RESONANCE AND THE PERCEPTION OF MUSICAL METER

Edward W. Large

John F. Kolen

The Ohio State University

Abstract

Many connectionist approaches to musical expectancy and music composition let the question of "What next?" overshadow the equally important question of "When next?". One cannot escape the latter question, one of temporal structure, when considering the perception of musical meter. We view the perception of metrical structure as a dynamic process where the temporal organization of external musical events synchronizes, or entrains, a listener's internal processing mechanisms. This article introduces a novel connectionist unit, based upon a mathematical model of entrainment, capable of phase- and frequency-locking to periodic components of incoming rhythmic patterns. Networks of these units can self-organize temporally structured responses to rhythmic patterns. The resulting network behavior embodies the perception of metrical structure. The article concludes with a discussion of the implications of our approach for theories of metrical structure and musical expectancy.

Connection Science, *6 (1)*, 177 - 208.

RESONANCE AND THE PERCEPTION OF MUSICAL METER

1 Introduction

> Embodied musical meaning is, in short, a product of expectation. ... If this hypothesis is correct, then an analysis of the process of expectation is clearly a prerequisite for the understanding of how musical meaning, whether affective or aesthetic, arises in any particular instance. (Meyer, 1956)

Meyer proposed that *expectation* is the key to understanding human intellectual and emotional response to music. Through artful patterning of the acoustic environment, composers and performers evoke expectations in their listeners. They skillfully manipulate these expectations, satisfying some and frustrating others, to arouse both affective and intellectual responses. Meyer argued that this is the property of musical experience that enables artistic communication. Since his proposal, a number of theorists have adopted Meyer's basic point of view, each exploring various types of expectancy in music perception. Simon and Sumner (1968) provide an analysis of music perception as a sequence extrapolation task, one in which the listener attempts to predict what patterns will follow based on analysis of the current pattern context. Narmour's implication-realization theory (1990) focuses on the innate expectancies that arise in response to the basic properties of individual melodic intervals and chains of melodic intervals. While these and other theoretical approaches differ in many important respects, one facet they share is a central concern with the musical question: "What Next?"[1].

Since time is the primary medium of musical communication, however, we cannot adequately characterize musical expectancy simply by considering *what* events a listener expects to occur. We must also take into account *when* a listener expects events to occur (Jones, 1981). In this regard, we find it useful to distinguish between sequence processing and temporal processing. In sequence processing, a system must predict the sequential ordering of future events ("What next?"). Thus, a sequence processing system must collect, organize, and use knowledge of sequential structure. On the other hand, a temporal processing system must predict when future events are likely to occur ("When next?") by exploiting knowledge of temporal structure.

Several connectionist models of music perception and production have focussed on issues of sequential structure using discrete-time recurrent neural networks trained with backpropagation (Bharucha & P. Todd, 1989; P. Todd, 1991; Mozer, 1991). Such architectures deal well with sequential information. In music, however, temporal organization includes periodic structure on multiple time scales and systematic expressive deviations from timing regularity. Simple discrete-time recurrent neural networks have difficulty capturing both forms of organization, thus, hampering their application to complex, temporally structured sequences such as music and speech (Cottrell, Nguyen, & Tsung, 1993; de Vries and Principe, 1992; Mozer, 1991; Mozer, 1993; P. Todd, 1991).

We believe that a more direct approach is called for. In order to address the problem of temporal structure in music, we focus on the perception of metrical structure: the perceived temporal structure of musical patterns that manifests itself phenomenologically as a sense of alternating strong and weak beats. Metrical structure provides listeners with a temporal framework upon which to build expectations for events. These expectations dramatically affect human perception, attention, and memory for the complex event sequences found in music (Jones & Boltz, 1989; Palmer & Krumhansl, 1990; Povel & Essens, 1985). It has been proposed (Jones, 1976; 1987; Jones & Boltz, 1989) that the perception of rhythm is a dynamic process in which the temporal organization of external musical events synchronizes, or entrains, a listener's internal rhythmic processes. Due to the absence of plausible mechanistic accounts, however, many implications of this theoretical position remain unclear. This article introduces a mathematical model of entrainment appropriate for modeling the perception of metrical structure. We present the model as a single abstract processing unit, amenable to connectionist implementation. This oscillatory unit phase- and frequency- locks to a single periodic component of a rhythmic pattern, embodying the notion of musical pulse, or beat. Models of meter perception will require interconnected networks of these units, whose self-organizing response to an incoming rhythmic pattern embodies a dynamic "perception" of metrical structure. We will analyze the behavior of both individual units and collections of units with the goal of understanding how to construct such a network.

1.1 Connectionist Approaches to Musical Expectancy

One approach to the problem of modeling musical expectancy equates event expectation with time series prediction (Dirst & Weigend, 1993). Connectionist approaches often employ recurrent networks (Elman, 1990; Jordan, 1986; Port, 1990), trained to predict the next event of a sequence, given a memory of past events. Bharucha and P. Todd (1989), for example, proposed a connectionist model of musical expectation to predict chords in a sequence using a recurrent neural network that stored previous sequence elements. These elements provided the necessary context to predict the next chord. This approach is appealing for a number of reasons. First, recurrent networks are simple; they process sequences one event at a time and assume no complex control mechanism. Second, recurrent networks fix no a priori limit on the size of the context that is used for prediction. Finally, recurrent neural networks can also be used for musical composition. By connecting the network's output units to its input units, the network can generate novel sequences that reveal what it has learned about musical structure (P. Todd, 1991; Mozer 1991).

Despite these incentives to implement general musical expectation engines using recurrent networks, certain problems arise in the processing of complex, temporally structured sequences such as music. First, the ability of discrete-time recurrent neural networks to learn and/or make use of temporal context information is appears to be limited (de Vries & Principe, 1992; Mozer, 1993). Specifically, recurrent networks have difficulty capturing relationships that span long temporal intervals, as well as relationships that involve very high order statistics (Mozer, 1993). Unfortunately, these are the sorts of relationships that a model must capture in order to adequately model musical expectancy or music composition (P. Todd, 1991; Mozer, 1991). Second, recurrent networks generalize poorly to novel presentation rates, relying upon absolute rate information to recognize temporal patterns (Cottrell, Nguyen, & Tsung, 1993). A network trained to recognize a melody played at 80 beats per minute, for example, may not recognize the same melody played at 90 beats per minute. McGraw, Montante, and Chalmers (1991) attempted to train various recurrent networks as simple “beat detectors”, but found that a network trained on one melody at three different tempos may not correctly respond to the same melody played at a fourth, intermediate tempo. The problems of temporal context and absolute rate dependence are symptoms of insensitivity to temporal structure. Musical rhythms display complex forms of temporal organization that listeners abstract and use to process complex musical event sequences. However, discrete-time recurrent neural networks fail to take advantage of this temporal information.

In this article we focus on issues of temporal structure in music, although our results may pertain to the processing of other complex, temporally structured sequences as well. The remainder of this paper is organized as follows. In the next section, we outline the problem of temporal structure in music. We review music-theoretic notions of rhythm and meter, and some recent literature on human rhythm perception. We then discuss current connectionist models related to meter perception and highlight some problems with these models. We introduce the mathematical concepts that play an important role in our theory, and then propose a mathematical model of entrainment appropriate for modeling aspects of human rhythm perception. We present the model as a single abstract connectionist processing unit. We assume that it will be necessary to compose a network of these abstract oscillatory units to model the perception of meter. We provide analyses of the model, revealing implications for theories of meter perception. Finally, we return to the problems of recurrent neural networks, connectionist time series prediction, and musical expectancy in general.

2 Rhythm and Meter

In this section, we explore issues related to the temporal structure of music and the processing of rhythmic sequences. We review music-theoretic notions of rhythm, beat, and meter, concentrating on recent cognitive proposals. We then review recent literature on human rhythm perception, highlighting the role of temporal structure in music perception. Finally, we discuss existing approaches to modeling beat tracking, entrainment, and meter perception.

2.1 Music-theoretic perspective

The term *rhythm* refers to the general sense of movement in time that characterizes our experience of music (Apel, 1972). Rhythm often refers to the organization of events in time, such that they combine perceptually into groups or induce a sense of meter (Cooper & Meyer, 1960; Lerdahl & Jackendoff, 1983). In this sense, rhythm is not an objective property of music, it is an experience that has both objective and subjective components. The experience of rhythm arises from the interaction of the various materials of music – pitch, intensity, timbre, and so forth – within the individual listener.

We will use the term rhythm in a second, more restricted sense – to refer to an objective component of musical rhythm. When we speak of “a rhythm” or “a rhythmic pattern” we will mean the pattern of inter-onset durations associated with a music sequence (Dowling & Harwood, 1986; Jones, 1987b). In music, a rhythm has an associated pattern of *phenomenal accents*, which is the physical patterning of events in the musical stream such that some seem to be stressed relative to others (Lerdahl and Jackendoff, 1983). Phenomenal accent can be conferred upon an event by the manipulation of many possible physical variables, including duration, pitch, and intensity.

By *beat*, we mean one of a series of perceived pulses marking subjectively equal units in the temporal continuum. Although the sense of beat is generally established and supported by objectively occurring musical events, beat is a subjective experience. Once a sense of beat has been established, it continues in the mind of the listener, even after the supporting stimulus has ceased. The experience of beat is necessary for the experience of meter (Cooper & Meyer, 1960). The term *tempo* refers to the rate (beats per unit time) at which beats occur. In this paper, we will generally refer to the reciprocal measure, the *beat period*, or the span of time between consecutive beats.

Meter, as it is traditionally defined, refers to the measurement of the number of beats between more or less regularly recurring accents. (Apel, 1972; Cooper & Meyer, 1960). There are two important implications here. First, in order for meter to exist, the listener must feel some beats to be accented relative to others. Accented beats are called *strong*, unaccented beats are called *weak*. Although phenomenal accents may correspond to strong and weak beats, metrical accent is subjective. Early research on the perception of rhythm indicated that even isochronous, unaccented pulse trains may elicit the experience of alternating strong and weak beats, a phenomenon called "subjective rhythm"[2] (Bolton, 1894; Woodrow, 1909). Second, the experience of *meter* implies the existence of at least two recurrent periodicities, describable as two separate levels of beats with related beat periods. Integer ratios usually characterize the beat period relationship (2:1 or 3:1 for example), so that meter is said to describe a nested grouping of beats (Lerdahl & Jackendoff, 1983).

Meterical organization usually exists on multiple time scales (Cooper & Meyer, 1960; Lerdahl & Jackendoff, 1983). Lerdahl and Jackendoff (1983) have proposed a construct that describes the temporal organization of a piece at all relevant metrical levels, called a *metrical structure*. The metrical structure of a piece can be transcribed as a grid (Figure 1). According to this notation, each horizontal row of dots represents a level of beats, and the relative spacing between dots of adjacent levels captures the relationship between the beat periods of adjacent levels of beats. A metrical structure describes one of the most important subjective components of rhythmic experience: the feeling of regularly recurring strong and weak beats called *metrical accent* (Lerdahl & Jackendoff, 1983). Points of metrical accent are captured, using a metrical grid, as temporal locations where the beats of many levels coincide. Points where many beats coincide are (subjectively) felt as stronger; points where few beats coincide are felt as weaker.

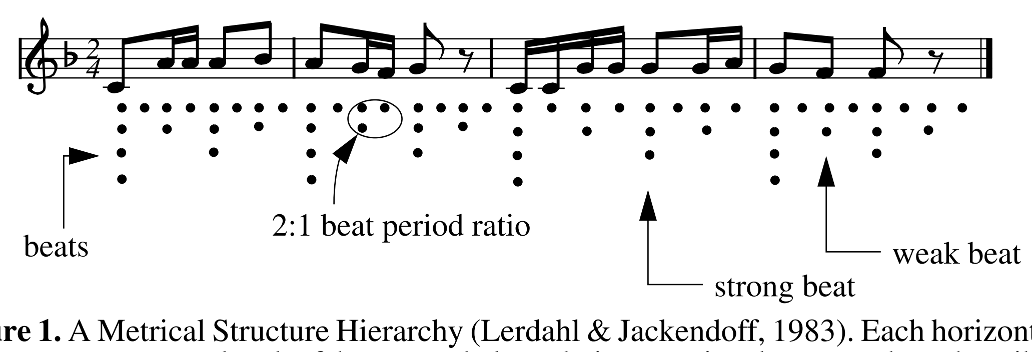

Figure 1. A Metrical Structure Hierarchy (Lerdahl & Jackendoff, 1983). Each horizontal row of dots represents a level of beats, and the relative spacing between dots describes the relationship between the beat periods of adjacent levels. Points where beats of many levels align describe points of metrical accent.

Lerdahl and Jackendoff's (1983) proposal describes certain aspects of music perception and cognition. A rhythm, with its pattern of phenomenal accent, functions as a perceptual "input" to metrical accent. Although phenomenal accent information may be missing or ambiguous, moments of musical stress in the raw signal are thought to serve as cues from which the listener may extrapolate a regular pattern of metrical accents (Lerdahl & Jackendoff, 1983). Lerdahl and Jackendoff (1983) have proposed a generative theory for the perception of metrical structure that is expressed as two sets of rules. A set of *well-formedness rules* describes legal metrical structure hierarchies. These rules restrict metrical structures to strictly nested hierarchies with beat-period ratios of either 2:1 or 3:1. Next, a set of *preference rules* describes which legal metrical structure an experienced listener would actually perceive for a given rhythmic pattern. These rules are concerned mainly with the placement of strong beats, as determined by the alignment of beats at adjacent levels in the metrical structure hierarchy.

Theories of metrical structure, such as the generative theory we have just described, have some limitations. First, the characterization of metrical structure as a hierarchical nesting of beats limits the scope of the theory. Only some music can be described in this way. Lerdahl and Jackendoff (1983) explicitly restrict their theory to Western tonal music of the common practice period. Much non-Western music, as well as contemporary Western Art music, Jazz, and popular music make use of *dissonant* rhythmic structures (Yeston, 1976), known as *polyrhythms*. A polyrhythmic relationship between two levels of beats is a relationship of beat-periods such that N beats at one level occupy the same amount of time as M beats at the next level. *Rational* ratios N:M, such that the integers N and M are relatively prime (3:2, 4:3, 5:4, and so forth), characterize polyrhythmic ratios. Hierarchical nestings do not adequately capture polyrhythmic structures, thus it is more general to think of metrical structures as being composed of layers, or *strata*, of beats at different time scales (Yeston, 1976).

A second limitation of current theories of meter is that they fall short of adequately explaining perception. Theories of metrical structure, as discussed above, apply to musical time as notated. It is well established, however, that musicians never perform rhythms in a perfectly regular, or mechanical, fashion. Instead, performers produce sound patterns that reveal both intentional and unintentional timing variability (Sloboda, 1983; Clarke, 1985; Shaffer, Clarke, & N. Todd, 1985; N. Todd, 1985; Palmer, 1989; Drake & Palmer, 1993). Current theories of metrical structure do not explain how listeners are able to perceive meter in rhythms that performers actually play (unless the performer is a computer). As we shall see, this is no small problem.

In summary, theories of metrical structure attempt to describe the perceived temporal organization of rhythmic patterns. A metrical structure is composed of layers, or strata, of beats that align with the onset of musical events. Theories of metrical structure address issues related to the beat period ratio and the relative alignment between adjacent levels of beats. Theories that require the layering of beats to describe a strictly nested hierarchy, however, are limited in scope. In order to include the polyrhythmic structures common in many forms of music, more complex relationships between adjacent levels must be allowed. Finally, because traditional theories do not deal with the issue of timing variability in music performance, they stop short of explaining the perception of metrical structure. As we shall see below recent psychological results implicate a class of mechanisms capable of linking traditional theories with the perception of metrical structure.

2.2 Psychological Perspective

Research into the human processing of complex, temporally structured sequences has provided some of the most intriguing results in the study of human perception and cognition. The temporal structure of sequences has been shown to dramatically affect human abilities to perceive, remember, and reproduce serial patterns. Recent results support theoretical proposals that implicate an entrainment response as one of the basic processes of human rhythm perception. Here we review some of the relevant psychological results. In the next section, we will explore in more detail mechanistic accounts that have been proposed to account for the perception of metrical structure.

Abstract knowledge of metrical structure has been shown to affect memory for temporal information in auditory sequences. In one study, memory for pitch sequences was found to be dependent on a perceived temporal frame. Pitch structures that coincided with temporal structures enhanced recall, while pitch structures that conflicted with temporal structures negatively affected recall (Deutsch, 1980). In a related finding, memory confusions of temporal patterns in a discrimination task were found to be consistent with a music-theoretic metrical structure hierarchy (Palmer & Krumhansl, 1990). Other studies have demonstrated similar memory constraints, by showing that the reproducibility of rhythms is affected by the patterns of phenomenal accentuation in the to-be-reproduced rhythm. The evidence suggests that sequences of events that imply a metrical organization are easier to memorize and reproduce than sequences lacking such organization (Essens & Povel, 1985; Povel & Essens, 1985).

These and related findings are often cited as evidence that listeners represent and/or remember rhythms in terms of metrical structure hierarchies. Essens and Povel (1985) have hypothesized that in perceiving a temporal pattern, listeners induce an internal clock that is subsequently used as a measuring device to code the structure of a temporal pattern. Rhythmic sequences are encoded in memory with respect to this clock, so that patterns that correspond well with an induced clock (metrical patterns) can be represented using simpler memory codes, and are therefore easier to remember and reproduce. Jones (1976; 1987) and Jones & Boltz (1989) offer a more comprehensive interpretation. They argue that the organization of perception, attention, and memory is inherently rhythmical. Music (and other rhythmic stimuli) *entrain* listeners' perceptual "rhythms", and these rhythms embody *expectancies* for when in time future events are likely to occur. Expectancies in turn guide *anticipatory pulses of attention* that facilitate perception of events that occur at expected points in time.

One source of evidence for the temporal expectancy hypothesis stems from studies that directly test listener attention rather than listener memory. These studies show that temporal pattern structure constrains the ability of subjects to attend to melodic sequences. For example, regularity of phenomenal accent placement has been shown to affect listeners' abilities to judge the temporal order of tones in a sequence (Jones, Kidd, & Wetzel, 1981). Listeners are also better able to identify pitch changes in sequences when these changes occur at points of strong metrical accent (Jones, Boltz, & Kidd, 1982). Additional evidence suggests that listeners' implicit knowledge of meter (beyond immediate sensory context) contributes to the perception of temporal sequences. Listeners' goodness-of-fit judgements for events presented in metrical contexts were shown to be consistent with multi-leveled metrical structure hierarchies (Palmer & Krumhansl, 1990).

Another source of evidence for the temporal expectancy hypothesis comes from psychophysical studies of time perception. It appears that the temporal structure of auditory patterns actually affects humans' abilities to perceive time. For inter-onset durations corresponding roughly to musical time scales, it can be shown that the ability to detect differences in temporal intervals approximately obeys Weber's law (Getty, 1975; Halpern & Darwin, 1982). That is, when subjects are asked to compare two intervals, the accuracy of their time discrimination judgement is related to the base length of the interval they are asked to judge. Adherence to Weber's law breaks down under certain circumstances, however. Temporal difference judgements improve as the number of reference intervals increases (Schulze, 1989; Drake & Botte, 1993). It has also been shown that sensitivity to time changes in sequences is best for metrically regular sequences (Yee, Holleran, & Jones, 1994), and that sensitivity to tempo changes degrades with the regularity of the stimulus (Drake & Botte, 1993). Some researchers have suggested that these results indicate perceptual synchronization of the listener to a perceived beat (e.g. Schulze, 1989; Yee, Holleran, & Jones, 1994).

In our view, the psychological literature offers strong support for the temporal expectancy hypothesis. In addition, the literature on motor coordination reveals a number of activities, including rhythmic hand movements and cascade juggling, to be consistent with mathematical laws governing coupled oscillations (e.g. Kelso & deGuzman, 1988; Schmidt et al., 1991; Treffner & Turvey, 1993; for a review of recent models see Beek et al., 1992). Shaffer (1981) has proposed that the performance of two-handed polyrhythms in music may be described as the entrainment of clocks. However the experimental literature often stops short of proposing specific mechanisms of entrainment. In the next section, we explore specific proposals that have been made relating to the perception of metrical structure.

2.3 Connectionist perspective

Metrical structure plays an important role in the organization of human perception. However, mechanisms for the perception of metrical structure are still poorly understood. Symbolic approaches relying on the parsing of temporal patterns have been proposed (e.g. Longuet-Higgins & Lee, 1982; Scarborough, Miller, & Jones, 1992), but like the generative theories of metrical structure upon which they are based, they fail to explain the perception of meter in musical performance. Entrainment, or synchronization to a perceived beat, may provide some answers. Connectionists, however, are only beginning to appreciate the power of this approach. We will discuss a number of models related to the perception of metrical structure, including previous connectionist approaches to entrainment, illustrating the problems entailed by the design of entrainment mechanisms for the perception of complex musical rhythms.

Scarborough, Miller, & Jones (1992) have described a model of meter perception called BeatNet, based on a parallel constraint satisfaction paradigm. Conceptually, the BeatNet network is a one-dimensional array of idealized low-frequency oscillators with different beat-periods that operate to align their output "ticks" with event onsets, producing a metrical grid of the style proposed by Lerdahl and Jackendoff (1983). A metrical structure emerges from local interactions between oscillators, rather than from the global effect of rule-based analysis. An advantage of this approach is that it handles the problem of metrical preferences through real time processing constraints, rather than by global evaluation of alternative constructs. This approach cannot deal with timing variability, however, because of the simplifying assumption of "idealized" oscillatory units.

According to Desain & Honing (1991), the problem of timing variability is the key problem for mechanistic accounts of meter perception. From this point of view, the relevant task is one of inferring from the inter-onset intervals that the performer *creates*, what inter-onset intervals the performer *intended* – a process called *quantization*. Desain and Honing have developed a connectionist quantizer to "clean up" messy timing data so that the meter may be inferred. The quantizer works to adjust durations so that every pair of durations is adjusted toward an integer ratio, if it is already close to one. A disadvantage of this approach is that it relies on a fixed input window, whose size may need to be adjusted depending on the input (Desain & Honing, 1991). Recently Desain (1992) extended this approach to present a theory of complex temporal expectancy.

An alternative approach to the problem of timing variability relies on a form of entrainment called beat-tracking (Allen & Dannenberg, 1989; Dannenberg, 1984; Dannenberg & Mont-Reynaud, 1987). This approach does not assume that the beat-period is static, rather the length of a beat is adjusted throughout the performance as the performer speeds up or slows down. Results reported to date indicate that this task is surprisingly difficult (Allen & Dannenberg, 1989). Longuet-Higgins and Lee (1982) have modelled the perception of meter as the parsing of inter-onset durations. Longuet-Higgins (1987) proposes a hybrid method that combines beat-tracking with metrical structure parsing. The program uses a static tolerance window, within which it will treat any onset as being "on the beat". Onsets which fall outside the window are interpreted as subdividing the beat (either into groups of two or three). This approach may solve some of the problems inherent in tracking a single level of beats, because separate mechanisms track separate levels of beats, reducing the potential for confusion.

Page (1993) acknowledges the problem that temporal structure poses for connectionist approaches to modeling music perception. He also suggests that entrainment may provide solutions to some of these problems, and proposes that a neural entrainment mechanism should operate analogously to a *phase-locked loop*, an electronic circuit commonly used in communications applications. We agree that entrainment mechanisms are important in addressing problems of temporal structure, however, the standard phase-locked loop not acceptable. Page's (1993) simulations detail the difficulties associated with this approach. The standard phase-locked loop design assumes that the input signal is periodic. This assumption places limitations on the circuits's ability to deal with the complex rhythmic structures of music. Because the phase-locked loop reacts to every input event, it cannot extract a "component periodicity" from a complex rhythmic pattern. Page attempts to deal with this problem by assuming that relevant periodicities are unambiguously marked in the signal via phenomenal accent information. In music, however, phenomenal accent information is often missing, ambiguous, or even misleading (e.g. syncopation).

Other connectionists have proposed entrainment mechanisms for meter perception as well (Large & Kolen, 1993; McAuley, 1993; 1994). An important research problem is to determine an appropriate type of oscillator for modeling musical beat. To illustrate the relevant issues, we begin with a simple model that has been used as a model of single cell oscillation in the nervous system, the integrate-and-fire oscillator (Glass & Mackey, 1988; Winfree, 1980). The simplest formulation of the integrate-and-fire model is shown in Figure 2. Activation increases (linearly) to a threshold, the unit “fires”, resets its activation to zero, and the process begins again. As can be seen in Figure 2A, the unit spontaneously oscillates with a period determined by the slope of the activation function and the height of the threshold. Figure 2B shows the unit phase-locking to a discrete periodic stimulus. Each discrete stimulus event temporarily lowers the unit’s threshold so that the oscillator may fire and reset earlier than would otherwise be the case. Figure 2B also illustrates one problem with phase-tracking oscillators as models of musical beat. When the stimulus ceases, or when an onset is missing, the oscillator immediately reverts to its original period, as though no stimulus had ever been present. In other words, the oscillator has no memory of the previous rhythmic context. Torras (1985) proposed a scheme for frequency tracking in a somewhat more complex integrate-and-fire model. In this formulation, an integrate-and-fire oscillator can phase-lock to a stimulus by adapting its threshold. This situation is shown for our simpler model in Figure 2C. McAuley (1993) proposed that a Kohonen map of Torras oscillators could memorize, categorize, and reproduce musical rhythms.

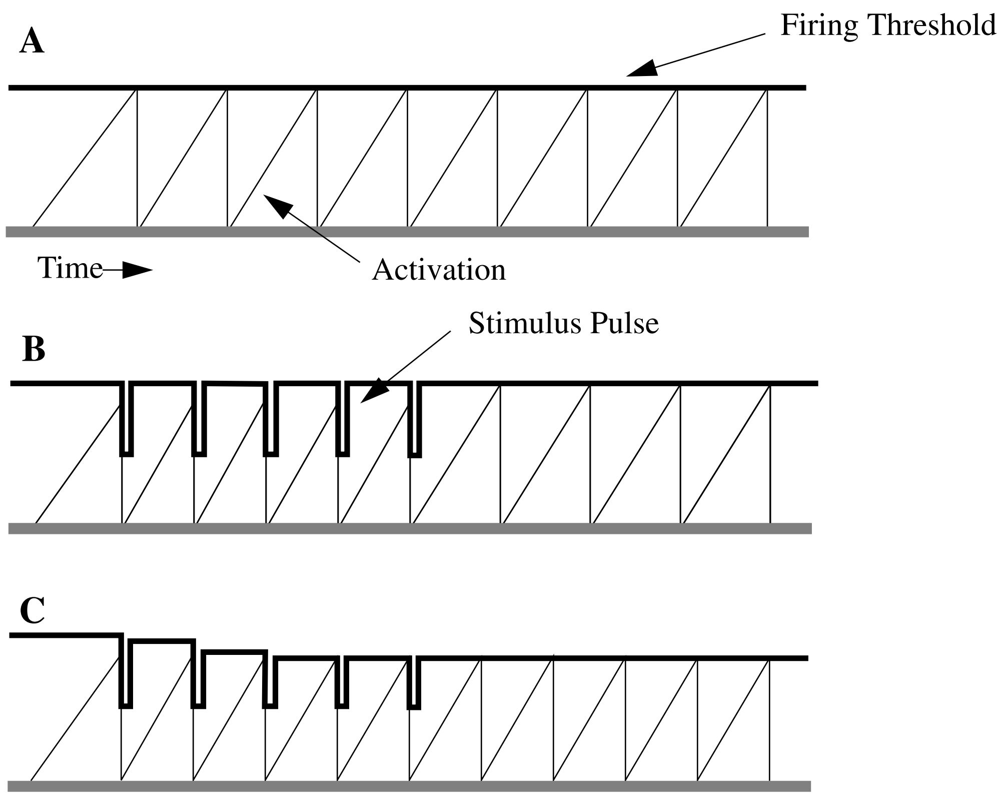

Figure 2. A Periodic Signal and the Response of an Integrate-and-Fire Oscillator
A) The oscillator in the absence of stimulation. When activation reaches the threshold, the oscillator “fires”. The period of the resulting oscillation is determined by the slope of the activation function and the height of the firing threshold. B) Phase-tracking. Discrete periodic stimulus effects the oscillator by lowering its firing threshold. The oscillator comes into phase and frequency lock with the periodic stimulus. The effect is temporary, however. When the stimulus is removed, the oscillator reverts to its intrinsic period. C) Frequency-tracking. By adjusting its firing threshold in response to stimulus, the unit may achieve permanent or semi-permanent frequency lock. When the stimulus is removed, the oscillator continues to fire at the stimulus period.

Integrate-and-fire units have their own set of problems in the domain of meter perception. For example, the discontinuity in the activation function constrains the oscillator to adjust its period only by speeding up (McAuley, 1994). We have proposed a continuous model (presented below, in a revised form) to avoid this problem, as well as the problems exhibited by phase-locked loop models (Large & Kolen, 1993). McAuley (1994) recently compared the performance of four different oscillatory units including two integrate-and-fire models, our earlier model (Large & Kolen, 1993), and a simplification of this model. McAuley (1994) prefers the simpler model, although in our view this simplification creates problems similar to those found in phase-locked loop models; both require strong assumptions about phenomenal accentuation to display appropriate behavior.

In summary, modeling the perception of metrical structure is difficult, in large measure because of problems arising from timing variability in musical performance. Entrainment remains an interesting possibility, despite the inadequacies of straight-forward approaches to entrainment such as phase-locked loop models. Entrainment models must have the ability to "pick" component periodicities out of a complex rhythmic pattern in spite of missing, ambiguous, or misleading phenomenal accent information. An entrainment model that provides such a capability would have important implications for theories of musical meter. We propose such a model below.

3 Mathematical Considerations

Musical rhythms afford the perception of a particular type of temporal organization called metrical structure. Both psychological evidence and connectionist analyses suggest that entrainment might serve as a useful tool in modeling this perception. The mathematics of entrainment describe many natural systems, and one of the goals of this paper is to add the perception of musical rhythm to this list. In this section, we briefly summarize some important mathematical concepts relevant to theories of entrainment and introduce the principles underlying our proposal. As we shall demonstrate, entrainment provides properties that map quite nicely onto the task of rhythm perception.

The swinging of a pendulum, the ticking of a metronome, and the firing of a neural pacemaker cell are examples of *oscillations*. Oscillations are periodic events – events that cycle, or repeat after some specific interval of time, called the *period* of the oscillation. Let us assume that the beginning of each cycle is identified by a discrete marker, and define the phase at this marker to be 0. Let us further assume that each cycle of the oscillation has intrinsic period T_0. We then define the *phase* at any time $0 < t < T_0$ to be

$$\phi = \frac{t}{T_0}.$$

As we define it here, phase lies between 0 and 1. Two oscillations are *synchronized* when they regularly come into phase, or begin their cycles together. A process by which two or more oscillators achieve synchronization is called *entrainment*. Entrainment occurs because a *coupling* between two or more oscillations causes them to synchronize. Coupling allows one oscillator to perturb another by altering its phase, its intrinsic period, or both.

One important type of entrainment is *phase-locking*. Phase-locking phenomena have been of interest in the connectionist community for some time, especially since the discovery of oscillations and synchronization behavior in the cat visual cortex (Eckhorn et al., 1989; Gray et al., 1989). It has been proposed that the oscillations of neurons in the cat visual cortex phase-lock to establish relations between features in different parts of the visual field (Gray, et al., 1989). It has further been suggested that the brain could be using synchronized oscillations as a general method of solving the binding problem (von der Malsberg & Schneider, 1986). Phase-locking may add an extra degree of freedom to neural network models, so that a number of different entities may be represented simultaneously using the same set of units, each by a different phase in an oscillatory cycle.

Our use of oscillatory units differs from that proposed in the literature on connectionist feature binding. First, rather than using coupled oscillations to describe a neural strategy for performing an implementation-level operation such as feature binding, we will use synchronization to describe how the brain may execute the relatively high-level cognitive function of meter perception. Consequently, the oscillatory units we propose will represent higher levels of neural abstraction than individual neurons. Second, we will be interested in dynamics that are more complex than 1:1 phase-locking. Therefore, we will need to take a moment to introduce the analytical tools used in subsequent sections.

Researchers since Poincaré have described entrainment phenomena using the mathematics of non-linearly coupled oscillators. The method we describe here assumes that the constituent processes are oscillatory, and the oscillations may be linear or non-linear. However, the coupling between the oscillators may exhibit various types of non-linearities. The Poincaré map, or circle map, summarizes the long term dynamics of a system of two oscillators. Consider the following mapping:

$$\phi_{i+1} = \phi_i + \frac{p}{q} + b \sin(2\pi\phi_i) .$$

This equation is a model circle map, called the *sine circle map*, that describes the dynamics of a system of two oscillators, a driving and a driven oscillator. The parameter q is the period of the driving oscillator, p is the period of the driven oscillator, and $b \sin(2\pi\phi_i)$ is a non-linear coupling term that describes the perturbations delivered to the period of the driven oscillator by coupling to the driver. ϕ_i is the phase of the driving oscillator at which the driven oscillator fires on iteration i. When b = 0 (no coupling), the behavior of the system is summarized by the ratio p/q, the so-called “bare-winding-number”. So, for example, if $p = 1$ and $q = 2$, the driven oscillator fires twice for each time the driver fires. As coupling strength, b, increases, another ratio, N:M, the so-called “dressed winding number”, describes the long-term dynamics of the system. In the dressed winding number, N is the period of the driven oscillator under the influence of coupling, and M is the period of the driver. If the coupling strength is high enough, even as p/q is perturbed away from 1/2, the system will still lock in a 1:2 relationship, because each time the driven oscillator fires, its phase is perturbed slightly by the coupling to the driving oscillator.

This locking behavior is highly structured. The dynamics of coupled systems like the sine circle map can be summarized in a *regime diagram*. Figure 3A shows a regime diagram for the sine circle map. The x-axis is the bare winding number, p/q, and the y-axis is coupling strength, b. The regime diagram identifies stable phase-locked states, also called attractors, mode locks, or resonances (Treffner & Turvey, 1993), for particular coupling strengths and driven/driver period ratios. The parameter regions that correspond to stable phase-locked states are known as Arnol’d tongues (Glass & Mackey, 1988; Schroeder, 1991). From this diagram we can see, for example, that for a bare winding number of 0.52, if coupling strength is high enough, the system will still phase-lock in a 1:2 relationship. We have labeled each tongue with a ratio corresponding to its locking mode. The width of each "tongue" reflects the stability of the corresponding mode lock for a given coupling strength, i.e. its sensitivity to noise in the p/q ratio. For example Figure 3 shows that, for a fixed coupling strength, 1:1 entrainment is more stable than 1:2 entrainment, which is more stable than 2:3 entrainment, and so forth. Depending upon the coupling strength, it can be shown that entrainment is possible at any frequency ratio, N:M where N and M are relatively prime integers (Glass & Mackey, 1988). The regime diagram is not arbitrarily organized. Rather, its structure can be summarized by a mathematical construct known as the Farey tree (Figure 3B). The Farey tree enumerates all rational ratios according to the stability of the corresponding mode lock in the coupled system. Its branching structure corresponds the structure of the Arnol'd tongues of the sine circle map, as well as to known bifurcation routes in other mathematical and natural systems (Schroeder, 1991).

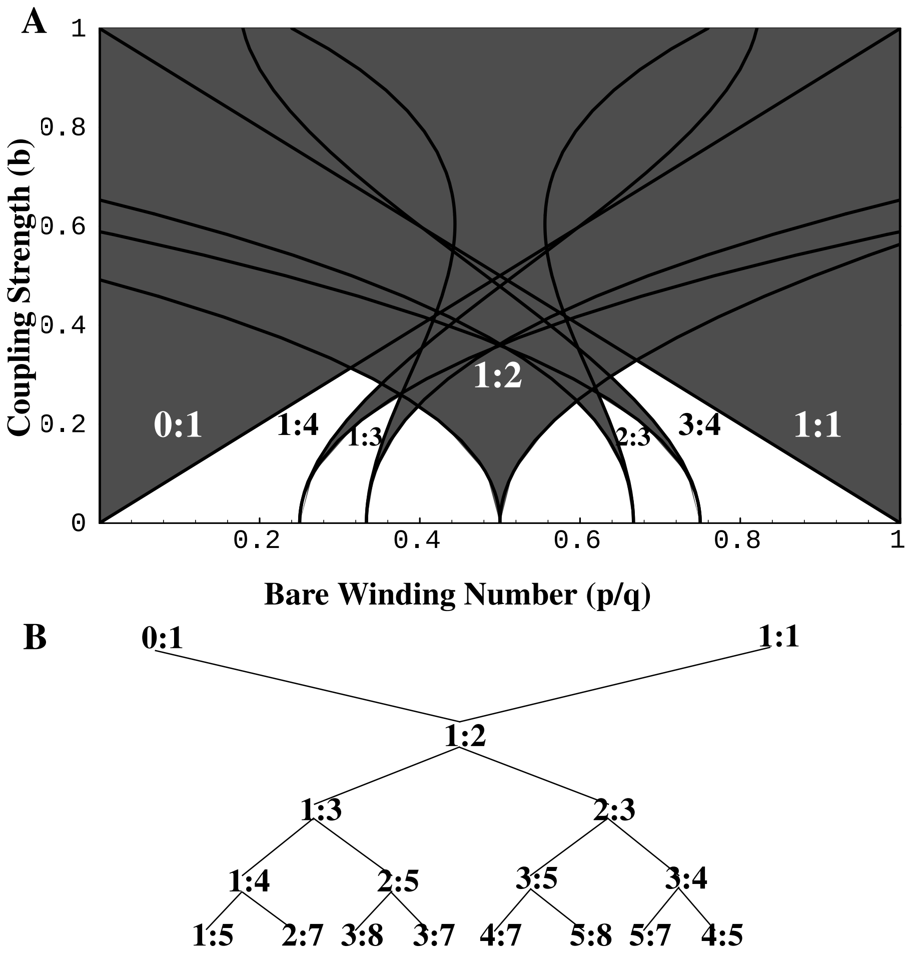

Figure 3. Entrainment. A) A regime diagram. The dynamics of a system of coupled oscillators may be summarized in a regime diagram. The parameter regions that correspond to mode-locked states are known as Arnol'd tongues (shaded). The width of each resonance tongue reflects the stability of the corresponding mode lock. For example, 1:1 entrainment is more stable than 1:2 entrainment, is more stable than 3:2 entrainment, and so forth. B) The Farey tree. The Farey tree is a mathematical object that summarizes the structure of the regime diagram. It provides an enumeration of all rational ratios according to the stability of the corresponding mode-lock in the coupled system. Its branching structure corresponds to known bifurcation routes in both mathematical and natural systems.

In phase-tracking systems, the frequency of the driven oscillator is altered because its phase is perturbed in every cycle. When the effect of the driving oscillator is removed, even for one cycle, the driven oscillator reverts to its intrinsic period. When the driver returns, a number of cycles may be required to reestablish phase lock. This behavior is unacceptable for the present purposes. In musical rhythms events do not necessarily occur on every beat. Thus, musical beat cannot be adequately modelled simply as phase-tracking entrainment. In order to model beat, the oscillator must somehow identify and "remember" the beat period. One way to do this is to allow frequency-tracking. *Frequency-tracking* entrainment occurs when coupling allows the driving signal to perturb the intrinsic period of the driven oscillator. A frequency-tracking oscillator can model musical beat because when the driving signal is removed, the oscillator continues at the driver's frequency, "expecting" the driver's eventual return.

In the model of musical beat that we propose, the driving signal (a rhythmic pattern) perturbs both the phase and the intrinsic period the driven oscillator, causing a (relatively) permanent change to the oscillator's behavior. In addition, the oscillator will adjust its phase and period only at certain points in the rhythmic pattern, effectively isolating a single periodic component of the incoming rhythm. With these assumptions, it will be possible to model the perception of metrical structure as a self-organizing process. What looks like a single macroscopic temporal pattern, a metrical structure, may emerge as the collective consequence of mutual entrainment among many constituent processes.We propose a network of oscillators of various native periods that entrain simultaneously to the periodic components of a rhythmic signal at different time-scales, and to the outputs of one another. We will use regime diagrams to analyze the mode-locking behavior of individual units, and we will examine the response of a system of units to an improvised musical performance, revealing the implications of this proposal for the perception of metrical structure.

4 An Entrainment Model

In this section, we present a model of entrainment that is suitable for explaining aspects of the perception of metrical structure. First, we develop the basic oscillatory unit, a unit capable of locking onto and tracking a single periodic component of a driving rhythm, to embody the musical concept of *beat*. Next, we analyze the dynamics of the unit using regime diagrams in order to understand its response to periodic stimulation. Finally, we examine aspects of the ability of a system of oscillators to self-organize a response to rhythmic patterns, embodying a dynamic "perception" of metrical structure. We discuss how our approach deals with problems that we have identified for theories of metrical structure. Finally, we address issues of neural implementation.

4.1 The Basic Oscillatory Unit

The basic unit has periodic output, and adjusts both its phase and period so that during stimulation the unit's output pulses become phase- and frequency-locked to a stimulus. The stimulus consists of a series of discrete pulses, $s(t)$, corresponding to the onset of individual events (e.g. notes). Event onsets may be derived from an acoustic representation of signal intensity (Marr, 1982; N. Todd, 1994), or alternatively, onsets may be extracted from a list of MIDI events. In this article, we assume that $s(t) = 1$ at the onset of an event, and 0 at other times.

The activation function of the unit is periodic:

$$a(t) = \cos\frac{2\pi}{p}(t - t_0) - 1, \tag{1}$$

where t is time, p is the period of the oscillation, and $t - t_0 \pmod p$ is the phase. The output of the unit is given by:

$$o(t) = 1 + \tanh(\gamma a(t)), \tag{2}$$

where γ is the output gain.

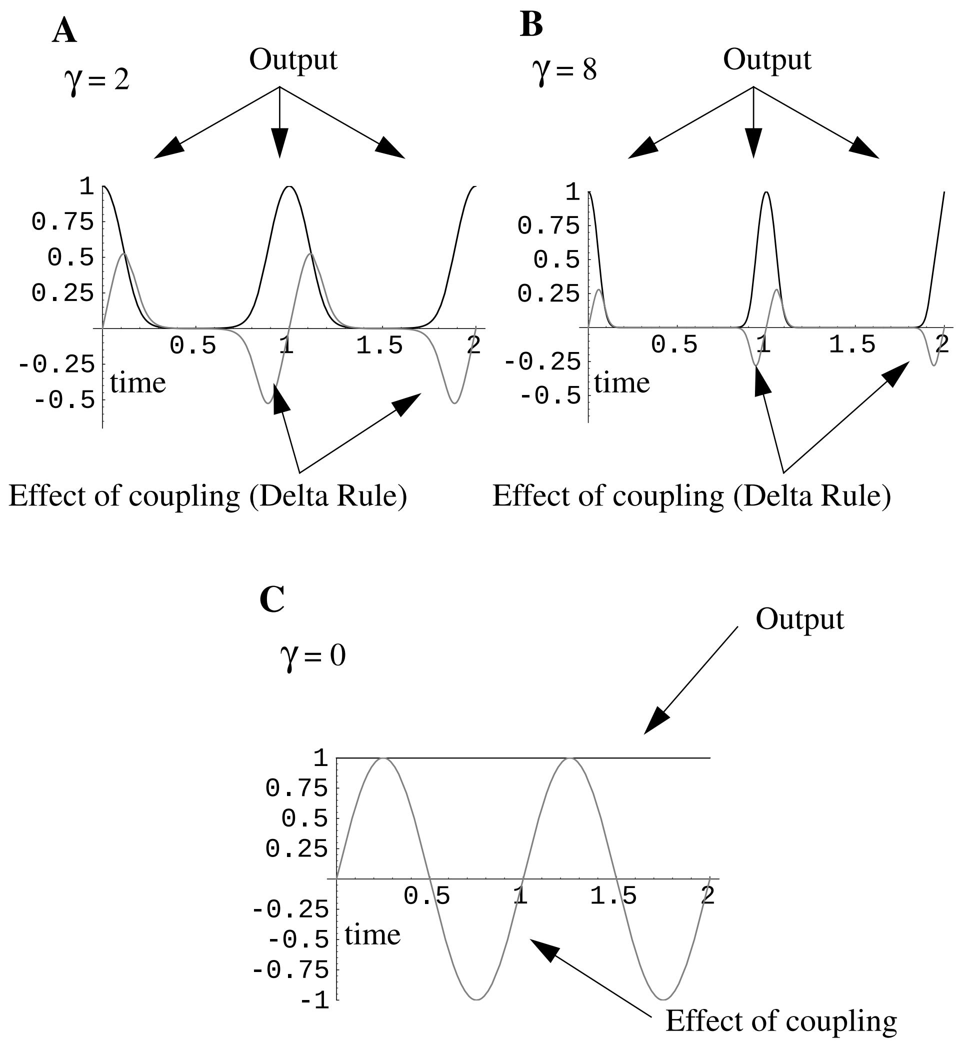

Figure 4. Schematics for the phase- and frequency- locking oscillator for different values of γ. The dark curves represent the output of the oscillator. The light curves summarize the effect of coupling on the phase and period of the driven oscillator. As gamma increases, the temporal receptive field shrinks. A) $\gamma = 2$. B) $\gamma = 8$. C) $\gamma = 0$.

Figures 4A and B show the output of the unit, in the absence of input, as a function of time. Output strength is maximum ($o(t) = 1$) at the beginning of each cycle (i.e. phase is 0), quickly falls to zero for the body of the cycle, then begins to rise again to a maximum as the cycle comes to a close. Output is only non-zero for a relatively small portion of the cycle, which we refer to as an *output pulse*. An output pulse defines a temporal receptive field for the unit, a region of *temporal expectancy.* After a unit has entrained to an input pattern, its output pulse marks a window of time during which it "expects" to see a stimulus pulse. As the unit entrains to stimulus pulses, it responds (i.e. adjusts its phase and period) only to pulses that occur within this temporal expectancy region; it ignores stimulus pulses that occur outside of this region. The parameter γ, the output gain, determines the width of this field. When γ is small, as shown in Figure 4A, the region is wide and temporal expectancy is relatively unfocused. When γ is large, as shown in Figure 4B, the region is narrow and temporal expectation is highly focused.

The unit entrains to the stimulus using a modified gradient descent procedure. That is, the unit adjusts its phase and period in such a way as to minimize an error function that measures the difference between when the unit maximally expects event onsets to occur, and when onsets actually do occur. Changes to phase and period are proportional to the partial derivative of the error function with respect to t_0 and p, respectively. We define the following error function:

$$E(t) = s(t)\,(1 - o(t))\,. \tag{3}$$

$E(t)$ has a non-zero value only when a stimulus is present ($s(t) = 1$), and single minimum value when the output strength is maximum ($o(t) = 1$). Therefore, minimization of the error function for both phase and period in response to a periodic input signal implies that the unit will act to align its points of maximum expectancy with the discrete event onsets in the signal.

To implement phase-tracking behavior, we minimize the error function by gradient descent on t_0. This yields the following delta rule for phase:

$$\Delta t_0 = -\eta_1 s(t) p \operatorname{sech}^2 \gamma a(t) \sin\frac{2\pi}{p}(t - t_0)\,, \tag{4}$$

where η_1 is the coupling strength (similar to b in the sine circle map), incorporating constants from the actual derivative of the error function. The factor p has been added to the "true" delta rule as a convenience. This has a scaling effect, so that this delta rule yields the same *proportional* phase adjustment regardless of the actual period of the oscillator. The light curves of Figures 4A and B show the shape of this curve, summarizing the effect of the delta rule in relation to the oscillator output. A stimulus pulse that occurs within the unit's expectancy region, but before the maximum of the output function causes a negative phase shift, because $\Delta t_0 < 0$. A stimulus pulse after the maximum of the output function causes a positive phase shift, because $\Delta t_0 > 0$. Thus, this delta-rule provides a non-linear coupling term implementing phase-locking entrainment. When $\gamma = 0$, as shown in Figure 4C, $\operatorname{sech}^2 \gamma a(t) = 1$, and the delta rule becomes the sine function. This is a significant special case because in this case the unit's phase-locking dynamics will become equivalent to those of the sine circle map. We will discuss this in more detail in the next section.

The preceding equations implement phase-tracking through a modified gradient descent strategy. We achieve frequency-tracking behavior using a similar strategy. For frequency-tracking, however, it is useful to limit the period of the oscillator to a fixed range between p_{min} and p_{max}. One way to do this is to introduce a frequency control parameter, Ω, according to the following relationship:

$$p = p_{min} + 0.5\,(p_{max} - p_{min})\,(1 + \tanh\Omega)\,. \qquad (5)$$

When Ω=0, then p takes on a value half way between p_{min} and p_{max}, and we refer to this as the *resting period* of the oscillator

$$p = \frac{p_{min} + p_{max}}{2}.$$

Because Ω determines p, we minimize the error function by gradient descent on Ω to implement frequency-tracking behavior. The change of Ω is proportional to the partial derivative of the error function with respect to Ω, which yields the following delta rule:

$$\Delta\Omega = -\eta_2 s(t)\,\mathrm{sech}^2\gamma a(t)\,\sin\frac{2\pi}{p}(t - t_0)\,\frac{\partial p}{\partial\Omega}, \qquad (6)$$

where η_2 is the coupling strength for frequency-tracking. Like the phase-tracking rule, this rule does not implement "true" gradient descent; we have taken some liberties to ensure quick, stable convergence over a range of frequencies. Note also that this delta rule is similar to the delta rule for phase-tracking, except for the term $\frac{\partial p}{\partial\Omega}$. Because of the similarity between the two delta rules, the shape of the light curves of Figures 4A and B also summarize the effect of the frequency-tracking delta rule. A stimulus pulse that occurs within the unit's receptive field, but before the maximum of the output function causes the unit to shorten its period, because $\Delta\Omega < 0$, whereas a stimulus pulse after the maximum of the output function causes the unit to lengthen its period, because $\Delta\Omega > 0$.

Figure 5 shows the output behavior of a unit with p_{min} = 600 ms, p_{max} = 800 ms, γ = 8, $\eta_1 = 0.2$, and $\eta_2 = 4.0$ exposed to a stimulus with a period of 660 ms. The oscillator initially fires at its resting period, $p = 700$ms. In response to input, it adjusts its phase and period so that it becomes synchronized to the stimulus within a few cycles. When the stimulus is removed, the oscillation continues with a period of 660 ms. As described above, the oscillation at this new period may be said to embody an "expectation" for events at these particular future times.

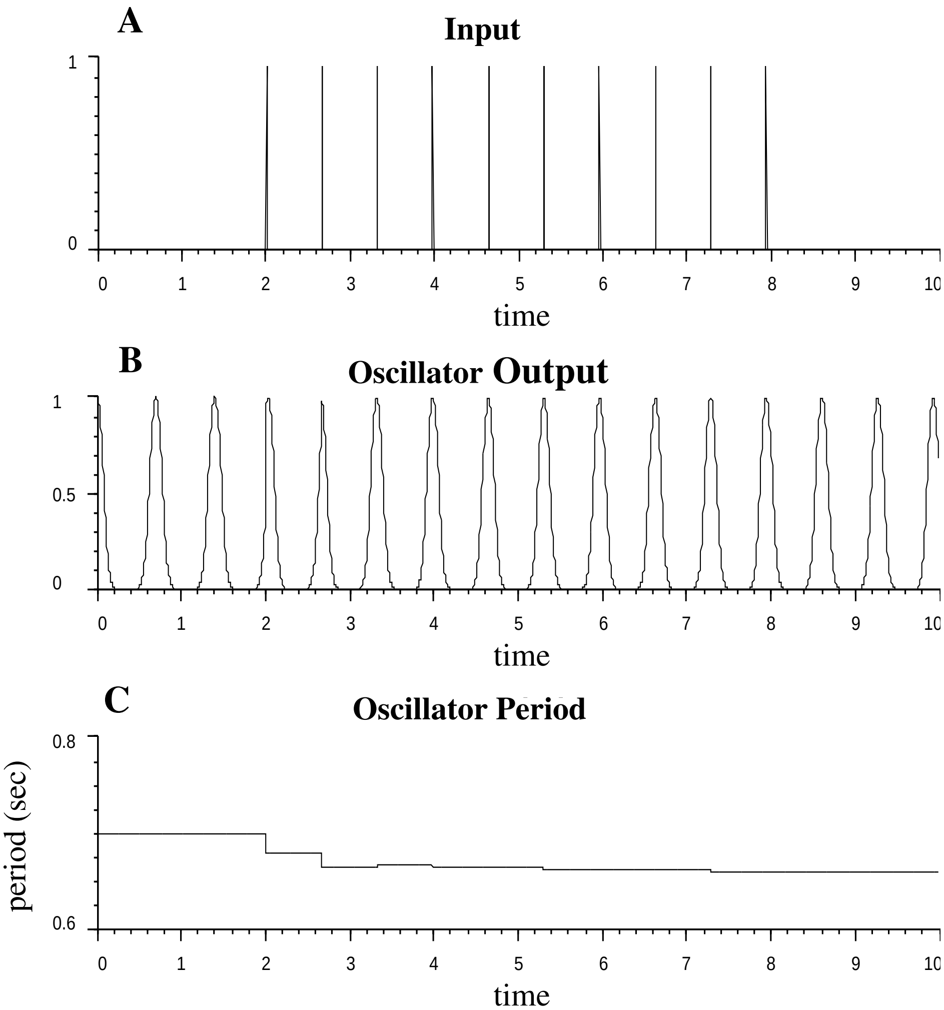

Figure 5. An oscillator responding to periodic stimulation at 660ms. Initially, the oscillator's period is 700ms. After a few stimulus cycles, the oscillator adjusts its period to 660ms. A) Periodic stimulus. B) Oscillator response. C) Oscillator period.

In summary, this single oscillatory unit synchronizes its output pulses to a periodic train of discrete event onsets. Each output pulse instantiates a temporal receptive field for the oscillatory unit – a window of time during which the unit "expects" to see a stimulus pulse. The unit responds to stimulus pulses that occur within this field by adjusting its phase and period, and ignores stimulus pulses that occur outside this field. The width of the receptive field can be adjusted by changing the unit's output gain. We have shown that the unit can entrain 1:1 to a simple periodic train of event onsets within its frequency range. A metrical structure, however, consists of levels of beats with different periods. To model the perception of metrical structure, we envision a network of units with different frequency ranges. Therefore, we must understand how the unit responds to periodic stimulation outside its frequency range. We address this issue by examining a regime diagram for the model. Next, we must understand how the unit will respond under conditions of complex rhythmic stimulation: Is the unit capable of isolating and responding to a single periodic component of a complex rhythmic stimulus?

4.2 Response of Single Units to Periodic Stimulation

If the unit is to be used within a network for self-organizing a perception of metrical structure, it will be exposed to stimulus frequencies that lie outside of its response range. Therefore, we must understand how the unit responds to periodic stimulation outside of this range. We can address this issue by examining a regime diagram for the model. This analysis will be useful in understanding how the unit will respond to any periodic driving stimulus, whether that stimulus arises from an external signal, or from the output of another oscillator in a network. Thus, the analysis will provide insight into several key aspects of our proposal for modeling the perception of metrical structure.

In order to understand the unit's response to periodic stimulation, we can formulate a circle map that summarizes the phase-locking behavior of the oscillator in response to a periodic signal. We assume that the driving signal has period q, and p is the resting period of the driven oscillator. We then use the delta rule to derive a non-linear coupling term, which gives the following circle map:

$$\phi_{i+1} = \phi_i + \frac{q}{p} - \eta_1 \operatorname{sech}^2\gamma(\cos 2\pi\phi_i - 1)\sin 2\pi\phi_i, \qquad (7)$$

where η_1 is the coupling strength for phase-tracking. ϕ_i represents the phase of the driven oscillation at which the driver fires on iteration i. This equation reveals the relationship between this circle map and the sine circle map, because when $\gamma = 0$, $\operatorname{sech}^2(\gamma\cos 2\pi\phi_i - \gamma) = 1$.

To create a regime diagram, rather than solving the model equations to analytically determine the boundaries of phase-locked states (as in Figure 3A), we repeatedly iterate this difference equation for different initial values of q/p and η_1, beginning with $\phi_0 = 0$ (i.e. we assume that the oscillators initially fire together). This allows us to calculate the number of cycles that it takes for the system to converge on stable phase-locked states, which is useful since time-to-convergence is an important factor in real-time processing.

Iteration of the equation yields the regime diagrams of Figure 6. Figures 6A-C show stable phase-locking modes for rational ratios, q/p such that $p \le 8$. Darker regions correspond to regions of faster convergence. Each individual picture corresponds to a different value of γ. Figure 6A, the regime diagram for our model with $\gamma = 0$, again shows the relationship between this circle map and the sine circle map (compare Figure 6A with Figure 3A). Diagrams B and C show entrainment zones for $\gamma = 2$ and $\gamma = 8$, respectively. As the diagrams show, the effect of increasing γ, thereby shrinking the oscillator's temporal receptive field, is to shrink the zones of 0:1 and 1:1 entrainment while widening the regions corresponding to more complex ratios. This allows the oscillator to acquire stable phase-locks in complex ratios with the stimulus more easily.

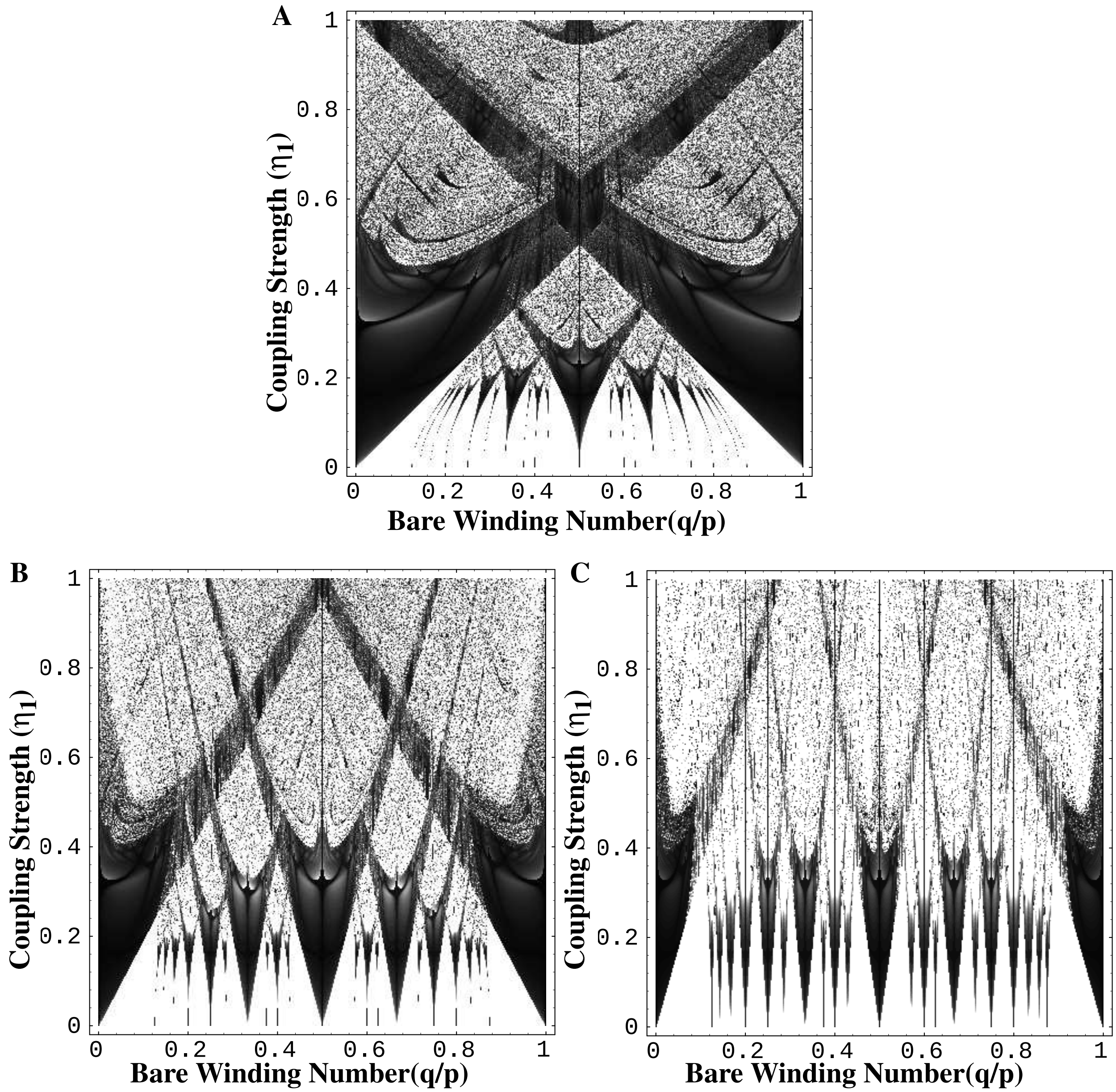

Figure 6. Regime diagrams summarizing phase-locking behavior for various values of γ. Darker regions correspond to parameter values that yield faster phase-locking. White regions are regions of quasi-periodic response. A) $\gamma = 0$. B) $\gamma = 2$. C) $\gamma = 8$.

Regime diagrams for the frequency-tracking version of the oscillator can also be developed. We add the equation:

$$\Omega_{i+1} = \Omega_i + \eta_2 \operatorname{sech}^2\gamma(\cos 2\pi\phi_i - 1)\sin 2\pi\phi_i \frac{dp}{d\Omega}, \tag{8}$$

and recalculate p_i according to Equation (5) at each iteration. Figure 7 shows the frequency-tracking resonance tongues. For easy comparison with Figure 6, we added Equation 8 to the model keeping η_2 fixed at a value of 2.5, and varied η_1 (in Equation 7) along the y-axis. Diagrams A, B, and C show resonance tongues for γ= 0, 2, and 8, respectively. The entrainment regions for the frequency-tracking oscillator are larger than the corresponding regions for phase-locking alone. Frequency-tracking causes widening of the resonance tongues. Therefore, not only does frequency-tracking act as a sort of memory, but it enhances the stability of the oscillator's response in the presence of timing deviations.

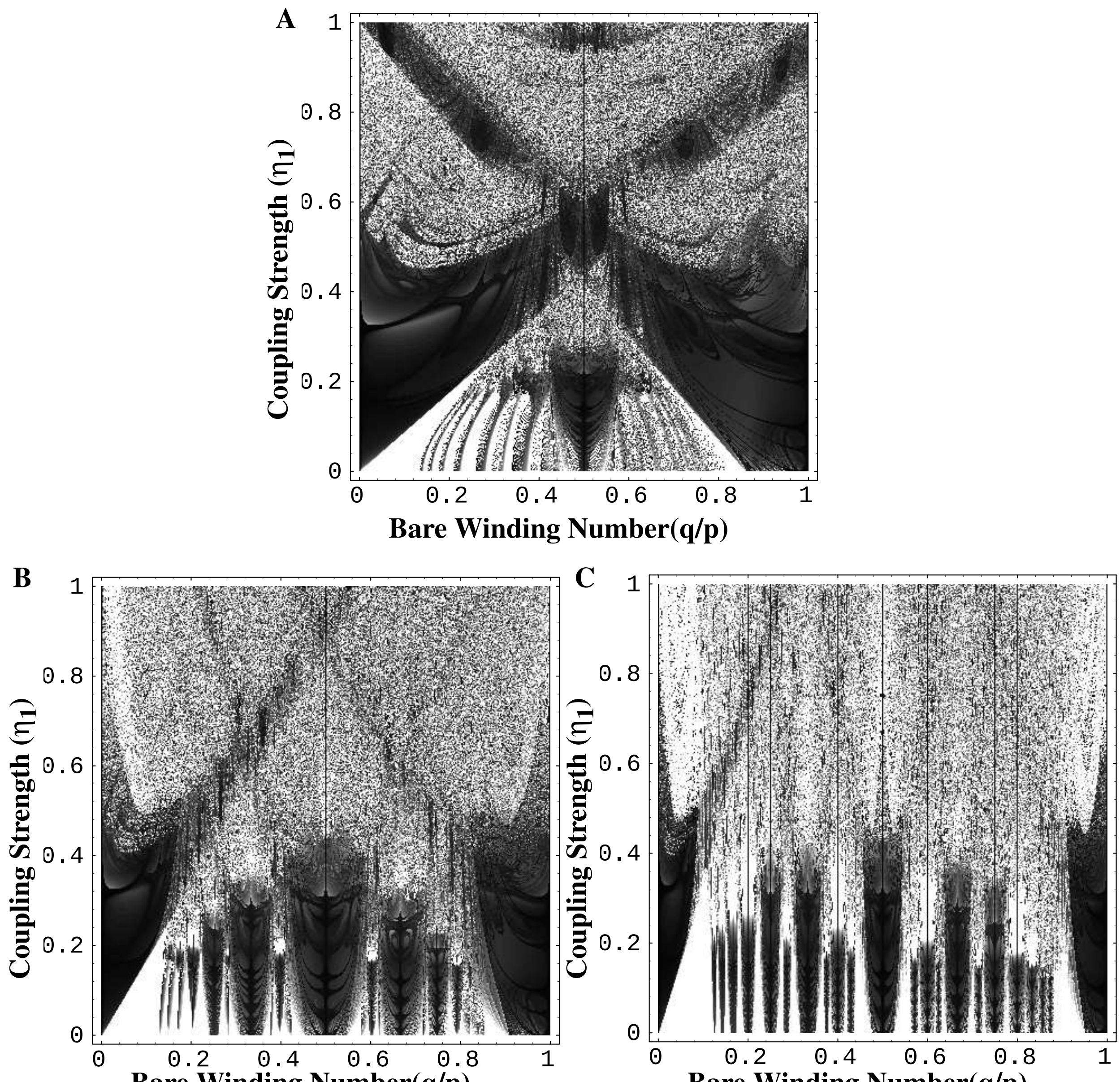

Figure 7. Regime diagrams summarizing phase-tracking behavior for various values of γ, with frequency-tracking turned on (compare with Figure 6). Frequency tracking strength is fixed at η_2 = 2.5. A) γ = 0. B) γ = 2. C) γ = 8.

The regime diagrams above detail the behavior of our oscillatory unit in response to periodic stimulation. In so doing, the analysis summarizes key implications of our single unit model for resonance-based theories of metrical structure. First, regime diagrams will capture the content of resonance theories regarding the well-formedness of metrical structures. Rather than formulate a set of rules summarizing allowable beat period ratios, we may specify the connectivity of a network and a set of parameters to each unit, and let the system dynamics enumerate allowable beat-period ratios. In principle any ratio of relatively prime integers may describe the relationship between two levels of beats. Thus, this approach affords a wider scope than conventional theories of metrical structure, incorporating not only integer ratios (such as 2:1 and 3:1), but also polyrhythmic ratios. Second, the width of each resonance tongue describes the stability of each beat period ratio in the face of timing variability. Therefore, the regime diagram provides an enumeration of allowable beat period ratios by sensitivity to timing variability. One implication of this observation is that in practice the number of mode-locks systems can achieve will be limited, preventing extravagant theoretical claims regarding the perception of polyrhythmic structure. This brings us to the issue of timing variability. The style of the above analysis raises the possibility that this entrainment theory may allow the formulation of theories of meter perception that adequately handle timing variability in musical performance. It also suggests that the units may be able to effectively isolate individual periodic components of complex rhythmic patterns. These and other empirical issues are examined in the following section.

4.3 Rhythmic Stimulation

As we have just seen, the response of a single unit to a simple periodic stimulus is quite complex. The response to a rhythmic performance will be even more complex. Therefore, rather than jumping to the complexity of a full-blown network of oscillators, we examine the behavior of individual units in more detail. In this section, we expose an unconnected system of oscillators to a performed musical rhythm, in order to address the response of individual units to complex temporal patterns. We demonstrate how individual units can isolate and track component periodicities in a complex rhythm, and also show how units realize natural metric preferences. We also address the problem of timing variability. Finally, the output of a system of multiple units may be interpreted as a perception of metrical structure. In fact, we will observe a system of unconnected units behaving quite reasonably in response to complex rhythmic stimulation.

We exposed a system of oscillators to one of the musical performances collected in an earlier study of musical improvisation (Large, Palmer, & Pollack, in press). This performance was collected on a computer-monitored Yamaha Disklavier acoustic upright piano. Optical sensors and solenoids in the piano allowed precise recording and playback without affecting the touch or sound of the acoustic instrument. The pitch, timing, and hammer velocity values (correlated with intensity) for each note event were recorded. The pianist performed and recorded an original melody, as presented in musical notation, five times. With the musical notation remaining in place, the pianist was then asked to play five "simple" improvisations. All performances were of a single-line melody only; the subject was instructed not to play harmonic accompaniment. The recording yielded a list of MIDI events, from which we extracted note-on times to use as input to our model. Figure 8A gives the performance in piano-roll notation, and 8B shows the input, $s(t)$, to the system of oscillators. Our metrical interpretation of this performance, in the form of a transcription, is given in Figure 9.

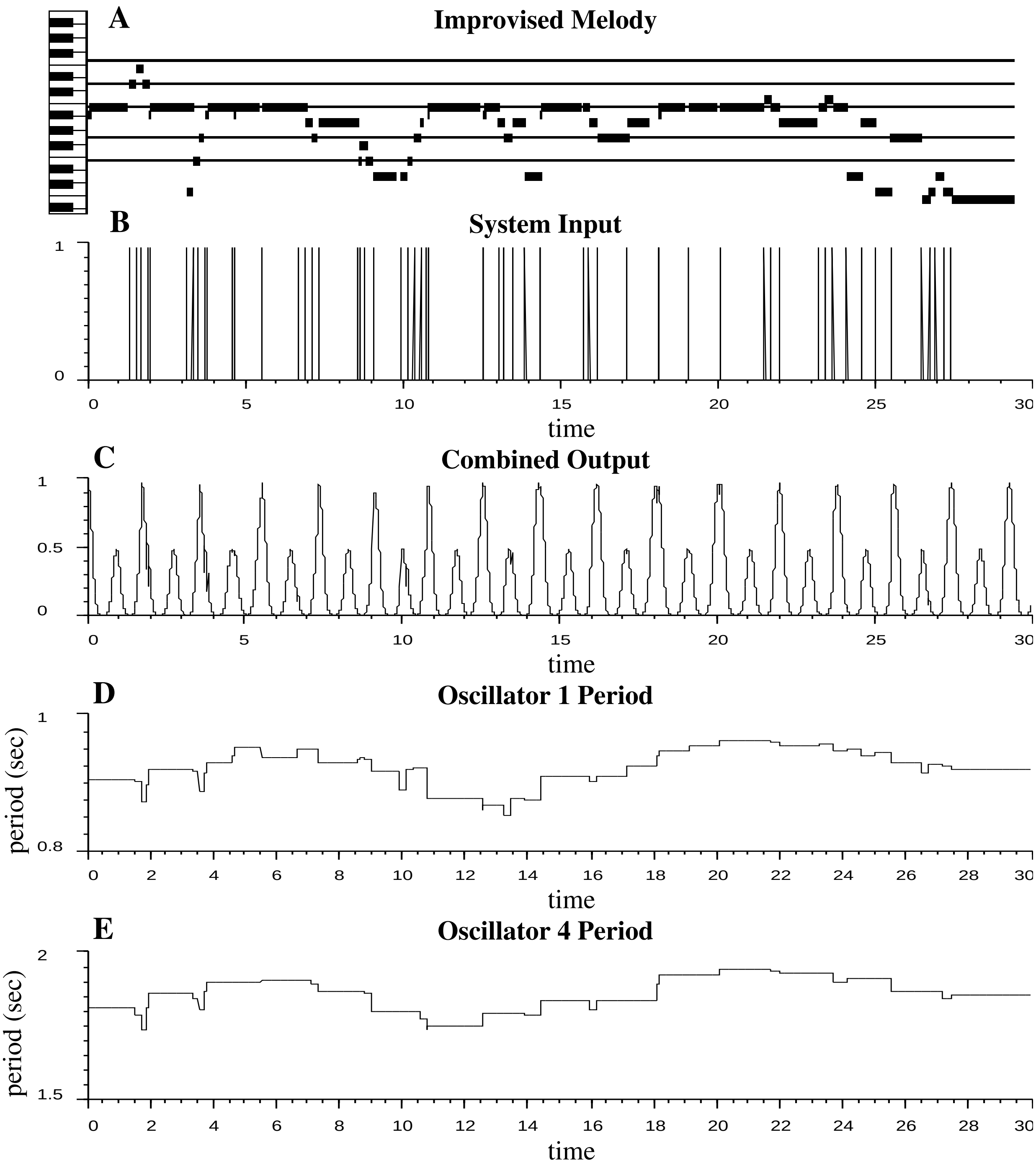

Figure 8. Two oscillators responding to an improvised melody. A) Piano roll notation for the melody. B) Input to the oscillators. C) Weighted sum of the oscillators' output. D) Oscillator 1's period tracking curve. E) Oscillator 4's period tracking curve.

For this study, we composed a system of oscillators with different frequency ranges. Such systems are useful for self-organizing metrical responses to rhythmic stimuli (Large & Kolen, 1993). We set each oscillator's period range according to the rule

$$p_{max} = \frac{4}{3}p_{min}.$$

We then spaced oscillators such that the relationship between the resting period of one oscillator and the next was given by

$$p_{i+1} = 2^{\sqrt[3]{p_i}}.$$

This relationship, in conjunction with the frequency range of each oscillator, provides for slight overlap in resonant frequency ranges between oscillators. We composed a system of two "octaves" of oscillators, six in all. The minimum period of the entire system was 600ms, and the maximum was 2560ms. For each oscillator we set $\eta_1 = 0.159$, $\eta_2 = 3.1416$.

We exposed the entire bank of oscillators to the performance, and each oscillator responded independently to the discrete event onsets. We assumed that the initial onset of the performance phase-reset all oscillators. Two of the six oscillators (oscillators 1 and 4) acquired stable mode-locks for this performance, the remaining oscillators never stabilized. Figure 8C shows the output of these two oscillators, combined according to the rule

$$\frac{o_1(t) + o_4(t)}{2}.$$

Figures 8D and E show the period of these two oscillators, respectively, as they track the expressive timing of the performance.

This single example provides much insight into the behavior of individual oscillators in response to musical rhythms. It also provides insight into the issues involved in using such units to build a network for meter perception. First, each of these oscillators is isolating a periodic component of the complex rhythm without any phenomenal accent information. The global response, as can be seen from the combined output of the two units in Figure 8C, shows that a stable metrical interpretation of the input rhythm emerges rather quickly, with strong and weak beats clearly observable. According to our metrical interpretation of this performance (see Figure 9), these two oscillators are correctly responding to the metrical structure at the quarter note and half note levels. Also, as Figures 8D and 8E show, the oscillators are tracking the performance over rather large changes in tempo.

Figure 9. Transcription of the improvised melody from Figure 8. Grace notes are not transcribed.

To understand how this happens, let us first examine how the units respond to timing variation in performance. Note that by the middle section of the performance ($14 < t < 17$), Unit 1 has locked into a periodic component of the rhythm, and is maintaining a relatively stable tempo. The next few onsets signal an audible *ritardando* in the performance. At each of these onsets the unit fires a bit early. The effect of the input in each case is to cause a slight positive phase shift, and a slight lengthening of the unit's period. The units can be seen to follow the systematic timing deviations that occur in this musical performance.

Next we examine how individual units isolate periodic components in complex rhythmic patterns. We also see how this process instantiates metrical preferences for individual units, which manifest themselves as the alignment between adjacent levels of beats. We are interested in understanding which onset of a group is interpreted by each unit to mark the beat at its beat period level. We can gain some insight by examining Figure 10, which shows a close-up for Units 1 and 4 as they respond individually in the early seconds of the performance. The panel for each unit

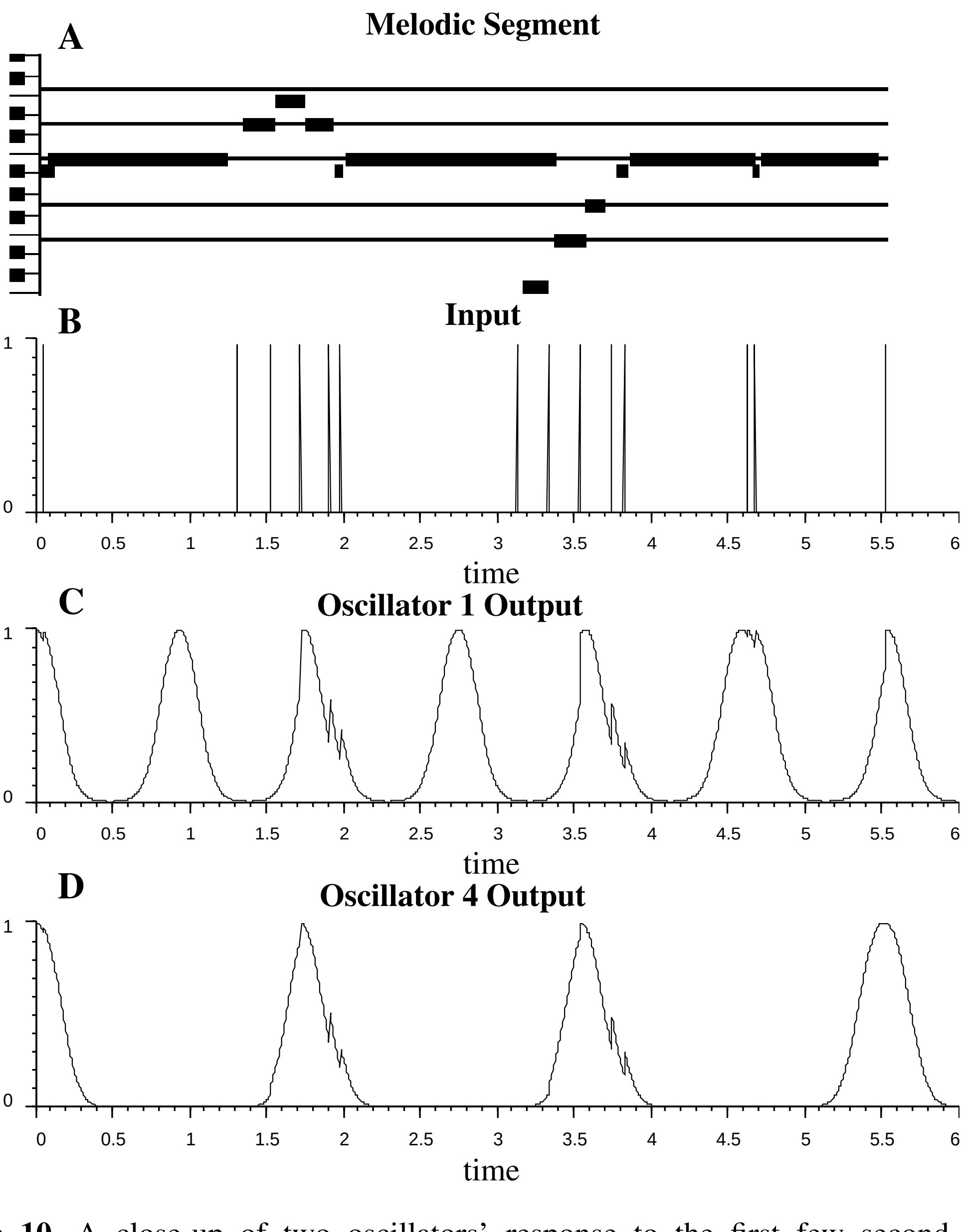

Figure 10. A close-up of two oscillators' response to the first few second of an improvised performance. A) Piano roll notation of the melody. B) Input to the oscillators. C) Oscillator 1's output. D) Oscillator 4's output.

shows the individual output pulses. Consider responses to the group of onsets between $t = 1$ and $t = 2$. The second onset (at $t \cong 1.5$) is ignored by Unit 1, because it does not fall within the unit's expectancy region. The same onset, however, causes a slight negative phase and period adjustment in Unit 4 (see also Figure 8E at $t \cong 1.5$). The third onset (at $t \cong 1.7$) causes Unit 1 to make a rather large adjustment and Unit 4 to make another small adjustment, eliciting coincident output pulses from both units. We interpret this onset as marking the second half of the 4/4 measure (see Figure 9), therefore both oscillators are responding correctly. However, the two onsets immediately following do cause some adjustment in both units. The response is still in flux.

Interestingly, both units respond almost identically to the next group of onsets, between times $t = 3$ and $t = 4$. This time, however, they respond maximally to the "wrong" onset (the third of the group). The fourth and fifth onsets actually mark the onset of the second measure according to our interpretation. These onsets do have some effect on both oscillators. However, it is the next two onsets (not counting the grace note accompanying the first onset) that clearly establish the beat at the quarter-note level. At this point, both oscillators lock into the rhythm, responding correctly to the remainder of the performance with little difficulty.

It is difficult to extract precise rules to explain individual unit preferences. However there are some general observations that can be made. First, each unit's choice is brought about by a subtle interplay between the unit's point of maximum expectancy (in the current cycle), the spacing of event onsets around this point, the width of the unit's expectancy region (determined by γ, the output gain), and the absolute amount of adjustment made to phase and period in response to each onset (determined by the coupling strengths, η_1 and η_2).

Assuming fairly high coupling strengths, as we used in the above example, consider a group of event onsets that surrounds a unit's point of maximum expectancy, such that the unit "expects" that an event somewhere between the first and last event of the group will mark the beat. If the spacing of the surrounding onsets is greater than the width of the unit's expectancy region, then the unit will simply ignore the surrounding events. However, as we squeeze the onsets closer together, encroaching on the unit's expectancy region, the unit will begin to respond to these onsets. If the onsets are very close together, the unit will continue to respond until reaching the end of the group. Thus, each unit tends to favor the end of a group of events, where "end of the group" is defined in relation to the width of the unit's expectancy region. Additionally, places where a number of events occur in rapid succession (grace notes above, or chords for example) act as points of greater "gravity" for all units, because they have additional impact via the delta rules.

This analysis has illustrated several important aspects of our proposal. First, the individual oscillatory units we propose can successfully pick out and lock onto periodic components of complex rhythmic patterns without making any assumptions about the structure of phenomenal accent patterns in the stimulus. This distinguishes our model from attempts to model musical beat using traditional phase-locked loop circuits. Second, the way in which units with different native periods accomplish this task effectively implements metrical preferences. Metrical preferences manifest themselves as the alignment, or relative phase, between adjacent levels of beats. This system of unconnected units correctly interpreted two levels of metrical structure for a complex performed rhythm, without any attempt to implement metrical well-formedness constraints. We assume that, in general, it will be necessary to implement well-formedness constraints using interactions among oscillators. However, results such as this suggest that the structure of rhythmic patterns, even for very complex performances, may contain more information than had previously been thought. Finally, this example demonstrates the proposed oscillatory units handling systematic timing deviations throughout a rhythmically free improvised performance. These results suggest that theories of metrical structure based on such models of entrainment may provide more complete theoretical accounts of metrical structure perception than have previously been offered.

4.4 Implementation Issues

We have described our model as a single abstract processing unit, in order to focus attention on the adequacy of the proposal for modeling the human response to musical rhythm. However, the issues surrounding implementation deserve some attention. Page (1993), in his proposal of an oscillatory connectionist network for tracking musical beats, recruits a relatively large network of traditional connectionist units into a neural implementation of a standard phase-locked loop. The heart of the network is a gated pacemaker circuit (Carpenter & Grossberg, 1983). Page then implements a Type-II phase detector and a low-pass filter using networks of connectionist processing units, to provide an error signal that controls adjustments of phase and period in the gated pacemaker. It seems likely that the oscillatory unit that we have proposed above could yield to a similar implementation strategy using a large network of simpler connectionist units with appropriate dynamics.

It is not clear, however, that a complex implementation strategy such as Page's (1993) is necessary. Consider that synchronization behavior has been proposed by other connectionist modelers to explain how the brain solves problems of binding. Often, a network of two units produces the oscillatory behavior of interest, and synchronization arises given simple couplings (e.g. Wang, 1993). Such proposals have received considerable physiological support, especially since the discovery of oscillations and synchronization in the cat visual cortex (Eckhorn et al. 1989; Gray et al. 1989). McAuley (1993; 1994), who has proposed entrainment models for the perception of rhythm, has suggested that behaviors relevant to this task, including frequency-tracking, may be found at the single neuron level. While we welcome this possibility, we do have some reservations. We have proposed a functional approach, not an implementation-level strategy. We have proposed oscillatory units to describe how the brain may execute the relatively high-level cognitive function of meter perception. Therefore we assume that our abstract, functional units represent higher levels of abstraction than individual neurons. Further, consider that the behaviors we have identified as necessary for rhythm perception may be implemented using large networks of simpler oscillatory elements. For these reasons, we suggest that the behavior of these abstract units may be plausibly regarded as the emergent behavior of a wide range of possible brain structures from simple neuronal substructures to large networks of oscillatory neurons.

5 General Discussion

The primary goal of our proposal has been to understand the implications of entrainment for theories of metrical structure and musical expectancy. To this end, we have proposed an abstract oscillatory unit that may be composed into networks for modeling the perception of metrical structure. The unit may synchronize its periodic output pulses with an incoming rhythmic pattern. The unit responds to event onsets that occur within its temporal receptive field by adjusting its phase and period, and ignores stimulus pulses that occur outside this field. The width of the receptive field can be adjusted using a parameter called output gain. This enables the unit to isolate single periodic components of complex rhythms.

Analysis of the behavior of a single unit in response to periodic stimulation reveals complex dynamics. In principle, the unit may mode-lock to a periodic stimulus in any one of an infinite number of rational ratios. Tuning the unit's temporal receptive field has the effect of adjusting the relative stability of mode-locking regions. Large temporal receptive fields result in a preference for simple ratios while finely tuned regions allow more complex ratios. These properties have important implications for entrainment theories of metrical structure. Regime diagrams will summarize the content of such theories regarding the well-formedness of metrical structures. The Farey tree enumerates the possible relationships between two levels of beats, while the corresponding regime diagram describes the stability of resulting metrical relationships.

The phase and frequency-locking behavior of individual units implicitly describes a set of metrical preferences – a set of preferred phase relationships between two levels of beats in a metrical structure, relative to the structure of the incoming rhythm. Such preferences may be best understood in terms of the characteristic response of an individual unit to a complex rhythmic pattern. Ultimately, in a network the influence of other units may mediate individual unit preferences, and subsets of units will respond to an afferent rhythmic pattern as a whole. We have also demonstrated that entrainment provides a robust approach to the perception of meter in musical performance. This allows us to account for the perception of meter in the face of timing variability. Thus, the model embodies a dynamic solution to the "quantization problem" (Desain & Honing, 1991). Finally, we demonstrated a system of oscillatory units correctly tracking two metrical levels in an improvised melodic performance. The rhythm of the improvised melody was complex, yet the simple system behaved quite reasonably. This indicates that our proposal may provide the basis for more comprehensive, robust, and parsimonious theories of the perception of metrical structure.

5.1 Future Work

The primary goal of this article has been to understand the implications of entrainment for theories of musical meter. We have stopped short of proposing a theory of musical meter. Before attempting to construct such a theory, we would need to resolve at least two issues. The first is the issue of phenomenal accent. We have seen how a tightly grouped set of onsets (a chord, or a melody note with an accompanying grace note) may impact the preference of individual units. The model, as it now stands, does not specify how other types of phenomenal accent (e.g. intensity, or large pitch leaps) may affect individual unit preferences regarding which events are interpreted as marking the beat. One possibility calls for a real-valued representation of event onsets, $s(t)$, to carry accent information. Because of the formulation of the delta rules, events with greater accent would cause greater phase and period adjustments. Although this technique appears promising, it does not provide a theory of accent such as N. Todd's (1994) "rhythmogram" model.

The second issue that stands between our entrainment mechanism and a theory of musical meter is the issue of network construction. In this article, we have concentrated on the behavior of individual units independently responding to afferent rhythms. Yet one would expect individual units within a network to interact, responding to the outputs of other units in the network. The important question is: Could a stable response emerge from such a network subjected to a musical event sequence? Our analysis of the single unit case suggests that subsets of units in a loosely coupled network could self-organize a coherent response to an afferent rhythm. In addition, the interaction would instantiate metrical well-formedness constraints. The major challenge facing this approach is to determine the nature of the interaction. We leave the issues of network construction and phenomenal accent unresolved, and we regard these as important areas for future exploration. We believe that an understanding of mechanisms of entrainment will result in theories of the meter perception that are wider in scope, and more parsimonious than those that have previously been offered. Entrainment and self-organization provide expressive power and useful physicalist constraints unavailable within more general purpose theoretical frameworks. At the same time, these principles offer greater robustness to deal with the problems associated with the perception of actual musical performances.

5.2 Implications for Connectionist Approaches to Expectancy

At the outset of this article, we pointed to two limitations of recurrent neural network approaches to musical expectancy. The first problem was the representation of temporal context. Mozer (1993) and deVries and Principe (1992) have suggested that an exponential recency gradient, inherent in most network architectures, limits the ability of recurrent networks to represent temporal context. The most recent items presented to the network carry more weight than previous inputs, inhibiting the network's ability to capture global structure. Other connectionist approaches have addressed this issue by using a system of short term memory delays to explicitly capture temporal context (Lang, Waibel, and Hinton, 1990; Unnikrishnan, Hopfield, and Tank, 1991; Bodenhausen and Waibel, 1991; deVries & Principe, 1992; for a review see Mozer, 1993). Delays may be hard-wired or learned during batch training, but during processing they remain fixed. The problem with a fixed memory delay solution for music processing should now be apparent. Music lacks fixed temporal structure. Musical signals display complex forms of temporal organization including expressive timing deviations and periodic structure on multiple time scales.

The second problem with recurrent neural network approaches to music expectancy was their inability to generalize to novel rates of presentation. Cottrell, Nguyen, and Tsung (1993) have attempted to solve this problem by implementing a strategy for rate invariant sequence recognition. They first trained a recurrent network to predict a target input signal presented at some "normal" rate. A typical recurrent network would be able to track the target signal at this rate, but would lose the signal at other rates. Cottrell et al augmented their network to control its own processing rate by adapting time constants and processing delays. Using prediction error, the recurrent network adapted its processing rate to match the rate of the current signal, much like a phase-locked loop varies its internal frequency to match the phase of an incoming signal. This approach appears to yield plausible explanations for some aspects of perception, including the perception of music. It is not a general solution to the problem of rate invariance in music, however, because it applies only to learned sequences.

Each approach described above addresses individual aspects of temporal sequence recognition and prediction, and leaves others unattended. Resonance-based approaches provide an alternative that combines the strengths of above mentioned proposals, while solving some remaining problems. The combined output of a system of oscillators, as shown in Figure 8C, can provide complex temporal control (i.e. pulses of attention) to affect the further processing of a musical sequence. A resonance-based system could adapt its processing rate according to the structure of the signal, without memorizing the signal in advance. Memory delays implemented as resonance-based components, could likewise adapt to the rate of the incoming signal. In addition, the structure of memory itself could adapt to reflect the temporal organization (e.g. the metrical structure) of the incoming signal. Thus, we feel that resonance mechanisms offer a particularly novel route toward understanding musical expectancy.

6 Conclusions

Lashley (1951) identified the problem of serial order, "... the logical and orderly arrangement of thought and action," as a central problem for psychologists, neurobiologists, and all those who ultimately wish to describe the phenomena of mind in terms of the mathematical and physical sciences. Lashley realized that the problem was not merely one of sequence processing. The temporal structure of human perception and action implies that the temporal structure of neural computation is extraordinarily complex (Lashley, 1951). In this regard, the study of music is invaluable to the understanding of neural computation. Music, unlike natural language, forces us to deal with all aspects of time: time is so fundamental to music that it cannot be conveniently and convincingly abstracted away. It may even be that composers and performers shape the temporal structure of music to reflect and to explore natural modes of temporal organization in the human nervous system.

For inherently temporal tasks, such as perception and motor coordination, we agree that resonance provides a more useful metaphor than general-purpose computation (Gibson, 1966; 1979; Treffner & Turvey, 1993). According to this view, the brain may be treated as a special purpose device, capable of temporarily adapting its function to specific perception-action situations (Kelso & deGuzman, 1988). In perception, the nervous system may adapt endogenous modes of temporal organization to external rhythmic patterns, controlling attention and memory (Jones, 1976). Other connectionists have noted the fundamental consonance of such dynamical systems approaches with modern connectionist cognitive modeling (e.g. van Gelder & Port, forthcoming). Ours is an attempt to bring the two closer together to overcome the limitations of current connectionist models. We have found music perception to be a fertile testing ground for this approach. Our current proposal attempts to explain the mechanisms underlying temporal adaptation in the human response to musical rhythms. We believe that this approach will lead to more robust and parsimonious theories of musical meter and musical expectancy.

References

Apel, W. (1972) *Harvard dictionary of music* (2nd ed.). Cambridge, MA: Belknap Press of Harvard University Press.

Allen, P. E. & Dannenberg, R. B. (1989) *Tracking musical beats in real time*. In Proceedings of the 1990 International Computer Music Conference. Computer Music Association.

Beek, P. J., Peper, C. E. & van Wieringen, P. C. W. (1992) Frequency locking, frequency modulation, and bifurcations in dynamic movement systems. In G.E. Stelmach and J. Requin (Eds.) *Tutorials in motor behavior II*. Elsevier Science Publishers B. V.

Bharucha, J. J. & Todd, P. M. (1989) Modeling the perception of tonal structure with neural nets. *Computer Music Journal*, 13, 44-53.

Bodenhausen, U. & Waibel, A. (1991) *The Tempo 2 algorithm: Adjusting time delays by supervised learning*. In R. P. Lippmann, J. Moody, & D. S. Touretsky (Eds.) Advances in Neural Information Processing Systems 3. San Mateo, CA: Morgan Kaufman.

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

Carpenter, G. A. & Grossberg, S. (1983) A neural theory of circadian rhythms: The gated pacemaker. *Biological Cybernetics*, 48, 35-59.

Clarke, E. (1985) Structure and expression in rhythmic performance. In P. Howell, I. Cross, and R. West (Eds.) *Musical Structure and Cognition*. London: Academic Press.

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

Cottrell, G. W., Nguyen, M. & Tsung, F. (1993) *Tau Net: The way to do is to be*. In Proceedings of the Fifteenth Annual Conference of the Cognitive Science Society. Hillsdale, NJ:Erlbaum Press.

Dannenberg, R. B. (1984) *An on-line algorithm for real-time accompaniment*. In Proceedings of the 1984 International Computer Music Conference. Computer Music Association.

Dannenberg, R. B., Mont-Reynaud, B. (1987) *Following an improvisation in real time*. In Proceedings of the 1987 International Computer Music Conference. Computer Music Association.

Desain, P. & Honing, H. (1991) The quantization of musical time: A connectionist approach. In P. M. Todd and D. G. Loy (Eds.) *Music and Connectionism*, 150-160. Cambridge: MIT Press.

Desain, P. (1992) A (de)composable theory of rhythm perception. *Music Perception*, 9, 101-116.

de Vries, B. & Principe, J. C. (1992) The gamma model – A new neural net model for temporal processing. *Neural Networks*, 5, 565-576.

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

Dirst, M. & Weigend, A. S. (1993) Baroque forecasting: On completing J. S. Bach's last fugue. In A. Weigend N. Gershenfeld (Eds.) *Predicting the Future and Understanding the Past*, 151-172. Reading, MA: Addison-Wesley.

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

Drake, C. & Palmer, C. (1993) Accent structures in music performance. *Music Perception*, 10, 343-378.

Drake, C. & Botte, M. (1993) Tempo sensitivity in auditory sequences: Evidence for a multiple-look model. *Perception and Psychophysics*, 54, 277-286.

Eckhorn, R., Reitboeck, H. J., Arndt, M. and Dicke, P. (1989) *Feature linking via stimulus evoked oscillations: Experimental results from cat visual cortex and functional implications from a network model*. In Proceedings of the International Joint Conference on Neural Networks. Washington.

Elman, J. (1990) Finding structure in time. *Cognitive Science*, 14, 179-211.

Essens, P. J. & Povel, D. (1985) Metrical and nonmetrical representation of temporal patterns. *Perception and Psychophysics*, 37, 1-7.

Getty, D. J. (1975) Discrimination of short temporal intervals: A comparison of two models. *Perception & Psychophysics*, 18, 1-8.

Gibson, J. J. (1966) *The senses considered as perceptual systems*. Boston: Houghton Mifflin.

Gibson, J. J. (1979) *The ecological approach to visual perception*. Boston: Houghton Mifflin.

Glass, L. & Mackey, M. C. (1988) *From clocks to chaos: The rhythms of life*. Princeton, NJ: Princeton University Press.

Gray, C. M., Konig, P., Engel, A. K. & Singer, W. (1989) Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. *Nature*, 338, 334-337.

Halpern, A. R. & Darwin, C. (1982) Duration discrimination in a series of rhythmic events. *Perception & Psychophysics*, 31, 86-89.

Handel, S. (1989) *Listening*. Cambridge, MA: MIT Press.

Jones, M. R. (1976) Time, our lost dimension: Toward a new theory of perception, attention, and memory. *Psychological Review*, 83, 323-335.

Jones, M. R. (1981) Music as a stimulus for psychological motion: Part I: Some determinants of expectancies. *Psychomusicology*, 1, 34-51.

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

Jones, M. R. (1987b) Perspective on musical time. In A. Gabrielsson (Ed.) *Action and Perception in Rhythm and Music*, 153-175. Stockholm: The Royal Swedish Academy of Music.

Jones, M. R., Boltz, M. & Kidd, G. (1982) Controlled attending as a function of melodic and temporal context. *Journal of Experimental Psychology: Human Perception & Performance*, 7, 211-218.

Jones, M. R., Kidd, G., Wetzel, R. (1981) Evidence for rhythmic attention. *Journal of Experimental Psychology: Human Perception & Performance*, 7, 1059-1073.

Jones, M. R. & Boltz, M. (1989) Dynamic Attending and Responses to Time. *Psychological Review*, 96, 459-491.

Jordan, M. (1986). Serial order. Technical Report 8604, Institute for Cognitive Science, University of California at San Diego, La Jolla, CA.

Kelso, J. A. S. & deGuzman, G. C. (1988) Order in time: How the cooperation between the hands informs the design of the brain. In H. Haken (Ed.) *Natural and Synergetic Computers*. Berlin: Springer-Verlag.

Lang, K., Waibel, A. H. & Hinton, G. E. (1990) A time-delay neural network architecture for isolated word recognition. *Neural Networks*, 3, 23-44.

Large, E. W. & Kolen, J. F. (1993) *A dynamical model of the perception of metrical structure*. Presented at Society for Music Perception and Cognition. Philadelphia, June.

Large, E. W., Palmer, C. & Pollack, J. B. (in press) Reduced memory representation for music. *Cognitive Science*.

Larson, S. (1993) *Modeling melodic expectation: Using three "musical forces" to predict melodic continuations*. In Proceedings of the Fifteenth Annual Conference of the Cognitive Science Society. Hillsdale, NJ:Erlbaum Press.

Lashley, K. (1951) The problem of serial order in behavior. In Jeffress (Ed.) *Cerebral mechanisms in behavior*. NY: Wiley.

Lee, C. S. (1985) The rhythmic interpretation of simple rhythmic sequences. In P. Howell, I. Cross, and R. West (Eds.) *Musical Structure and Cognition*. London: Academic Press.

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

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

Longuet-Higgins, H. C. (1987) *Mental Processes*. Cambridge: MIT Press.

Marr, D. (1982) *Vision*. New York: W. H. Freeman.

McAuley, J. D. (1993) Learning to perceive and produce rhythmic patterns in an artificial neural network. Technical Report, Department of Computer Science, Indiana University.

McAuley, J. D. (1994) *Finding metrical structure in time*. In M. C. Mozer, P. Smolensky, D. S. Touretsky, J. L. Elman & A. S. Weigend (Eds.) Proceedings of the 1993 Connectionist Models Summer School. Hillsdale, NJ: Erlbaum Associates.

McGraw, G., Montante, R. & Chalmers, D. (1991) *Rap-master network: Exploring temporal pattern recognition with recurrent networks*. Technical Report No. 336. Computer Science Department, Indiana University.

Meyer, L. (1956) *Emotion and meaning in music*. Chicago: University of Chicago Press.

Mozer, M. C. (1991) Connectionist music composition based on melodic, stylistic and psychophysical constraints. In P. M. Todd and D. G. Loy (Eds.) *Music and Connectionism*, 195-211. Cambridge: MIT Press.

Mozer, M. C. (1993) Neural net architectures for temporal sequence processing. In A. Weigend N. Gershenfeld (Eds.) *Predicting the Future and Understanding the Past*, 243-264. Reading, MA: Addison-Wesley.

Narmour, E.(1990) *The analysis and cognition of basic melodic structures: The implication-realization model*. Chicago: University of Chicago Press.

Page, M. P. A. (1993) Modelling aspects of music perception using self-organizing neural networks. Unpublished doctoral dissertation, University of Cardiff.

Palmer, C. (1989). Mapping musical thought to musical performance. *Journal of Experimental Psychology: Human Perception & Performance*, *15*, 331-346.

Palmer, C. & Krumhansl, C.L. (1990) Mental representations of musical meter. *Journal of Experimental Psychology: Human Perception & Performance*, 16, 728-741.

Port, R. F. (1990) Representation and recognition of temporal patterns. *Connection Science*, 2, 151-176.

Povel, D. & Essens, P. J. (1985) Perception of temporal patterns. *Music Perception*, 2, 411-440.

Scarborough, D. L., Miller, P. & Jones, J. A. (1992) On the perception of meter. In M. Balaban, K. Ebcioglu, & O. Laske (Eds.) *Understanding Music with AI: Perspectives in Music Cognition*, 427-447. Cambridge: MIT Press.

Schmidt, R. C., Beek, P. J., Treffner, P. J. & Turvey, M. T. (1991) Dynamical substructure of coordinated rhythmic movements. *Journal of Experimental Psychology: Human Perception & Performance*, 17, 635 - 651.

Schroeder, M. (1991) *Fractals, Chaos, Power Laws*. New York: W. H. Freeman and Company.

Schulze, H. H. (1989) The perception of temporal deviations in isochronic patterns. *Perception & Psychophysics*, 45, 291-296.

Shaffer, L. H. (1981) Performances of Chopin, Bach, and Bartok: Studies in motor programming. *Cognitive Psychology*, 13, 326-376.

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

Simon, H. A. & Sumner, R. K. (1968) Pattern in music. In B. Kleinmuntz (Ed.) *Formal Representation of Human Thought*. New York: Wiley.

Sloboda, J.A. (1983) The communication of musical meter in piano performance. *Quarterly Journal of Experimental Psychology*, 35, 377-396.

Todd, N. P. M. (1985) A model of expressive timing in tonal music. *Music Perception*, 3, 33-59.

Todd, N. P. M. (1994) The auditory primal sketch: A multi-scale model of rhythmic grouping. *Journal of New Music Research*, *23*, 25-69.

Todd, P. M. (1991) A connectionist approach to algorithmic composition. In P. M. Todd and D. G. Loy (Eds.) *Music and Connectionism*, 173-194. Cambridge: MIT Press.

Torras, C. (1985) *Temporal pattern learning in neural models*. Berlin: Springer-Verlag.

Treffner, P. J. & Turvey, M. T. (1993) Resonance constraints on rhythmic movement. *Journal of Experimental Psychology: Human Perception & Performance*, 19, 1221-1237.

Unnikrishnan, K. P., Hopfield, J. J. & Tank, D. W. (1991) Connected-digit speaker-dependent speech recognition using a neural network with time-delayed connections. *IEEE Transactions on Signal Processing*, 39, 698-713.

von der Malsberg, C. & Schneider, W. (1986) A neural cocktail-party processor. *Biological Cybernetics*, 54, 29-40.

van Gelder, T. & Port, R. (forthcoming) Introduction. In T. van Gelder & R. Port (Eds.) *Mind as Motion*. Cambridge: MIT Press.

Wang, D. (1993) Pattern recognition: Neural networks in perspective. *IEEE Expert*, 8, 52-60.

Winfree, A. T. (1980) *The geometry of biological time*. New York: Springer-Verlag.

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

Yee, W., Holleran, S. & Jones, M. R. (1994) Sensitivity to event timing in regular and irregular sequences: Influences of musical skill. *Journal of Experimental Psychology: Human Perception & Performance*, *55*, 1-11.

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

Acknowledgments

This research was partially supported by the Ohio State University Presidential Fellowship to the first author, by NIMH grant 1R29-MH45764 to Caroline Palmer in support of the first author, and by ONR Grant N00014-92-J115 to Jordan B. Pollack in support of the second author. The authors would like to thank Mari Jones for many helpful discussions, and for comments on earlier drafts of this paper. We would also like to thank Caroline Palmer, Jordan Pollack, Timothy Walker, Kevin Lenzo, Flip Phillips, Shannon Campbell, and Gregory Saunders for helpful comments and discussions. Thanks also to two anonymous reviewers for providing thoughtful and thorough comments on an earlier draft of the paper. Reprint requests should be addressed to Edward W. Large, Department of Computer and Information Science, 2036 Neil Avenue, The Ohio State University, Columbus, OH 43210.

Reference Notes

[1] "What Next" is also the name of a computer program that models musical expectancy (Larson, 1993).

[2] The term "subjective rhythm" is a misnomer according to modern terminology. According to conventional modern usage, this phenomenon would be called "subjective meter".