miller, G.A., & Heise, G.A. (1950). The trill threshold. *The Journal of the Acoustical Society of America*, Vol. 22, No. 5, pp. 637-638.
- Mirka, D. (2010). *Metric Manipulations in Haydn and Mozart: Chamber Music for Strings, 1787-1791*. New York: Oxford University Press.
- Palmer, C., & Kelly, M.H. (1992). Linguistic prosody and musical meter in song. *Journal of Memory and Language*, Vol. 31, pp. 525-542.
- Parncutt, R. (1994). A perceptual model of pulse salience and metrical accent in musical rhythms. *Music Perception*, Vol. 11, no. 4, pp. 409-464.
- Patel, A.D. (2007). *Music, Language, and the Brain*. New York: Oxford University Press.
- Patel, A.D., & Daniele, J.R. (2003). An empirical comparison of rhythm in language and music. *Cognition*, Vol. 87, No. 1, pp. B35-B45.
- Repp, B.H. (2001). Phase correction, phase resetting, and phase shifts after subliminal timing perturbations in sensorimotor synchronization. *Journal of Experimental Psychology-Human Perception and Performance*, Vol. 27, No. 3, pp. 600-621.
- Repp, B.H. (2006a). Musical Synchronization. In E. Altenmüller, M. Wiesendanger, & J. Kesselring, (Eds.), *Music, Motor Control and the Brain*. New York: Oxford University Press, pp. 55-76.
- Repp, B.H. (2006b). Rate Limits of Sensorimotor Synchronization. *Advances in Cognitive Psychology Special Issue: Music Performance*, Vol. 2, Nos. 2-3, pp. 163-181.
- Repp, B.H. (2007a). Hearing a melody in different ways: Multistability of metrical interpretation, reflected in rate limits of sensorimotor synchronization. *Cognition*, Vol. 102, No. 3, pp. 434-454.
- Repp, B.H. (2007b). Perceiving the Numerosity of Rapidly Occurring Auditory Events in Metrical and Nonmetrical Contexts. *Perception & Psychophysics*, Vol. 69, No. 4, pp. 529-543.
- Repp, B.H., London, J.M., & Keller, P.E. (2005). Production and synchronization of uneven rhythms at fast tempi. *Music Perception*, Vol. 23, No. 1, pp. 61-78.
- Swain, J. (1997). *Musical Languages*. New York, W. W. Norton.
- van Noorden, L., & Moelants, D. (1999). Resonance in the perception of musical pulse. *Journal of New Music Research*, Vol. 28, No. 1, pp. 43-66.
- Windsor, W.L. (1993). Dynamic accents and the categorical perception of metre. *Psychology of Music*, Vol. 21, pp. 127-140.
- Woodrow, H. (1932). The effect of rate of sequence upon the accuracy of synchronization. *Journal of*

Experimental Psychology, Vol. 15, No. 4, pp. 357-79.

Zatorre, R.J., Chen, J.L., & Penhune, V.B. (2007). When the brain plays music: Auditory-motor interactions in music perception and production. *Nature Reviews Neuroscience*, Vol. 8, No. 7, pp. 547-558.