Psychomusicology, 10, 88-106

RHYTHM AND METER: A THEORY AT THE COMPUTATIONAL LEVEL

Philip N. Johnson-Laird
Princeton University

This paper presents a theory of rhythm and meter. The theory treats a rhythm as a sequence of onsets of events created and perceived within a metrical framework. Although there are many distinct rhythms within a given meter, they fall into certain natural families.These families derive from an underlying rhythm (an "UR-rhythm," akin to a prototype) that depends on three distinct classes of events. They are stated here in their order of significance: (a) a syncopated note, (b) a note that has its onset at the start of a beat, and (c) every other sort of music event, including a rest, a continuation of a note, and a note that is not syncopated but that occurs within the interval from one beat to the next. An UR-rhythm is defined by the first significant event for each beat in each measure of the music phrase. The theory analyzes music phrases in terms of the cues to their boundaries. Its fundamental claim is that rhythms, unlike tonal harmonic sequences, require only weak computational power, i.e., they can be created and perceived using only a limited working memory for the results of intermediate computations.

Music is rhythmic. This almost universal characteristic poses a challenge to cognitive science: why should music be organized in this way, and what does it tell us about the psychological principles underlying the production and perception of music? The aim of this paper is to try to answer these questions by proposing a theory of rhythm. This theory deals with the creation of rhythms, particularly their improvisation and their perception. It considers the need for, and the nature of, metrical frameworks, and it analyzes the rhythms of music phrases only within such frameworks. It is a theory at the computational level, that is to say, it concerns what the mind has to compute rather than how the mind carries out the computations (see Marr, 1982). However, to understand how the mind carries out any task, we need first a good account of what it is doing. Elsewhere, I have reported studies of the nature of creativity that led to the development of computer programs that can improvise music (see Johnson-Laird, 1991).

There is a long tradition of theoretical and experimental investigations of rhythm (see e.g., Stetson, 1905; Woodrow, 1909; Weaver, 1939). One common feature of many theories is that rhythmic organization exists on several levels (e.g., Lerdahl & Jackendoff, 1983). Such theories can perhaps be traced back to the work of Heinrich Schenker (see e.g., Forte, 1959). Some authors base the deepest level on patterns derived from prosody (Cooper & Meyer, 1960); others have repudiated this view and argued that the underlying "strata" are musical (see Yeston, 1976, p. 31). It is not my intention to review these theories: the account that I shall develop is in debt to some of them, but, as I hope will be clear, it offers a different account of the underlying basis of rhythmical prototypes.

Johnson-Laird, P. N., *Rhythm and Meter: A Theory at the Computational Level* ,
Psychomusicology, 10:2 (1991:Fall) p.88

Rhythm as a Metrical Phenomenon

A priori definitions seldom help much in science, where the aim is to reach explanations. Nevertheless, the Shorter Oxford Dictionary provides us with a useful working definition:

> *rhythm*: (mus) That feature of music composition which depends on the systematic grouping of notes according to their duration.

Rhythmic music certainly depends on music events that tend to be grouped together. For example, if you tap your pencil on your desk with the following time intervals (where "x" denotes a tap):

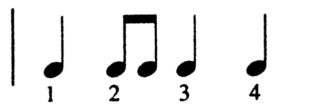
```
       1s          .5s       .5s          1s
x <————————> x <————> x <————> x <————————> x
```

then it is hard for you not to perform (and for others to perceive your performance) in a rhythmical way. As many theorists have pointed out (e.g., Povel, 1984), the critical feature of a rhythm is the sequence of onsets of its notes. Hence, if you tap the rhythm of a familiar piece, such as "We all live in a yellow submarine," then listeners will be able to identify it. Tapping, of course, provides information only about onset times, but a rhythm is not just a sequence of events with different onset times. This claim may seem strange at first because if you program your computer to produce a sequence of tones with the onset times above, you will probably perceive its output as a rhythm—and you may even want to add the two notes that complete its familiar pattern. In fact, you are bringing to the computer's output something that is not actually to be found in its performance—you are perceiving the sequence of events within a metrical framework. I will present some evidence for this claim presently, but first I need to explain the nature of meter.

The events making up a rhythm—the taps in the example above—are perceived within a framework of "beats," Each beat occurs at a regular interval in time, and the sequence forms a group of two, three, four, or more beats, which is repeated over and over with an emphasis on the first beat in each group. Thus, your tapping of the example above is likely to be perceived within the following grouping of beats:

```
      1s          .5s       .5s          1s           1s
 x <————————> x <————> x <————> x <————————> x <————————>
| 1             2                3             4           |
```

The vertical lines indicate the beginning of each new group (the so-called "measure" or "bar"), and the numbers indicate each of the equally-spaced beats in the measure. This regular meter provides the framework for rhythm. The conventional notation of rhythm accordingly relies on vertical lines to demarcate the measures. The rhythm above is written as follows (without the numbers denoting the beats, of course):



In some music—jazz, for example—the metrical structure tends to be made explicit: certain instruments, such as the double bass and drums, play a note on each beat of nearly all measures. Much music, however, relies on musicians and listeners alike to perceive the metrical pulse even though it is not explicitly marked by an event on each beat. The perception of this framework is therefore something that individuals bring to the music: they perceive it within a framework which is cognitively constructed, and which is not necessarily part of the objective events that occur in the music (see also, e.g., Lee, 1991). This point is so obvious to musicians that they seldom comment upon it: a common sign of elementary musicianship is to clap on the beats of a given piece of music. Although musically-untutored individuals may not be able to carry out this task, there are grounds for supposing that they nevertheless perceive music within a metrical framework. My colleagues Jung-Min Lee and Malcolm Bauer have observed at my suggestion that if one counts, "1 2 3 4" in a regular way so as to establish a meter, and then claps the following simple rhythm in the same tempo:

listeners will judge all four notes to be of equal duration. But, if one counts, "1 2 3 4", and then claps:

they will judge that last note to be shorter than its predecessor. These judgments are remarkable because the claps themselves are all of the same brief duration. So, why should the last note in the first case be judged as longer than the last note in the second case? The answer is that listeners must perceive both rhythms as having a meter of four beats to the measure—as suggested by the experimenter counting "1 2 3 4," and they must assume that there will be a note on the first beat of the next measure. In the first case, the final clap is therefore perceived correctly as a 1/4 note[1] whereas, in the second case, it is perceived correctly as an 1/8 note. If musicians are asked to carry out the same task without trying to notate the sequences in their mind, they are likely to judge that the following rhythm:

ends with a longer note. Musically-naive individuals may be confused by this case because its last note is a syncopation—a notion that I will explain presently. In short, listeners make a default assumption about the metrical location of the next event: it will occur on the first beat of the following measure. This phenomenon is probably the simplest way to convince a skeptic of the psychological reality of metrical structure.

The Nature of Meter

Meter seems so natural, and its lack in certain modern compositions so unnatural, that it clearly has an important function. But what is this function—why should music have a meter? The answer is: organization. The cognitive and emotional impact of music depends on the temporal organization of pitches, intensities, and timbres. Organization is, by definition, a matter of relations among elements, and, whatever its domain, its principles do not allow all possible relations among all possible elements. Its constraints decrease uncertainty (in the sense of statistical information theory) and consequently increase predictability. In music, for example, the tonal relations established by a particular key make it easier to predict the next note of a melody: certain notes will be favored by the key, certain notes will be unlikely, and still other notes will be musically impossible (see Krumhansl, 1991). The temporal aspect of music events is similarly organized so that their sequence is constrained and thus to some extent predictable. If the onset and duration of each event is entirely random (as it is in certain aleatory music), then listeners are likely to hear a series of notes with no discernible structure. They are unable to predict what will happen next unless the statistical properties of the events have some obvious constraints; performers have great difficulty in performing such a score, and improvisers are unlikely to extemporize music of this sort because they have difficulty in generating genuinely random events.

A first step towards temporal organization is the synchronization of certain music events. Human beings, I assume, have access to an internal clock[2] that can be used to generate a 'pulse' demarcating roughly equal intervals of time:

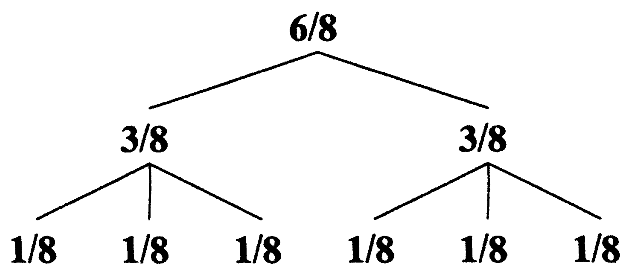
Such a pulse can serve to organize music events: each event can be synchronized with a pulse. The result, of course, is totally predictable and thus monotonous. For variety, some events should not occur at the start of a pulse. But when precisely should they occur? Once again, random onsets between pulses will be difficult for musicians to time accurately. What is more natural is to subdivide each interval (or measure) between pulses by a separate sequence of pulses. These are the beats of the measure:

This principle can be applied recursively and the interval between one beat and the next subdivided in turn, and so on:

In this case, the measures at the top level are subdivided by beats into three equal intervals, and then each of these intervals is subdivided into two.

There are limits to the number of subdivisions that are likely to yield perceptible rhythms, and there are also limits on the number of subdivisions at any level. A meter is readily perceptible if the measures are subdivided by two, three, or four beats; and occasionally the measures may be subdivided by a greater number of beats. The maximum number of beats that can be apprehended as making up an undivided measure is probably 7 ± 2—Miller's (1956) magical number—beyond that the meter is subdivided. Even with seven beats to a measure, musicians are likely to divide the measure into a pattern such as: 3 + 2 + 2.

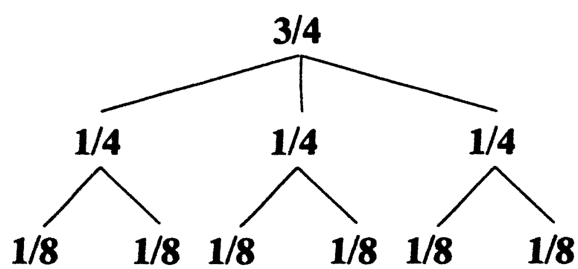
This analysis brings out the point that meter is not merely the number of beats in the measure, but has a more complex structure: each beat can also be subdivided in a regular way. A good example is the contrast between two beats to the measure that are each subdivided into three, i.e., 6/8 time, and three beats to the measure that are each subdivided into two, i.e., 3/4 or waltz time. Each measure of these two meters contains the same total number of units, i.e., six 1/8th notes in both cases. The important point is that the *structures* of the two sorts of measure are different. A measure of 6/8 has the following structure:



Such a tree diagram represents a grouping of events, which can equally well be representing by the following bracketing:

(((1/8)(1/8)(1/8)) ((1/8)(1/8)(1/8)))

A measure of 3/4, however, has a different grouping of the same six events:



Leonard Bernstein's song "America" (from *West Side Story*) wittily alternates measures of these two sorts.

Where the intervals between beats are subdivided regularly into equal intervals, listeners have an impression of one meter nested within another. Such simultaneous meters, however, generally occur only at two adjacent levels. The mental principles underlying meter seem able to cope with two or three levels at most. Likewise, they do not easily grasp two independent meters at the same level. It takes work to learn to tap simultaneously, say, four beats to the measure with one hand and three beats to the same measure with the other hand. Musicians typically first learn the overall rhythmic pattern that the two meters make, and then how to divide it appropriately between the two hands. West African music, which is polyrhythmic in this way, is also acquired by a similar technique of learning the overall pattern of the simultaneous meters.

We can use a grammar to capture the grouping principles of meter. A grammar is a finite set of rules for a domain of symbols (or language) that characterizes all the properly formed constructions, and provides a description of their structure. There is a general relation between grammars and computer programs: the output of any program can be captured by a grammar. For computationally more powerful programs, grammars of increasing power are needed in order to characterize their outputs. Perhaps surprisingly, computational power depends solely on memory for intermediate results. The weakest computational devices are so-called finite-state machines: they need a memory only for a finite number of intermediate results, which they represent either by states of the machine or by a working memory of a fixed capacity. The grammar that corresponds to a finite-state machine is known as a "regular" grammar, and we will presently encounter an example of such a grammar. The power of a finite-state device can be enhanced by equipping it with an indefinite amount of working memory in the form of a stack, where access to this memory is limited to whatever is the topmost symbol on the stack. The more powerful grammar that corresponds to such a device is known as a "context-free grammar." Such a grammar contains a finite set of rules that each allow a single symbol to be rewritten as other symbols regardless of the context in which the symbol occurs. The reason for invoking grammars is simple: the human cognitive system has limited resources, and so a major aim of cognitive science is to pin down the amount of computational power that is needed for various cognitive tasks (for an elementary introduction to these ideas, see Johnson-Laird, 1988, Ch. 3).

Longuet-Higgins (1978) has shown that the grouping principles of meter can be captured by a context-free grammar. His grammar contains the following sort of rules, which generate the tree above:

3/4	→	1/4 1/4 1/4
	→	rest
	→	𝅗𝅥.
1/4	→	1/8 1/8
	→	rest
	→	♩

The arrow means that the symbol on the left-hand side can be rewritten by choosing one of the alternative expansions on the right-hand side. For example, the symbol 3/4 could be rewritten as: 1/4 1/4 1/4 (as in the tree above); it could be re-written as "rest," which signifies a silence for the corresponding number of notes; or it could be rewritten as 𝅗𝅥., which signifies a note lasting for three 1/4 notes. Likewise, the symbol 1/4 can also be rewritten in several ways, e.g., as ♩, which signifies a 1/4 note.

A context-free grammar makes a strong structural claim about the grouping of music events, namely, that two or more events can be grouped together to make up a single higher-order constituent—in this case, the measure. Thus, the bracketing for the measure:

((1/4)(1/4)(1/4))

shows that three 1/4 notes make up a single measure. This claim could not be made by postulating the weaker hypothesis of a "regular" grammar. The case for context-free power rests on the fact that musicians group notes in the way shown by the bracketing.

The processing of context-free grammars, as I mentioned earlier, calls for a working memory in the form of a stack. Access to the stack is solely by the topmost item. In other words, the first item into memory is the last item out. For example, consider a simple procedure for checking that the brackets match in expressions, such as:

((1/4) ((1/8)((1/16)(1/16))) (1/4))

They could be parsed using the following principles:

1. Read each symbol (from left to right), and ignore everything except the brackets.
2. If a symbol is a left bracket, then put it on the stack.
3. If a symbol is a right bracket, then remove the topmost symbol from the stack.

Well-formed expressions are precisely those for which the parse begins and ends with an empty stack, and for which there is never any attempt to remove an item from an empty stack. Unlike natural language, there is a bound on the degree of *self-embedding* in the sequences of elements making up meters and rhythms. Hence, for the production or perception of a meter, there is no real need for a stack of unlimited size. A working memory of limited capacity will suffice for a proper conception of metrical structure. Let us now turn from meter to the topic of rhythm proper. I will begin with the phenomenon of syncopation because it can be elucidated only by a metrical conception of rhythm.

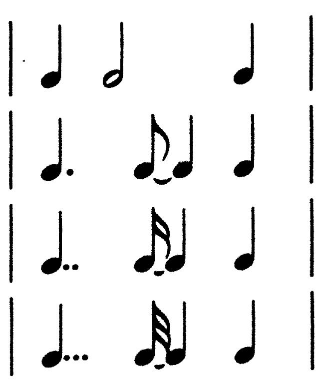
Syncopation

A syncopation, roughly speaking, is the occurrence of a relatively long note at an unexpected place—as though it starts in anticipation of its proper place in metrical structure. In the following example in 4/4 time:

the final note in the first measure is syncopated. Its onset is after the third beat but before the fourth beat, and so it begins on a metrical unit of less importance than the fourth beat on which there is no onset of a note. A syncopation therefore depends on two factors: the onset of the syncopated note and the onset of the note following it—the note in the second measure in the example above. Between these two notes there occurs a metrical unit of greater importance than the metrical unit corresponding to the onset of the syncopated note[3]. In the more precise terms of Longuet-Higgins and Lee (1984), a syncopated note is one that has an onset on a metrical unit of lesser importance than one that intervenes prior to the next note. The importance of a metrical unit is given by its height in the tree generated by Longuet-Higgins's metrical grammar, or equivalently its position in the corresponding labeled brackets. For example, consider a bar consisting of five notes in 4/4 meter with the following structure:

(((1/4)(1/4)) ((1/4) ((1/8)(1/8))))

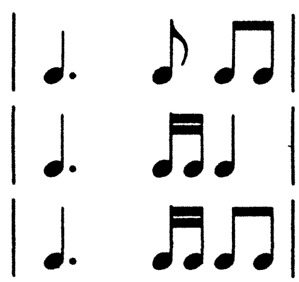
The first note occurs on the most important metrical unit because its onset is co-incident with the bracketing at the highest level, i.e. the bracket that precedes the whole measure. The third note is on the next most important metrical unit because its onset is co-incident with the bracketing at the next highest level, i.e. the bracket that precedes the second half of the measure. The second and fourth notes are on the lesser important metrical units co-incident with bracketing one further level down. Finally, in this example the last note is on the least important metrical unit in the measure—one co-incident with the lowest level of bracketing.

According to Longuet-Higgins and Lee's principle, the following are examples of increasingly marked syncopations of the second note:

In theory, the closer the onset of the syncopated second note to the onset of the third beat, the smaller the importance of its own metrical unit, and so the greater the degree of its syncopation. Eventually, as the series above continues, listeners begin to perceive the syncopation as merely a slight perturbation in the performance of an unsyncopated rhythm. In fact, the principal perceptual distinction appears to be between the minor syncopation that occurs in the first example, where all the notes start on a beat, and those cases in which the syncopated note clearly occurs between the two beats. This intuition, which has yet to be confirmed empirically, suggests a way in which to define *families* of rhythms based on the same underlying prototype.

Families of Rhythms

There are many possible rhythms, but different rhythms vary in the degree to which they resemble one another. Indeed, some rhythms that are different nevertheless appear to depend on the same prototype both for their creation and their perception. Listeners can detect such resemblances, and musicians who improvise can readily extemporize variants of the same rhythm. This similarity reflects an underlying cognition of rhythms, and so it makes psychological sense to consider rhythms as falling into different families. In this section, I will propose a theory of the prototypes for families of rhythms.

The first assumption of the theory is that a family of rhythms depends solely on onsets. An onset can occur either on a beat or else between two beats, and an onset can occur either as a syncopation or as an ordinary note. The second assumption of the theory is that onsets differ in their significance, and that a family of rhythms is defined by the first significant onset, if any, that occurs for each beat. There is a simple rank order of significance: the most significant event is a syncopation, the next most significant event is a note on the beat, and, finally, all other events are of equal minimal significance. If there is a note on the beat, then, for example, it does not matter whether it is followed by one or more notes prior to the next beat: the result is a different rhythm, but one within the same family. Hence, each of the following rhythms in a meter of 3/4 is a member of the same family:



The family is based on the following prototype of significant events:

This simple prototype can be treated as an *underlying representation* of a rhythm (an UR-rhythm) from which all members of the family are derived. Any variant on the UR-rhythm is allowed provided that no change occurs to the first significant event in each beat. Hence, a rhythm such as:

is not included in the family above, because it contains a note on the second beat of the measure—that is a significant event, whereas a note half way through is not. The following example:

is also not a member of the family, because it contains a syncopation after the first beat.

UR-rhythms correspond to phrases, which can last for any number of measures, or parts of measure, and they depend on the three categories of significance, which I will label as follows:

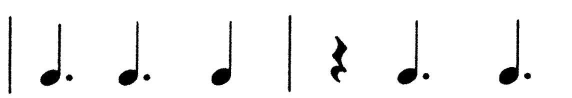
Sync: A syncopation starts after the beat, and so by definition a note cannot occur on the next beat.
Note: A note occurs at the start of a beat.
Other: This includes all possibilities apart from a Sync or a Note, e.g., there is a rest, or one or more unsyncopated notes occur after the start of the beat.

These principles may be too coarse, and it may be necessary to refine them so that they apply at more than one level of metrical analysis, i.e., for units smaller than the beat.

An UR-rhythm of two measures in 4/4 can take the form:

| Note Sync Other Note | Other Note Sync Other |

The simplest realization of this rhythm (the Bossa Nova) is:



which is derived in the following way. The first item in the UR-rhythm, Note, must be realized by a note on the first beat of the measure. (It may, or may not, be followed by other notes prior to the next beat.) The second item in the UR-rhythm, Sync, must be realized by a syncopation at some point after the second beat, and it cannot be followed by a note on the third beat. The third-item in the UR-rhythm, Other, allows the syncopated note to continue through the interval from the third to the fourth beat. The fourth item in the UR-rhythm, Note, must be realized by a note on the fourth beat, and so on.

The set of possible UR-rhythms for two measures can be captured by a simple regular grammar based on two sorts of symbols: the *terminal* symbols consist in Sync, Notes, Other, and the bar line; and the *nonterminal* symbols are the theoretical categories of the grammar. The grammar has the following sort of rules, in which the terminal symbols have been italicized:

First-measure	→	*|*	1st-beat
1st-beat	→	*Sync*	After-sync-on-1st-beat
	→	*Note*	2nd-beat
	→	*Other*	2nd-beat
After-sync-on-1st-beat	→	*Other*	3rd-beat
	→	*Sync*	After-sync-on-2nd-beat
2nd-beat	→	*Sync*	After-sync-on-2nd-beat
	→	*Note*	3rd-beat
	→	*Other*	3rd-beat

and so on.

Each symbol on the left-hand side can be rewritten in alternative ways, e.g., the 1st-beat could be a Sync, and in this case the next symbol is After-sync-on-1st-beat, and the rules for rewriting this symbol constrain the next terminal symbol to be either Sync or Other, i.e. after a syncopation there cannot be a Note on the next beat. A regular grammar, by definition, is one in which the right-hand side of each rule must contain at least one terminal symbol in a fixed position, and no more than one nonterminal symbol.

A regular grammar is the weakest possible sort of grammar. It corresponds to the claim that UR-rhythms can be generated and parsed without more than a finite working memory for the results of intermediate computations. As in the case of meter, it may be necessary to use a context-free grammar to capture the groupings of notes that accord with musicians' intuitions. Yet, the fundamental claim of the theory is that rhythms are computationally simple. Their generation is simpler than, say, the generation of tonal chord sequences, which call for considerable computational power. Rhythms can be created and perceived within the constraints of a small finite working memory. Tonal harmonic sequences cannot be generated within such a memory, and accordingly depend on notation and other methods of extending the capacity of working memory, e.g., the use of long-term memory (see Johnson-Laird, 1991).

A given UR-rhythm has many different realizations, which are the different members of its family. These realizations can be generated by a regular "transducer"—a procedure that takes an UR-rhythm as input and then rewrites it using a regular grammar. The form of these rules for a meter of 4/4 is as follows:

Input symbol	Rule of grammar			
|	Measure	→	|	1st-beat
Note	1st-beat	→		2nd-beat
		→		2nd-beat
		→		2nd-beat

and so on.

The rules expanding Sync allow for the syncopation to occur either after a rest or after a preceding note on the beat, for example:

Input symbol	Rule of grammar			
Sync	1st-beat	→		2nd-beat
		→		2nd-beat

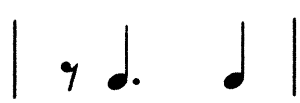
and so on.

One of the consequences of the present analysis of families of rhythms is that 1/8th note pickups, and other such lead-ins, are of no consequence. Both they and rests (i.e., lack of a pickup) are members of the same category, Other. It follows that the phrase:

is from the same family as:

A more subtle point is that the occurrence of a syncopation may, or may not, be preceded by a note on the beat: both cases count as a syncopation, since this category is the most significant event that can occur. It follows that the phrase:

is a member of the same family as:



In short, the category, Sync, indicates that a syncopation occurs and says nothing about whether or not it was preceded by any notes on or after the beat.

So far, I have been talking as though the same rules would apply to all genres of music. This assumption is false. The UR-rhythm for the Bossa Nova is impossible for a piece written in the eighteenth century. Hence, both the principles that specify families of UR-rhythms and the transducer principles that realize their actual forms are specific to genres. The following UR-rhythm is likely to occur in many genres:

| Note Note Note Note | Note Other Other Other |

Its simplest realization is:

which occurs in many genres of music. But, its particular realizations may differ from one genre to another. Certain UR-rhythms also appear to be specific to genres. The following rhythm, for example, is unlikely to occur in classical music but is common in jazz:

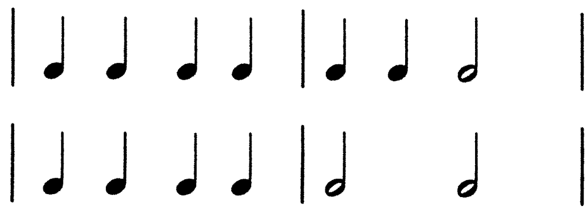
| Note Note Note Note | Sync Other Other Other |

where it might be realized as:

The essence of the theory of production and perception that I have outlined is that rhythms occur within a metrical framework, and that the prototype underlying a family of rhythms depends on a simple tripartite classification of the events that occur from one beat to the next. Events are ranked according to their significance: syncopations, notes that start on a beat, and other events such as rests or unsyncopated notes. The theory is speculative, but it is easily tested: variants of the same UR-rhythm should be judged as more similar to one another than those based on different UR-rhythms. For example, the opening phrase of "We all live in a yellow submarine" should be judged as similar to the opening phrase of the Christmas carol "Good King Wenceslaus," but dissimilar to the opening phrase of Beethoven's Fifth Symphony.

Phrases and their Constituents

A melody consists in one or more melodic phrases, each of which has a rhythm. The rhythm of an individual phrase is derived from an UR-rhythm, and so in order to give a fuller account of UR-rhythms I need to consider what counts as a phrase from the standpoint of rhythm. For the player of a wind instrument, a phrase is typically played in a single breath. Cognitively speaking, listeners need to get their breath back too. Hence, in my view, a phrase is the music analog of a sentence. It allows time for the listener to make music sense of the events that precede its ending, and it allows time for musicians to focus on the upcoming events. If you listen to a melody, you will usually have a clear intuition about where one phrase ends and another starts. If you tap the following rhythm (from "Good King Wenceslaus"), you will probably agree that it divides into two phrases:

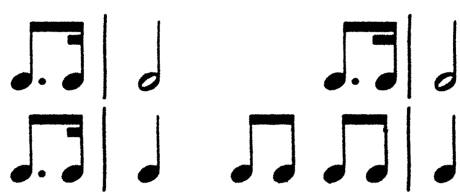


The first two measures make up one phrase, and the second two another. The boundary between one phrase and another appears to depend on several cues. The principal cue concerns rhythm (though aspects of key are also important in certain cases). Thus, as in Wenceslaus, there is a lengthier interval from the onset of the last note in one phrase to the onset of the first note in the next phrase.

Phrases can have subphrases as constituents. Consider, for example, the minuet in Beethoven's Piano Sonata, Op 49, no. 2. Here is the rhythm in 3/4 of the opening theme:

one naturally groups this into two subphrases that end on the long notes. These two constituents together make up a higher order phrase. The piece continues with the phrase:

and a repeat of the opening with a slight modification at the end:

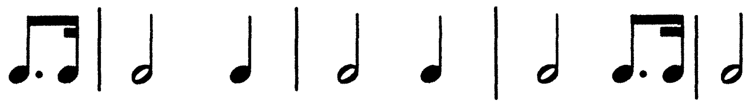


At this point, a new section begins, which acts as a "bridge" leading back to the opening. This bridge starts with two subphrases:

which are followed by the phrase:

that brings us back to the opening.

What factors underlie these intuitions about the division of the melody into phrases? One cue, as I have noted, is that a phrase often ends with a note that is of longer duration than any within the phrase. Another factor is the use of rests, which serve to emphasize the interval between onsets. Suppose that Beethoven's minuet had instead started like this:



then the division between phrases might occur in the third measure. The cue here is the repeat of the rhythm that begins the first phrase. It follows that the initial notes of the Sonata:

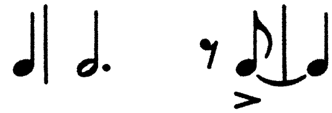
are ambiguous. What determines that they are a subphrase in the Sonata is the sequence of notes that occurs afterwards. Even the cue of a relatively longer onset between phrases depends on the melody as a whole and its harmonic sequence. The opening phrase of Dizzy Gillespie's theme, "Night in Tunisia," for example, has the following rhythm:

The long note in the first measure is not the end of the phrase, presumably because it is a weak syncopation and because the interval from the last note of the phrase to the start of the next phrase is even longer. The opening phrase of "Walkin'," a composition by Miles Davis, has the following rhythm:

The first phrase ends, not with the first note in the second measure, but with the accented syncopation at the end of that measure, which is tied through to the next measure. Onset is a slight cue here, but it is helped by the tonality of the piece: the first note in the second measure is an augmented fourth—an unlikely

Johnson-Laird, P. N., *Rhythm and Meter: A Theory at the Computational Level* ,
Psychomusicology, 10:2 (1991:Fall) p.88

note with which to end a phrase—and it leads to the fifth, which is the syncopated note that ends the phrase. In an alternative version of the theme, it is realized as:



which bears out the earlier analysis of families of rhythms. Both versions are realizations of the same UR-rhythm:

Other Note | Note Other Other Sync | Other

At the highest level of organization, a melody is a sequence of phrases. They can vary in length from less than a measure to several measures, and a good musician aims for variety. At this point, the distinction between perceiving a melody and creating one becomes pertinent. Improvisers tend to have little conception of what notes to play beyond the current phrase, and they may even be surprised by what they play. To paraphrase E.M. Forster, the improviser's maxim is: how do I know what I am improvising until I hear what I play? Hence, what holds a series of improvised phrases together is unlikely to be any complex architectural design. In the case of modern jazz, for example, an improvisation is based on the framework of a chord sequence—typically a 12, 16, or 32 bar sequence, which is repeated chorus after chorus as the basis for generating ideas. No explicit music plan governs the rhythmic structure of an improvisation above the level of individual phrases. There seem to be no strong aesthetic principles at stake above the level of phrases, and no strong intuitions about a good organization other than that it should contain variety. Although the principles underlying rhythmic phrases are best described in two stages, as I have done, one of the tasks of the musician learning to improvise is to master the principles in a way that minimizes the load on working memory. Both the generation of UR-rhythms and their transduction into actual phrases can be carried out by finite-state machines, and, in fact, the two devices can be collapsed into a single finite-state machine that creates rhythms. The arbitrary choice of one alternative principle over another yields phrases of various lengths. The improvisations of rhythm can therefore be carried out using the weakest possible computational resources.

The Expressive Nature of Rhythmical and Metrical Timing

Rhythm and meter are cognitive aspects of the human conception of temporal events. They readily enter into the performance of any repetitive activity: I hazard a guess that rhythm is particularly important in the unconscious and automatic performance of any repetitive sensory-motor skill (see Wason, 1954). Rhythms also exist in the minds of musicians, and performance may convey them to the minds of listeners. A mathematically exact meter, however, is unnatural. It sounds mechanical, and indeed empirical studies show that the real-time properties of music performance—the nuances of phrasing—introduce many significant departures from a strictly timed realization of the score. The performance of waltzes by Viennese orchestras depends on a specific metrical distortion that yields a characteristic lilt (see Bengtsson & Gabrielsson, 1983; Gabrielsson, Bengtsson, & Gabrielsson, 1983). An elusive aspect of jazz timing is "swing": a propulsive rhythmic feeling that impels all the best performances but that is difficult to discern in the actual durations of notes (see Johnson-Laird, 1991). Classical musicians prolong the durations of certain notes in order to convey phrasing and metrical structure (see e.g., Longuet-Higgins & Lee, 1984; Shaffer, 1984; Shaffer, Clarke, & Todd, 1985). In my view, the real-time phenomena of performances are part of a tacit cultural tradition. They are at the service of communicating different sorts of relations between rhythms and their metrical framework. They function as one final transducer through which a music conception passes on its way to an expert realization.The ultimate nature of this transducer has not been discovered yet.

Conclusions

The theory that I have described can be summarized in a few simple principles:

1. Music is rhythmic, but only onset times matter in our conception of rhythm, not whether a note is prolonged or followed by a rest.
2. Rhythmical music has a metrical pulse and it is performed and perceived within this framework, which depends on a recursive subdivision of time into small groups of equal intervals (for organizational purposes).
3. A syncopated note has an onset on a metrical unit of lesser importance than a metrical unit that intervenes prior to the next note, or prior to the start of the next measure if the syncopated note is the last in the phrase (see Longuet-Higgins & Lee, 1984).
4. There are families of rhythms based on prototypes.
5. A prototype, or underlying rhythm (UR-rhythm), depends on the first significant event occurring on each beat in each measure, according to the following rank order: a syncopation, a note occurring at the start of the beat, any other event.
6. Melodies are divided into rhythmic phrases by various cues, including a relatively longer interval between onsets.
7. Rhythms can be created and perceived within the constraints of a small finite working memory. The rhythmic system of the human mind is computationally weaker than the harmonic system.

In short, rhythm is a psychological phenomenon that depends on the way in which the mind interprets events in time: human cognition has a natural tendency to organize them metrically and to relate them to underlying prototypes.

References

Bengtsson, I., & Gabrielsson, A. (1983). *Analysis and synthesis of musical rhythm.* Mimeo: University of Stockholm.

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

Forte, A. (1959). Schenker's conception of musical structure. *Journal of Music Theory, 3*, 1-30.

Gabrielsson, A., Bengtsson, I., & Gabrielsson, B. (1983). Performance of musical rhythm in 3/4 and 6/8 meter. *Scandanavian Journal of Psychology, 24*, 193-213.

Johnson-Laird, P.N. (1988). *The computer and the mind.* Cambridge, MA: Harvard University Press.

Johnson-Laird, P.N. (1991). Jazz improvization: A theory at the computational level. In P. Howell, R. West, & I. Cross, (Eds.), *Representing music structure.* (pp. 291-325). London: Academic.

Krumhansl, C.L. (1991). Music psychology: Tonal structures in perception and memory. *Annual Review of Psychology, 42*, 277- 303.

Lee, C.S. (1991). The perception of metrical structure: Experimental evidence and a model. In P. Howell, R. West, & I. Cross, (Eds.), *Representing music structure* (pp. 59-127). London: Academic.

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

Longuet-Higgins, H.C. (1978). The grammar of music. *Interdisciplinary Science Reviews, 3*, 148-156.

Longuet-Higgins, H.C., & Lee, C.S. (1984). The rhythmic interpretation of monophonic music. *Music Perception, 1*, 424-441.

Marr, D. (1982). *Vision.* San Francisco: W.H. Freeman.

Miller, G.A. (1956). The magical number seven, plus or minus two. *Psychological Review, 63*, 81-97.

Povel, D-J. (1984). A theoretical framework for rhythm perception. *Psychological Research, 45*, 315-337.

Shaffer, L.H. (1984). Timing in solo and duet piano performances. *Quarterly Journal of Experimental Psychology, 36*A, 577-595.

Shaffer, L.H., Clarke, E.F., &Todd, N.P. (1985). Metre and rhythm in piano playing. *Cognition, 20*, 61-77.

Stetson, R.B. (1905). A motor theory of rhythm and discrete succession. *Psychological Review, 12*, 250-270, 293-350.

Wason, P.C. (1954). Soap wrappers' "Jig." *British Journal of Industrial Medicine, 11*, 279-283.

Weaver, H.E. (1939). Syncopation: A study in musical rhythms. *Journal of General Psychology, 20*, 293-350.

Woodrow, H. (1909). A quantitative study of rhythm. *Archives of Psychology, 14*, 1-66.

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

Author Notes

I am grateful to my colleagues Malcolm Bauer and Jung-Min Lee for many interesting discussions about rhythm and meter, for their hard work in carrying out experiments and computer simulations, and for critical readings of an earlier version of this paper. I thank Mari Jones and two anonymous referees for their helpful criticisms. Finally, I am also indebted to two old friends, Christopher Longuet-Higgins and Mark Steedman, who first convinced me that music raises deep problems for cognitive psychology. Requests for reprints should be sent to P.N. Johnson-Laird, Department of Psychology, Princeton University Green Hall, Princeton, N.J. 08544.

Footnotes

[1]A note that lasts for one beat of a bar is known as a "quarter note" in the USA and as a "crotchet" in the UK. I will use the US nomenclature.

[2]This clock may be based on an internal pulse-generator analogous to a mechanical clock. Alternatively, it may be based on a mechanism that depends on assessing the decay of information over time. People can readily tap equal intervals of time, where the rate ranges from roughly 4 taps per second up to 1 tap every two seconds. The fastest rate is bounded by the physiology of tapping; the slowest rate is bounded by a sense of the "phenomenal" present. Of course, one can tap at much slower rates, but it becomes necessary to count (at equal intervals) in order to demarcate them. The phenomenal present may depend on a capacity to estimate the decay of a recent event: if, say, the decay is exponential but roughly constant, then its current value will be proportional to elapsed time. Hence, in this way, decay could serve as the basis of meter.

[3]Strictly speaking, the very last note of a phrase could be a syncopation provided that it lasts through an important metrical unit, viz., a beat.