Periodicity, Pattern Formation, and Metric Structure

Edward W. Large

Center for Complex Systems and Brain Sciences

Florida Atlantic University

Running Head: Periodicity and Pattern

Address correspondence to:

Edward W. Large
Assistant Professor
Center for Complex Systems and Brain Sciences
Florida Atlantic University
777 Glades Road, P.O. Box 3091
Boca Raton, FL 33431-0991, USA

large@walt.ccs.fau.edu
tel: 561.297.0106
fax: 561.297.3634

July 3, 2001

In Press: *Journal of New Music Research*

Abstract

This article describes an approach to metrical structure focussing on its role as an active listening strategy. The theory postulates that metrical structure is a self-organized, dynamic structure composed of self-sustaining oscillations. The emergence of this structural representation is modeled as a pattern formation process whose the neural correlate is the formation of a spatiotemporal pattern of neural activity. The primary function of the dynamic structure is attentional: it enables anticipation of future events thus, targeting of perception, and coordination of action with exogenous events. Stability and flexibility properties arise through nonlinearities in the underlying pattern-forming dynamics. Furthermore, this dynamic representation functions in musical communication. Transient stimulus fluctuations observed in musical performance (e.g. rate changes, intonation) are not noise, but rather communicate structural information, intention, and affect. These communicative gestures are recognized as deviations from temporal expectations embodied in the metrical structure. Experiments are reviewed that investigate stimuli of varying complexity, from simple isochronous tone sequences to performed music, and the model's success at capturing these data is assessed.

I: Periodicity, Pattern Formation, and Metric Structure

Listeners bring many different types of expectations to the musical experience. These include expectations about patterns of melody, harmony, compositional structure, social context, and even behavioral responses. Performers and composers create patterns of sound with reference to these expectations; in this way the listener can be said to play a causal role in the process of musical creativity. Thus, to understand the structure of music, it is helpful to understand the structure of listeners' expectations. Perhaps the most basic of these is the expectation of periodicity and, more generally, the expectation of stratified, multi-periodic structures. Structured temporal expectation, often called metrical structure, provides a framework around which other aspects of musical expectation are choreographed in time.

The question of metrical structure is an intriguing one because the notion of meter implies a perceptual structure that is stable enough to accommodate the astonishing temporal complexity of musical patterns, yet one that is flexible enough to adapt to local temporal fluctuations such as rubato, and to reorganize itself when stimulus structure changes. Furthermore, the challenge of articulating a viable theory of meter perception goes beyond that of describing a method for objective analysis of the musical material. The perception of metrical structure is not merely an analysis of rhythmic content, rather it shapes an active listening strategy in which the listener's expectations about future events can play a role as important as the musical events themselves. The percept of musical meter reflects a dynamic attending capacity, it is the experience of an inherently rhythmic component of auditory attention (Jones, 1976; Large & Jones, 1999).

This article considers the hypothesis that the temporal structure of listeners' expectations is a dynamic, self-organizing multi-periodic structure. It describes how the emergence of such structural representation can be modeled as a pattern formation process whose neural correlate is the formation of spatiotemporal patterns of neural activity. A primary function of this structure is attentional: it allows anticipation of future events, enabling perceptual targeting, and coordination of action with musical events. This article focusses on four main questions. How do metrical patterns form? How do structures reorganize in response to structural change? How do structures accommodate the temporal fluctuations found in music performance? How do metrical structures support communication between performer and listener?

II: A Model of Meter Perception

Let us begin by stating some background assumptions. The term *rhythm* will be used to refer to the pattern of timing and accentuation of a complex sequence, as shown in Figure 1. *Accent* means the phenomenal accent (cf. Lerdahl & Jackendoff, 1983) associated with each sounded event. Phenomenal accent arises through a combination of acoustic variables including pitch, loudness, timbre, and duration. In the study of rhythm perception and performance, a conceptual division is usually made between meter, a perceived temporal–accentual structure, and grouping, the segmentation of the auditory event into meaningful structural units, such as phrases. Relative timing and phenomenal accent both provide perceptual cues to both meter and grouping. Figure 1 illustrates both forms of rhythm organization.

Insert Figure 1 about here.

The basic element of meter is a beat, a series of perceived pulses marking subjectively equal units of time. Perceived beat is an inference from the acoustic stimulus (Cooper & Meyer, 1960; Lerdahl & Jackendoff, 1983), and functions as an expectation for when events are likely to occur in the future (cf. Large & Kolen, 1994). Meter is a more complex temporal accent scheme that describes the perception of regularly alternating strong and weak beats, sometimes described as the simultaneous perception of beats on multiple time scales (Lerdahl & Jackendoff, 1983; Yeston, 1976). Temporal locations at which beats of several levels coincide are termed *strong* beats, those at which few levels coincide are *weak*. Meter is a complex form of temporal expectation: Acoustic events are more strongly expected at strong beats, and less expected on weak beats (cf. Desain, 1992; Large & Jones, 1999). Grouping structure refers to the way the acoustic stream is parsed into meaningful structural units (Lerdahl & Jackendoff, 1983), as notated above the musical staff of Figure 1. A significant body of research has investigated cues to grouping, which include global pattern structure, local patterns of timing and accent, and performance cues (see e.g. Garner & Gottwald, 1968; Povel, & Okkerman, 1981; Palmer, 1997).

Formation and Stability of Metrical Structures

According to Cooper & Meyer (1960) a perception of beat generally arises in response to a periodicity present in the musical signal, but it is stable in the sense that once perceived it may continue even when the periodicity is interrupted in some way. To model this phenomenon, *self-sustained oscillation* is proposed as the basic model of musical beat. A self-sustained oscillation is autonomous in the sense that, once activated, it can persist, even after stimulation ceases or changes in significant ways. However, it also has the property that it entrains to incoming rhythmic signals. The simplest mathematical form for self-sustained oscillation is called the *Hopf oscillator* (Arrowsmith & Place, 1990), which can be written as a differential equation in the complex variable $z(t)$.

$$\dot{z} = \alpha z + i\omega z - z|z|^2 \tag{1}$$

Here the dot operator denotes the differentiation with respect to time (e.g. $\dot{z} = dz/dt$), α is an energy parameter, and ω is the eigenfrequency (inverse of period) of the oscillator. The behavior of this system is best understood by transforming to polar coordinates using the identity $z(t) = r(t)\exp i\phi(t)$, and separating real and imaginary parts. The transformation yields a system of two differential equations, describing the time evolution of amplitude and phase respectively.

$$\dot{r} = r(\alpha - r^2) \tag{1a}$$

$$\dot{\phi} = \omega \tag{1b}$$

The transformation reveals that the amplitude and phase of the Hopf oscillator are independent. Thus, by setting $\dot{r}$ to zero, we can find the stable states of the system. Such an analysis reveals two steady state behaviors; the value of the energy parameter determines which behavior is observed, as shown in the Figure 2. When α is less than zero (energy is being removed from the system), the system has a stable fixed point with r = 0, and the system behaves as a damped oscillator. For parameter values greater than zero (energy is being added into the system), a stable limit cycle develops; the system generates a self-sustained oscillation. The point α=0 is called a *bifurcation point*: It is the point at with the behavior changes qualitatively from damped oscillation to self-sustained oscillation. This type of bifurcation is called a Hopf bifurcation.

Insert Figure 2 about here.

Representation of metrical structure can be modeled by a network of oscillators with different periods, competing for activation through mutual inhibition. Oscillations that are most consonant with the input (see Large, 2000a) tend to deactivate those that do a poorer job of predicting the incoming rhythm. In response to a rhythmic pattern a few oscillations achieve stable activation: those that best summarize temporal structure of the stimulus. This pattern-formation process is captured with the addition of an interaction term into Equation 1:

$$\dot{z}_n = \alpha_n z_n + i\omega_n z_n - z_n|z_n|^2 - \sum_{m \neq n} \gamma_{mn} z_n |z_m|^2 \qquad (2)$$

where γ is an interaction matrix, $\gamma_{mn} \geq 0$. As above, transformation to polar coordinates shows that in Equation 2 amplitude and phase are independent.

$$\dot{r}_n = r_n(\alpha_n - r_n^2) - \sum_{m \neq n} \gamma_{mn} r_n r_m^2 \qquad (2a)$$

$$\dot{\phi}_n = \omega_n \qquad (2b)$$

As shown in Figure 3A, activation of one oscillator effectively moves the bifurcation point of the other oscillator. The γ matrix specifies the new activation thresholds relative to the energy parameters of competing oscillators. For example, in a two-oscillator network, when one oscillation is active, the second oscillator’s energy must exceed a certain proportion of the active oscillator’s energy in order for it to activate; the required proportion is given by the strength of the inhibition parameter. Figure 3B shows the four metrical patterns that can arise in such a network, and the conditions – in terms of relative parameter values – under which each exists and is stable.

Insert Figure 3 about here.

The pattern forming dynamics described by Equation 2 provides a number of interesting properties. First, it describes the formation of metrical structures. The simple analysis provided for the two oscillator network extends in a straightforward way to any number of oscillators, and it is possible to tie the parameters to the rhythmic stimulus in such a way that distinct metrical representations arise within the first few beats of a musical rhythm (Large, 2000a). It also makes predictions about the stability of metrical representations. Once a metrical pattern forms in the network it tends to persist, to remain stable even when a rhythmic pattern comes into conflict with its predictions. Among other things, this provides a basis for understanding phenomena such as syncopation. Given rhythmic contradiction that is strong or consistent enough, the network may also reorganize its metrical representation.

The formation and stability of patterns in such a system are best studied under conditions of *multistability.* Referring back to Figure 3B, note that the stability conditions for patterns 2 and 3 are not mutually exclusive[1]. The network cannot occupy both states simultaneously, however. This situation is called bistability, and in a larger network, multistable states are also possible. In its most general form, the theory predicts that certain rhythmic patterns can support more than one stable metrical representation (Large, 2000b). An observable consequence of multistability is *hysteresis*, the persistence of a structural representation despite stimulus parameters that would normally favor an alternative structural interpretation. In section III, evidence regarding the formation and stability of metrical representations is reviewed.

Synchronization and Communication

How does the rhythmic stimulus influence the network? Details of how musical rhythms are input are discussed in (Large, 2000a). In this article, a simplified model in considered, one that captures the phase dynamics of a single active oscillation in the network, driven by an external rhythm (Large & Kolen, 1994). This type of discrete time dynamical model is called a phase attractive circle map (cf. Kelso, deGuzman, & Holroyd, 1990). The series of phases produced by this circle map represents the phases of the oscillation at which auditory events occur. From a musical point of view, the primary feature of interest is that the oscillation is *phase coupled* to the stimulus, such that it synchronizes with certain stimulus events in the external signal.

$$\phi_{i+1} = \phi_i + \frac{IOI_i}{p} - \eta_\phi X_i F(\phi_i, \kappa) \qquad (mod_{-0.5,\,0.5} 1) \tag{3}$$

Equation 3 maps event onset times onto the phase of the internal oscillation as depicted in Figure 5, where p is the period of the oscillation and IOI_i is the i^{th} inter-onset interval ($t_{i+1} - t_i$) of the auditory event sequence. Incorporation of stimulus IOI's enables the model to predict responses to any rhythmic pattern (Large & Kolen, 1994). The phase coupling term, $-\eta_\phi X_i F(\phi_i, \kappa)$, models entrainment of the oscillator with the signal, where η_ϕ is coupling strength, X_i is the amplitude of the i^{th} event, and $F(\phi_i, \kappa)$ is a coupling function (Large & Kolen, 1994; Large & Palmer, 2001). This model can maintain synchrony with a musical rhythm even in the presence of small random timing fluctuations.

Insert Figure 4 about here.

Phase coupling alone, however, is not sufficient to model entrainment in the presence of large, systematic timing fluctuations, such as musical rubato. If the tempo of a musical performance changes too much, for example, the oscillator will simply lose synchrony. To maintain synchrony listeners must track tempo fluctuations. To explain this, it is necessary to assume that the period of the internal oscillation adapts to a changing stimulus tempo (Large & Kolen, 1994):

$$p_{i+1} = p_i + p_i \eta_p X_i F(\phi_i, \kappa) \tag{4}$$

Equation 4 serves as a simplified model of period adaptation, which corresponds to smooth movement of peak activations in the oscillator network of Equation 2 (cf. Figure 5A, below).

Once the musical stimulus enters the picture, it is possible to make predictions about the real-time tracking of auditory events. This model predicts synchronization with temporally complex stimuli and flexible adaptation to natural temporal fluctuations of the type found in performed music. Specific patterns of phase and period adaptation, should be observed in response to temporal stimulus perturbations. For simple metronomic stimuli, these tracking predictions are similar to those of linear error correction models (e.g Vorberg & Wing, 1996). However, the oscillator network model is more general. It includes nonlinear phase correction, models adaptation of period as well as phase, and accommodates rhythmically complex sequences. Finally, the model makes predictions about the perception of expressive timing. Transient tempo fluctuations observed in speech and musical performance (e.g. rate changes, intonation) communicate structure, intention, and affect. The theory models communication by assuming that listeners recognize communicative temporal gestures as deviations from temporal expectations embodied in attentional structures. In section III, I describe experiments that evaluate both aspects of the model's predictions about responses to rhythmic patterns.

III: Some Examples

In the previous section, a model was introduced that made a number of predictions. These included predictions about the formation of temporally structured representations, the persistence or stability of representations, real-time temporal tracking of stimulus fluctuations, and musical communication based on expressive performance timing. In general, these predictions can be evaluated with many different types of stimuli. Some of the most interesting involve perceiving and attending to natural, performed musical rhythms. Thus, one approach to understanding these issues is to investigate perception of music performances directly. In addition, using simpler sequences, one can make more controlled observations regarding, for example, temporal tracking. In this section I describe how both types of stimuli are used to evaluate the predictions of the nonlinear pattern formation approach to the perception of metrical structure.

Formation of Metrical Structures

One way to assess the formation of metrical structures is to ask participants to explicitly describe their metrical perceptions. In one such study, Snyder and Krumhansl (2000) investigated the nature of mental pulse-finding using a synchronization tapping task. They presented musicians with excerpts of eight ragtime piano pieces, and asked them to tap the most comfortable pulse of each excerpt on a piano keyboard. In Snyder and Krumhansl's study, the musical excerpts were generated by computer so that the timing of the individual events was metronomically precise. The pieces were played to participants via MIDI, using a digitally sampled piano timbre. To control different types of pulse-finding cues that were available, four versions of each piece were generated. First, full versions (both LH and RH parts) and right-hand only (RH) versions were created. The elimination of the left hand part created a version of the music that was more syncopated, and potentially more metrically confusing than the full version, because fewer events occurred on the downbeats. Next, based upon each of these two versions, two additional stimuli were created from which non-temporal cues were eliminated. In these, all pitches were changed to middle C, creating monotonic versions of the excerpts. The elimination of pitch information created versions of the music in which pitch cues could not be used to resolve metrical ambiguity, thus these versions served as a gauge of the role of pitch information in structure formation.

To test the model, Snyder & Krumhansl's MIDI recordings were played on a Kawai 950 digital piano, and the acoustic signal was passed through a simple auditory model, approximating processing in the earliest stages of the auditory system: frequency filtering by the cochlea, mechanical to neural transduction at the level of the primary afferents, and onset responses, which can be observed as early as the cochlear nucleus (cf. Scheirer, 1998; Todd, 1994). This process produced a rhythmic signal comprised of event onsets. The onsets drove the network through a multiplicative coupling that tied the energy parameter, α, of each oscillation to its success in predicting upcoming events in the rhythm (Large, 2000a). Using this approach, model predictions were compared with the results of Snyder & Krumhansl's (2000) experiment.

The network makes several predictions, two of which are shown in Figure 5. First, a pattern of oscillations arises in response to the musical rhythm, forming a dynamic representation of metrical structure. This can be seen in Panel A, which summarizes network performance for the entire excerpt as average amplitude. Four peaks in network amplitude are evident, corresponding to oscillations at the 1/16, 1/8, 1/4, and 1/2 note metrical levels. This is one of many different metrical patterns that can be stabilized within this network. The spectral properties of the rhythm in conjunction with the inhibitory network interactions gave rise to this four-leveled, duple metrical structure.

Insert Figure 5 about here.

But amplitude is only half of the story, the sustained oscillations also synchronize to the rhythm, predicting phase (Panel B). Snyder & Krumhansl measured the phase – the location within the measure – at which people tapped along with the excerpts, i.e. whether they tapped in the downbeat, upbeat, neither (at some other metrical location), or in an aperiodic fashion. The musicians almost always tapped on the downbeat for the full pitched versions, but tapping performance broke down as information was removed from the stimulus. To model this, the highest amplitude oscillation (the 1/4 note level) was used to predict the period at which people would tap along with the rhythm. The network predicted tapping phase well for natural musical signals, and although as a whole it produced taps less often on the downbeat than the musicians, the network produced a similar pattern of breakdowns as accompaniment (LH) and pitch information was removed from the stimulus.

Overall, when stimulated with rhythmic input, the pattern-forming dynamical system of Eq 2 gave rise to metrically structured patterns of oscillation. It provided reasonable predictions of human performance for natural music, and network performance deteriorated similarly to human performance as stimulus information is degraded. For both the model and for the musicians, increased syncopation was more disruptive to synchronization than lack of pitch information. Thus, this test suggests that the nonlinear pattern formation approach may be on the right track. However, these stimuli did not contain any temporal fluctuation, thus predictions about tempo tracking cannot be assessed. Furthermore, the musicians tapped at only one level metrical level, leaving the network's predictions about multiple structural levels untested. Using other techniques, however, it is possible to gauge other aspects of the network's predictions.

Stability of Metrical Structures

How can we assess which of several possible metrical interpretations is heard for a given rhythm? To ask this question, it is necessary to identify a task that directly measures the formation of multi-leveled structures. Evidence suggests that the categorization of rhythmic patterns provides such a task, one that taps directly into listeners' dynamic representations of metrical structure. Clarke (1987) demonstrated that ambiguous temporal ratios (between 1:1 and 2:1) were more likely to be categorized as 2:1 in the context of triple meter, whereas these same ratios were more likely to be categorized as 1:1 in the context of duple meter. Clarke interpreted his results in terms of two processes operating in rhythm perception. One places temporal intervals into categories depending on the prevailing metrical context, the other interprets deviations from categorical precision as musically expressive gestures.

Furthermore, in dynamical systems terms, Clarke's data provide evidence of *hysteresis* in meter perception, the persistence of a percept (e.g. a duple meter) despite a change in the stimulus that favors an alternative pattern (e.g. a triple meter). As discussed above (see Figure 3B) hysteresis occurs in situations of multistability. Multistability and hysteresis are fundamental predictions of the nonlinear pattern formation theory, thus tests of these predictions provide key assessments of this approach. In one study, a rhythmic pattern composed of three event onsets was gradually changed such that its metrical interpretation would be altered (Large, 2000b). On *increasing* trials, a 1:1 time interval ratio, implying a duple metrical structure, was gradually changed to a 2:1 ratio, favoring a triple metrical structure (Figure 6A); on *decreasing* trials, the 2:1 ratio gradually changed to a 1:1 ratio. Participants were asked to categorize the rhythmic pattern as either *duple* or *triple*. Because it was possible that some intermediate patterns would be heard as neither duple nor triple, a third response category, *neither,* was included. Also, the potential for confounding perceptual hysteresis with response hysteresis was minimized using the modified method of limits procedure (Hock, Schöner, & Kelso, 1993). Changes occurred within the context of a rhythmic sequence, such that the model predicted that a two-leveled metrical structure would be perceived, as shown in Figure 6B (an *increasing* trial is pictured).

Insert Figure 6 about here.

To produce model predictions, it was assumed that when two oscillations were active, one with a period corresponding to the base interval, and the second with one-half the period of the base interval, the model (i.e. Equation 2) responded “duple”. When the second oscillation had one-third the period of the base interval, the model responded “triple”. When only one oscillation – corresponding to the base interval – was active, the model responded “neither”. All three oscillations competed with one another. Figure 6C compares representative predictions of the model with experimental results. Here, the perceptual boundary between the categories “duple” and “not duple” are compared in the decreasing context (subject was asked to respond as soon as pattern was clearly duple) and the *increasing* context (subject was asked to respond as soon as pattern became anything other than clearly duple) and for one subject. For both the model and the subject hysteresis was observed. On *increasing* trials, both persisted in the perception of “duple” for quite some time; on *decreasing* trials both persisted in judgements of “not duple” until the stimulus is very nearly true 1:1. This result provides strong evidence of multistability, supporting the nonlinear pattern formation approach.

The results of this study are consistent with the general prediction of hysteresis in meter perception, and specific patterns of results can be used to fit model parameters for each subject. Other issues remain to be investigated, however. For one, studies such as this one simply assume that categorization is based on an underlying representation of metrical structure. In principle, however, this assumption could itself be tested. Additionally, if some rhythms truly do possess bistable metrical interpretations, then even more dramatic effects should be observed in more complex musical sequences. Finally, this study addressed only the first of the two processes proposed by Clarke (1987), categorization of an imperfect temporal interval according to a stable metrical framework. But what happens when the tempo of a sequence changes? How do people track the temporally fluctuating stimulus? How do temporal fluctuations communicate musical expression?

Tracking Temporal Fluctuations

Real-time tracking is most directly measured using tasks such as synchronization tapping, in which people are required to coordinate bodily movements with auditory patterns. The ability to coordinate motor behaviors with auditory stimuli has been widely studied (e.g. Bartlett & Bartlett, 1959; Dunlap, 1910; Fraisse, 1956 Kohlers & Brewster, 1985; Michon, 1967; Stevens, 1886; Vos, Mates, & van Kruysbergen, 1994; Woodrow, 1932). Most studies have investigated synchronization to simple, structurally isochronous sequences. These have uncovered several robust behavioral phenomena, and various models have been proposed to account for these findings, which include rate dependence, anticipation bias, patterns of variability, compensation for phase perturbations, and learning (e.g. Chen, Ding, & Kelso, 1997; Hary & Moore, 1987; Kelso, DelColle, & Schöner, 1990; Mates, Radil, & Pöppel, 1992; Pressing & Jolley-Rogers, 1997; Semjen, Vorberg, & Schulze, 1998; Vorberg & Wing, 1996, Zanzone & Kelso, 1992). However, such models have yet to be convincingly extended to more complex and naturalistic stimuli, such as music. Even the simplest musical performance presents considerably more complexity than the ticks of an auditory metronome. And while a great deal is known about synchronization with simple acoustic sequences, relatively little is known about how people accomplish the similar feat of synchronization with more complex auditory stimuli.

In one study, Large, Fink, & Kelso (2001) addressed issues of synchronization to rhythms that approached musical complexity. In two experiments, synchronization to simple and more complex rhythmic sequences was investigated. Their first experiment examined responses to phase and tempo perturbations within simple, structurally isochronous sequences, presented at different base rates. Phase and tempo perturbations were embedded, at random intervals, in otherwise isochronous tone sequences. Subjects were asked to synchronize finger taps to these sequences. Figure 7 shows model predictions (top) and representative data for a single subject (bottom) responding to a phase perturbation (left) and a tempo perturbation (right), where perfect synchrony is indicated by $\phi = 0$, an early tap by $\phi < 0$ and a late tap by $\phi > 0$. The overshoot in the response to the phase perturbation combined with the smoother relaxation profile for the tempo perturbation are evidence for period adaptation, and are nicely predicted by the model (Equations 3 & 4).

Another experiment investigated responses to similar perturbations embedded within more complex, metrically structured sequences; participants were explicitly instructed to synchronize at different metrical levels (i.e. tap at different rates to the same rhythmic patterns) on different trials. Large, et. al. (2001) found evidence that people can synchronize with unpredictable, metrically structured rhythms at different metrical levels, with qualitatively different patterns of synchronization seen at higher versus lower levels of metrical structure. Furthermore, intrinsic tapping frequency adapts in response to temporal perturbations in these complex rhythms, and synchronization at each tapping level reflects information from other metrical levels.

Insert Figure 7 about here.

This study provided evidence for a dynamic and flexible internal representation of the sequence's metrical structure. The unpredictability of the patterns in Experiment 2 rules out the possibility that coordination of action depended upon learning of specific rhythmic patterns (e.g. Vorberg & Hambuch, 1984; Pressing, 1999). Furthermore, successful recovery from phase and tempo perturbations in unpredictable rhythms disconfirms the hypothesis that temporal tracking is based solely upon prior learning of specific sequences of auditory events (cf. Cottrell, Nguyen, & Tsung, 1993). Taken together with the observation that musicians and non musicians performed equivalently, the basic finding points to a natural, inherently rhythmic faculty, that is not based on learning of specific event sequences, rhythmic patterns, or response strategies. But what about perception: how do temporal fluctuations function in musical communication?

Musical Communication

The temporal fluctuations observed in naturally performed music are not arbitrary temporal perturbations as in the above study. Rather, the transient fluctuations observed in musical performance communicate structure, intention, and affect. Performers use temporal fluctuations to communicate musical interpretation, including interpretations of phrase structure, melody, and meter (for a recent review, see Palmer, 1997). Indeed violation of musical expectations, in general, is thought to be critical in the communication of emotion and meaning in music (e.g. Meyer, 1956;

Narmour, 1990). Furthermore, certain aspects of temporal structure have been linked to communication of emotion and affect, even across performers and listeners of different musical cultures (Balkwill & Thompson, 1999).

Deviations from a regular beat or pulse in music performance, sometimes called rubato, are often largest near phrase boundaries. One study of music performance examined the effects of phrase structure on temporal fluctuations in piano performances (Palmer & van de Sande, 1995). In this study, performances of polyphonic music by Bach (two- and three-part inventions) which contained multiple voices were collected on a computer-monitored acoustic piano. Pianists performed the same musical pieces in terms of three different phrase structures as marked in different versions of the music notation; in a control condition, there were no marked phrase boundaries. Palmer & van de Sande (1995) found that performers reliably used rubato at intended phrase boundaries.

This result provides two challenges to the oscillator model. First, beat tracking in the presence of rubato provides a test of the model's ability to adapt to a changing tempo in a real performance. Second, the model's sensitivity to deviations from its temporal expectations can be compared with performers' structural intentions to further test how the temporal fluctuations in music performance may be informative for listeners. To do this, the model must both categorize temporal intervals in the presence of musical rubato, while at the same time registering deviations from expected timing.

Phrase boundary detection was accomplished using an extension of the model (Equations 3 & 4) described by (Large & Jones, 1999). Events occurring at expected times are processed more accurately and efficiently than events occurring at less expected times (Large & Jones, 1999). This is captured as a concentration of attentional resources about expected time points called an attentional pulse. A parameter called attentional focus indexes the concentration of expectancy about the expected time point ($\phi = 0$), and focus adapts dynamically based upon the accuracy of the oscillator's predictions. The salience of a temporal deviation (for example, an event performed later than expected) increases not only with the size of the deviation but also with the temporal regularity of the performance. Large & Jones successfully used this framework to predict performance in time discrimination tasks.

In an application to musical performance (Large & Palmer, 2001), a two-oscillator model first tracked the fluctuating temporal structure, categorizing event onsets according the meter, and then determined the probability that individual events marked phrase boundaries as shown in Figure 7. The model performed remarkably well in finding intended phrase boundaries, predicting perceived phrasing, thus modeling communication between performer and listener.

Insert Figure 8 about here.

IV: Conclusions and Future Directions

The theoretical approach described in this article model the perception of metrical structure as a nonlinear pattern formation process that can be driven by a rhythmic stimulus. Many features of the patterns that form in the network described here match those that were reported by musicians, at least those that were tested in the case of ragtime piano music. Stability properties of perceived metrical structure – a measure of independence from the rhythmic signal that drove formation of the original pattern, the tendency to remain stable in the face of rhythmic conflict – are captured in this framework by nonlinearities in the pattern forming system. These same properties allow the structure to yield when rhythmic conflict becomes strong or consistent enough, as seen in the categorization study. In any such system, the basic signature of a nonlinear mechanism is hysteresis – persistence of a percept despite a change in the stimulus that favors the alternative percept. Hysteresis was reported here in one experiment, and musical intuition suggests that many more interesting cases of rhythmic multistability remain to be investigated in the laboratory.

Rubato also provides an interesting demonstration of the flexibility of metrical percepts. As Clarke (1987) observed, people possess the ability not only to categorize imperfect temporal intervals, following the meter despite tempo changes, but they also respond meaningfully to deviations from temporal expectations, understanding the meaning of musically expressive temporal gestures. The theory described here accommodates rubato by postulating a period adaptation mechanism, while at the same time capturing the response to expressive timing using a simple time discrimination model. The period adaptation model accurately predicted the response to artificial tempo perturbations, and also fared well in tracking temporal fluctuations in expressively timed musical performances. In the latter, the model was also able to reliably detect intended phrase boundaries, capturing one aspect of musical communication between performer and listener.

Although aspects of this approach seem muscially intuitive, and several predictions stand up to empirical scrutiny, the models presented in this article are merely steps along the path to a more complete understanding of rhythm perception. The mathematical models described here are formulated at a level that captures the phenomenology of meter perception directly, without modeling the details of the neural structures that give rise to this experience. The Hopf oscillator, for example, is a normal form dynamical system, a mathematical abstraction that discards the details of particular physical systems, succinctly capturing the basic properties that are shared among a family of more complex differential equations. Normal form dynamical systems are, in a rather specific mathematical sense, the simplest systems that capture the basic phenomenon under study (Arrowsmith & Place, 1990). Thus the pattern forming system of Equation 2 is, in a sense, the simplest dynamical system that will exhibit the properties that we have exploited here to model meter perception.

There are two main advantages to modeling at the phenomenological level. The model systems that are developed are susceptible to mathematical analysis, and predictions are made at a level that is appropriate to behavioral investigation. Thus, for example, the predictions of hysteresis in meter perception (Figure 3B) can be worked out analytically, such that behavioral data can be used to fit model parameters. The circle map model (Equations 3 & 4) is a further simplification of the normal form model, taking external input into account (see Large, 2001). The discrete-time phase equation is a straightforward discretization of the continuous phase dynamics, and period adaptation replaces amplitude dynamics with the assumption that the period of each network oscillation can adapt smoothly in response to tempo changes. These forms of the model support numerical analysis (e.g. Large & Kolen, 1994, deGuzman & Kelso, 1991), and they are particularly convenient for modeling synchronization performance, time discrimination, and the perception of expressive timing.

Future modeling work should proceed along two fronts. The first concerns the use of information that goes beyond event onsets. The current models rely only on onset timing, amplitude, and frequency information, which is recoverable from acoustic recordings, and has been used as input to continuous-time models such as Large (2001) of Scheirer (1998). However, such models do not make use of all of the information relevant to human meter perception, which includes information about pitch, melody, and harmony. Such information is not yet fully resolvable from acoustic recordings, thus other avenues will have to be investigated. Here, the discrete-time formulation offers some advantages over its continuous-time cousin, for example the ability to work directly with MIDI recordings. Modeling using MIDI signals could take advantage of information that is available to the auditory system, without first solving the equally difficult problem of how the auditory system resolves such information For example, Toiviainen (1998), uses MIDI information to drive a continuous-time model that takes event duration into account.

Second, theories of meter perception should move toward the development of more neurally realistic models. Single neuron Hogkin-Huxley type equations are currently being investigated for modeling of interval selective auditory neurons (e.g. Large & Crawford, 2001; Crawford, 1997). The disadvantage of such models is that their analytical and computational complexity makes predictions regarding high-level phenomena such as meter perception difficult to directly evaluate. Recently, Eck (2000) has described a model of meter perception based on a single neuron model known as the FitzHugh-Nagumo oscillator. The FitzHugh-Nagumo oscillator is essentially a mathematical simplification of the Hogkin-Huxley model, one that is more tractable, both analytically and computationally. Other possibilities for neural modeling include so-called neural field equations (e.g. Jirsa & Haken, 1996; Amari, 1977), which attempt to capture population-level dynamics of neural ensembles. Such model will become especially relevant as neural imaging data becomes available for musical rhythm perception (cf. Mayville, et. al., 1999; Meaux, 2000).

Another promising area in the study of rhythm perception is, of course, is language. Similar descriptions of meter and grouping have been advanced by linguists and music theorists (e.g. Hayes, 1984; Lerdahl & Jackendoff, 1983; Liberman, 1975; Selkirk, 1984; Yeston, 1976), where direct analogies are often made between the rhythmic organization of music and speech. Although simple approaches to timing in language (e.g. “stress” versus “syllable” timing; Abercrombie, 1967; Pike, 1945) have not received strong empirical support (Hoequist, 1983; Roach, 1982), timing in music is significantly more complex and flexible that is commonly assumed. In fact, it is remarkable that listeners are able to perceive durational categories corresponding to the eighth-notes, quarter-notes, half-notes, and so forth, of musical notation because the actual durations measured in music performance deviate greatly from notated categorical durations (Clarke, 1987; Longuet-Higgins & Lee, 1982). Temporal fluctuation is commonly observed in all auditory communication, (*time-warping* in speech, *expressive timing* in music), and temporal perturbations are not noise. Rather they communicate information about things like grouping in both speech and music (Lehiste, 1977; Price et. al., 1991; Palmer, 1989; Shaffer, Clarke & Todd, 1985). Overall, transient stimulus fluctuations (temporal fluctuation, phenomenal accentuation) are quite important in auditory communication, distinguishing pragmatic categories of an utterance (e.g. statement vs. question), signalling focus, marking the boundaries of structural units, and communicating affect. Recent studies have already begun to investigate the relevance of dynamical models for linguistic rhythm (e.g. Cummins & Port, 1999).

To sum up, from the listener’s point of view, meter is more than an objective analysis of a rhythmic input signal. The perception of metrical structure is a stable dynamic pattern that underlies a dynamic attentional strategy. Attention, in this view, is not a filter that protects a limited capacity information processing mechanism (cf. Broadbent, 1958), rather is a process of selection for action (cf. Allport, 1989). Dynamic attending, as conceived here, aids in selection of temporally coherent auditory events for interaction, whether it be participating in an improvisation, coordinating toe-taps with a drummer, taking turns in a conversation, or simply focussing on a piano accompaniment for more careful scrutiny. The theory explains how people maintain a stable attentional focus over temporally extended events while flexibly adapting to transient temporal fluctuations. It provides mathematical models of dynamic structural representation, meaningfully extending previous approaches to auditory attending. It makes predictions about general properties of the neural correlates of auditory representation, attention, and communication. Finally, it applies to complex, temporally structured event sequences, explaining how people respond to the auditory complexity of the real world.

Abercrombie, D. (1967). *Elements of general phonetics*. Edinburgh: Edinbutgh Univ. Press.

Allport, A. (1989). Visual attention. In *Foundations of Cognitive Science*, pp. 631-682. In M.I. Posner (Ed.). Cambridge, MA: MIT Press.

Amari, S. (1977). Dynamics of pattern formation in lateral-inhibition type neural fields. *Biological Cybernetics*, 27, 77-87.

Arrowsmith, D.K. & Place, C.M. (1990). *An Introduction to Dynamical Systems*. Cambridge: Cambridge University Press.

Balkwill, L. L. & Thompson, W. F. (1999). A cross-cultural investigation of the perception of emotion in music: Psychophysical and cultural cues. *Music Perception, 17 (1)*, 43-64.

Bartlett, N. R., & Bartlett, S. C. (1959). Synchronization of a motor response with an anticipated sensory event. *Psychological Review*, 66, 203-218.

Broadbent (1958). Perception and Communication. London: Pergamon Press.

Chen, Y., Ding, Mingzhou, & Kelso, J. A. S. (1997). Long Memory Processes (1/fa Type) in Human Coordination. *Physical Review Letters*, *79* (*22*), 4501-4504.

Clarke, E. F. (1987). Categorical rhythm perception: An ecological perspective, In A. Gabrielsson (Ed.) Action and Perception in Rhythm and Music, pp 19-33. The Royal Swedish Academy of Music, 55.

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.

Crawford, J. D. (1997) Feature detection by auditory neurons in the brain of a sound-producing fish. *J. Comp. Physiol. A*, *180*, 439-450.

Cummins, F. & Port, R.F. (1996). Rhythmic constraints on stress timing in English. Journal of Phonetics, 26, 145-171.

deGuzman, G. C., & Kelso, J. A. S. (1991). Multifrequency behavioral patterns and the phase attractive circle map. *Biological Cybernetics*, *64*, 485-495.

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

Dunlap, K. (1910). Reactions to rhythmic stimuli, with attempt to synchronize. *Psychological Review*, 17, 399-416.

Eck, D. (2000). Meter Through Synchrony: Processing Rhythmical Patterns with Relaxtion Oscillators. Unpublished doctoral dissertation, Indiana University, Bloomington, IN.

Fraisse, P. (1956). *Les structures rhythmiques*. Louvain: Publication Universitaires de Louvain.

Garner, W. R., & Gottwald, R. L., (1968). The perception and learning of temporal patterns. *Quarterly Journal of Experimental Psychology, 20*, 97-109.

Hary, D. & Moore, G.P. (1987). Synchronizing human movement with an external clock source. *Biological Cybernetics, 56*, 305-311.

Hayes, B. (1984). The phonology of rhythm in Ecnglish. *Linguitic Inquiry, 15*,. 33-74.

Hock, H. S., Schöner, G., & Kelso, J. A. S. (1993). Bistability and hysteresis in the organization of apparent motion patterns. Journal of Experimental Psychology: Human Perception & Performance, 19, 63-80.

Hoequist, C. (1983). Syllable duration in stress-, syllable- and mora-timed languages. *Phonetica, 40*, 203-237.

Idson, W. L. and Massaro, D. W. (1976). Cross-octave masking of single tones and musical sequences: The effects of structure on auditory recognition. Perception and Psycholphysics, 19, 155-175.

Jirsa, V. K., & Haken, H. (1996). Derivation of a field equation of brain activity. *Journal of Biological Physics, 22*, 101-112.

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

Kelso, J. A. S., & DelColle, J. D., & Schöner, G. (1990). Action perception as a pattern formation process. In *Attention and Performance XII*I (pp. 139-169). Hilldale, NJ: Erlbaum.

Kelso, J. A. S., deGuzman, G. C., & Holroyd, T. (1990). The self-organized phase attractive dynamics of coordination. In A. Babloyantz (Ed), *Self Organization, Emerging Properties, and Learning* (pp. 41-62). NATO ASI Series B: Physics, Vol. 260.

Kohlers, P. A., & Brewster, J. M. (1985). Rhythms and responses. *Journal of Experimental Psychology: Human Perception and Performance,* 11, 150-167.

Large, E. W. (1994). *Dynamic representation of musical structure*. Unpublished Ph.D. dissertation. The Ohio State University.

Large, E. W. (1996). Large, E. W. (1996). *Modeling beat perception with a nonlinear oscillator.* In Proceedings of the Eighteenth Annual Conference of the Cognitive Science Society.

Large, E. W. (2000a). On synchronizing movements to music. Human Movement Science, *19*, 527-566.

Large, E. W. (2000b). Rhythm categorization in context. In *Proceedings of the International Conference on Music Perception and Cognition*, August.

Large, E. W., & Kolen, J. F. (1994). Resonance and the perception of musical meter. *Connection Science, 6,* 177 - 208.

Large, E. W., and Jones, M. R. (1999). The dynamics of attending: How we track time varying events. *Psychological Review, 106* (*1*), 119-159.

Large, E. W., and Palmer, C. (2001). Perceiving Temporal Regularity in Music. In Press *Cognitive Science*.

Large, E. W., and Palmer, C. (2001). Perceiving Temporal Regularity in Music. In Press *Cognitive Science*.

Large, E. W., Fink, P. & Kelso, J. A. S. K. (2001). Tracking simple and complex rhythms. *Psychological Science*, (under review).

Large, E.W., and Crawford, J.D. (2001). Auditory temporal computation: Interval selectivity based on post-inhibitory rebound. Submitted to Journal of Computational Neuroscience.

Lehiste, I. (1977). Isochrony reconsidered. *Journal of phonetics, 5*, 253-263.

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

Liberman, A. M., Harris, K. S., Hoffman, H. S., Griffith, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries. Journal of Experimental Psychology, 54, 358 - 368.

Liberman, M. (1975). *The intonational system of English*. Unpublished Ph.D. Dissertation, MIT.

Longuet-Higgins, H. C., & Lee, C. S. (1982). The perception of musical rhythms. *Proceeding of the Royal Society of London B*, *207*, 187-217.

Mates, J., Radil, T., & Pöppel, E. (1992). Cooperative tapping: Time control under different feedback conditions. *Biological Cybernetics*, *70*, 463-473.

Mayville, J. M., Bressler, S. L., Fuchs, A., Kelso, J. A. S. (1999). Spatiotemporal reorganization of electircal activity in the human brain associated with a timing transition in rhythmic auditory-motor coordination. *Exp Brain Research*, *127*, 371-381.

Meaux, J. M. (2000). Ph.D. Dissertation, Center for Complex Systems and Brain Sciences, Florida Atlantic University (in preparation).

Meaux, J. M. (2000). Ph.D. Dissertation, Center for Complex Systems and Brain Sciences, Florida Atlantic University (in preparation).

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

Michon, J. A. (1967). *Timing in temporal tracking*. Assen, NL: van Gorcum.

Naatanen, R., & Winkler, . (1993). The concept of auditory stimulus representation in cognitive science. *Psychological Bulliten*, *125*, 826-859.

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

Palmer, C. & van de Sande, C. (1995). CP: Need Ref.

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

Palmer, C. (1997). Music performance. *Annual Review of Psychology, 48*, 115-138.

Palmer, C., & Holleran, S. (1994). Pitch, harmonic, Melodic, and frequency height influences in the perception of multivoiced music. *Perception and Psychophysics*, *56 (3)*, 301-312.

Patel, A. D., Peretz, I., Tramo, M., and Labreque, R. (1998). Processing prosodic and musical patterns: A neuropsychological investigation. Brain Lang, 61, 123-144.

Peretz, I., & Morais, J. (1989). Music and modularity. *Contemporary music review, 4*, 279-293.

Pike, K. (1945). *The intonation of American English*. Ann Arbor: University of Michigan Press.

Pinker, S. (1997). *How the mind works*. W. W. Norton & Company.

Povel, D., & Okkerman, H. (1981). Accents in equitone sequences. *Perception& Psychophysics*, *7*, 565-572.

Pressing, J. & Jolley-Rogers, G. (1997). Spectral Properties of human cognition and skill. *Biological Cybernetics* 76 (5) 339-347.

Pressing, J. (1999). The Referential Dynamics of Cognition and Action. *Psychological Review*, 106 (4), 714-747.

Price, P. J. & Ostendorf, M., Shattuck-hufnagel, S & Fong, G. (1991). The use of prosody in syntactic disambiguation. *Journal of the Acoustical Society of America, 90*, 2956-2970.

Roach, P. (1982). On the distinction between "stress-timed" and "syllable-timed" languages. In D. Crystal (Ed.), *Linguistic controversies: Essays in linguistic theory and practice in honour of F. R. Palmer.* London: Arnold.

Scheirer, E.D. (1998). Tempo and beat analysis of acoustic musical signals. *Journal of the Acoustical Society of America*, *103*, 588-601.

Selkirk, E. (1980). *Phonology and Syntax: The relation between sound and structure*. Cambridge, MA: MIT Press.

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

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

Snyder, J. & Krumhansl, C. L. (2000). Tapping to ragtime: Cues to pulse-finding. *Music Perception*, (in press).

Stevens, L. T. (1886). On the time sense. *Mind, 11*, 393-404.

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.

Toiviainen, P. (1998). An interactive MIDI accompanist. *Computer Music Journal, 22*, 63-75.

Vorberg, D. & Wing, A. (1996). Modeling variability and dependence in timing. In H. Heuer & S. W. Keele (Eds) *Handbook of Perception and Action, Volume 2: Motor Skills.* (pp.181-262). London: Academic Press.

Vorberg, D., & Hambuch, R. (1984). Timing of two-handed polyrhythmic performance. In J. Gibbon and L. Allan (Eds.) Timing and Time Perception. New York Academy of Sciences, New York, pp 390-406.

Vos, P. G., Mates, J., & van Kruysbergen, N. W. (1995). The perceptual centre of a stimulus as the cue for synchronization to a metronome: Evidence from asynchronies. *The Quarterly Journal of Experimental Psychology, 48A*, 1024-1040.

Woodrow, H. (1932). The effects of rate of sequences upon the accuracy of synchronization. *Journal of Experimental Psychology, 15,* 357-379.

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

Zanzone, P. G., & Kelso, J. A. S. (1992). Evolution of behavioral attractors with learning: Nonequilibrium phase transitions. *Journal of Experimental Psychology: Human Perception & Performance*, *18,* 403-421.

Acknowledgments

This research was supported in part by National Science Foundation grant SBR-9809446.

Reference Notes

[1]For example, when $\alpha_1 = \alpha_2 = 1$ and $\gamma_{2,1} = \gamma_{1,2} = 1.5$, pattern 2 and pattern 3 are both stable.

Figure Captions

Figure 1. Opening Aria from Bach's Goldberg Variations. The metrical structure (below staff) is triple: strong-weak-weak. The grouping structure (above staff) marks two phrases.

Figure 2. A Hopf oscillator models activation of a self-sustained oscillation (Equation 1). For parameter values less than zero, the system generates a damped oscillation; for values greater than zero a stable limit cycle develops: the system generates a self-sustained oscillation. α=0 is the bifurcation point, the parameter value at with the behavior changes qualitatively.

Figure 3. Competition (Equation 2) in a two-oscillator network. (A) Competition from oscillator 2 moves the bifurcation point of oscillator 1. (B) Metrical patterns and associated stability conditions for the two-oscillator network. Four metrical patterns can arise this simple network. The stability condition for patterns 2 and 3 are not mutually exclusive, thus this parameter region is bistable.

Figure 4. Relative phase and the circle map. Time, *t*, is mapped onto phase, ϕ, by Equation 3 such that the expected onset time, $t = t_x$ is transformed to an expected phase $\phi = 0$. Circle maps work in relative phase directly, mapping points on the circle to new points on the circle ($\phi_n \rightarrow \phi_{n+1}$). Adapted from (Large & Jones, 1999).

Figure 5. (A) A temporal pattern of oscillations arises in response to a musical rhythm, dynamically representing its meter: a binary structure with 4 metrical levels (amplitude peaks are shown). The most prominent pulse level (the highest peak) predicts the period at which people will tap along with the rhythm. (B) The oscillators also synchronize to the rhythm, predicting the phase at which people will tap, and the patterns of phase instability that occur as pitch information is removed (Monotonic) and as rhythmic complexity is increased (RH). Data from Snyder & Krumhansl (2000).

Figure 6. The categorization experiment: stimuli and results. (A) 1:1 ratios imply duple metrical structure, 2:1 ratios imply triple structure. Rhythmic patterns were gradually changed from duple to triple, and subjects were asked to categorize them as *duple*, *triple* or *neither.* (B) An increasing trial: the duple rhythmic pattern gradually changes toward a triple rhythmic pattern. (C) The model predicts hysteresis in the boundary between *duple* and *not duple*. Data from one subject also shows hysteresis at this boundary.

Figure 7. The model (Equations 3 & 4) predicts motor tracking responses to phase and tempo perturbations in isochronous sequences. Both model and data show overshoot in response to phase perturbations, and smoother relaxation from tempo perturbations, indicating an internal period correction process.

Figure 8. (A) Categorization of note events (dark = strong beat, light = weak beat), and (B) probability of lateness (see Figure 4) relative to categories. The intended phrasing (instructions to performer) is shown above. The network predicts perceived phrasing, modeling communication between performer and listener (Adapted from Large & Palmer, 2001).

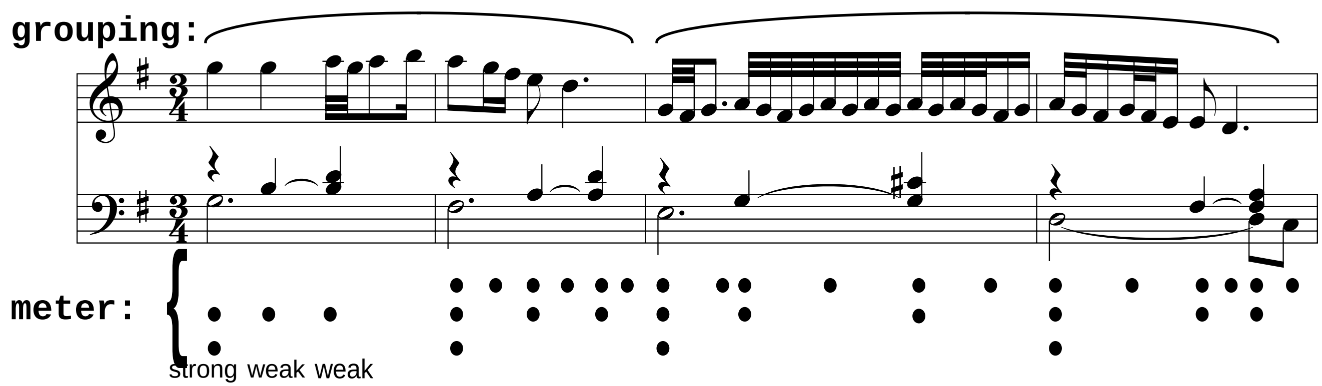


Figure 1

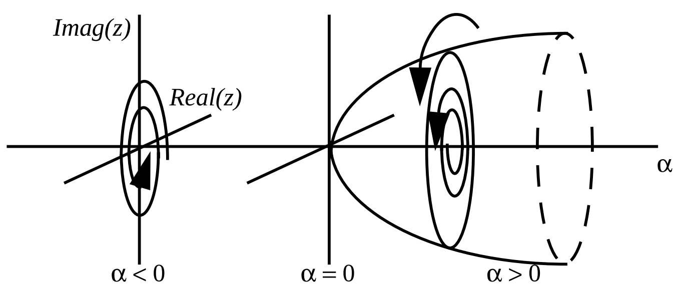


Figure 2

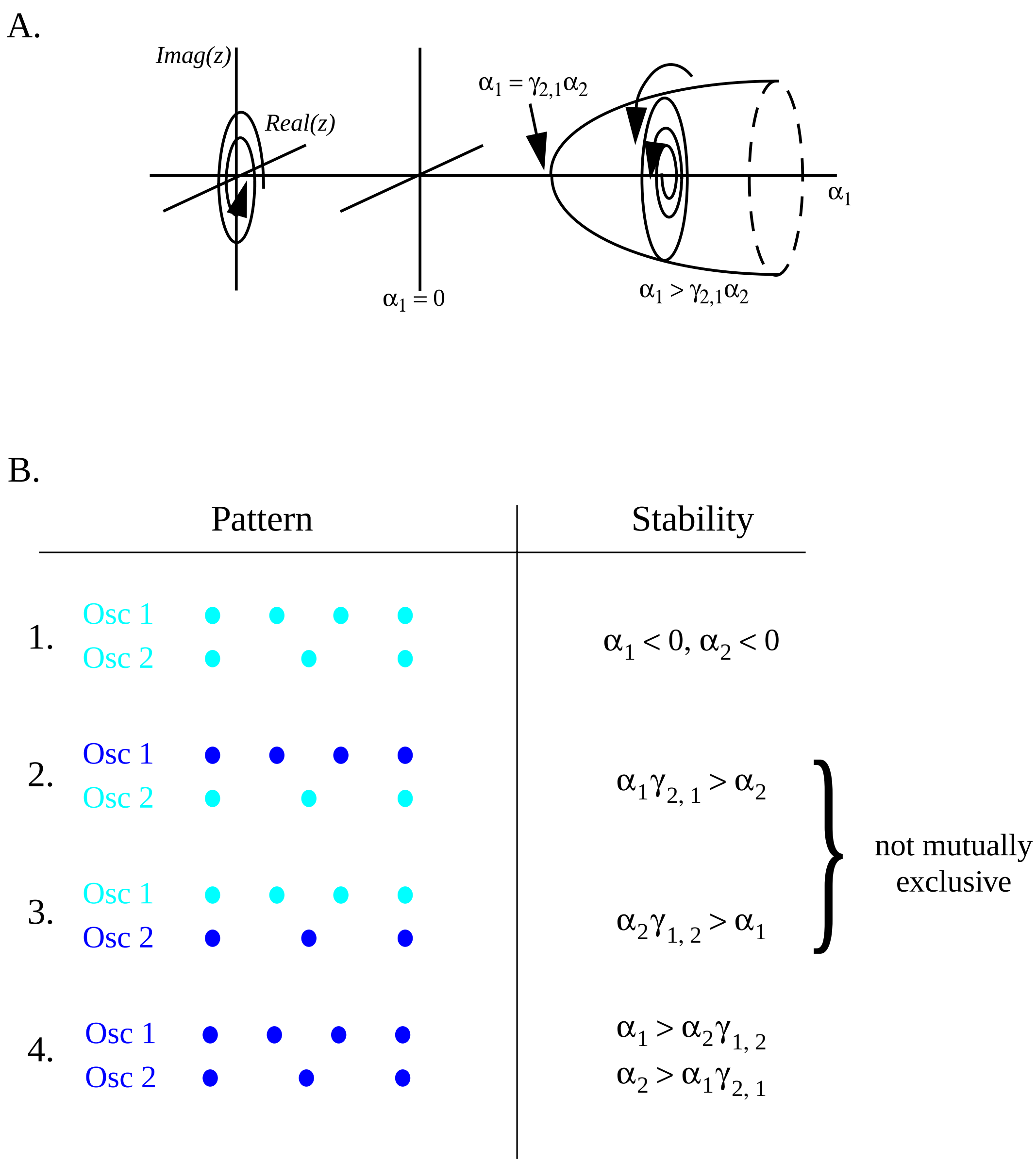
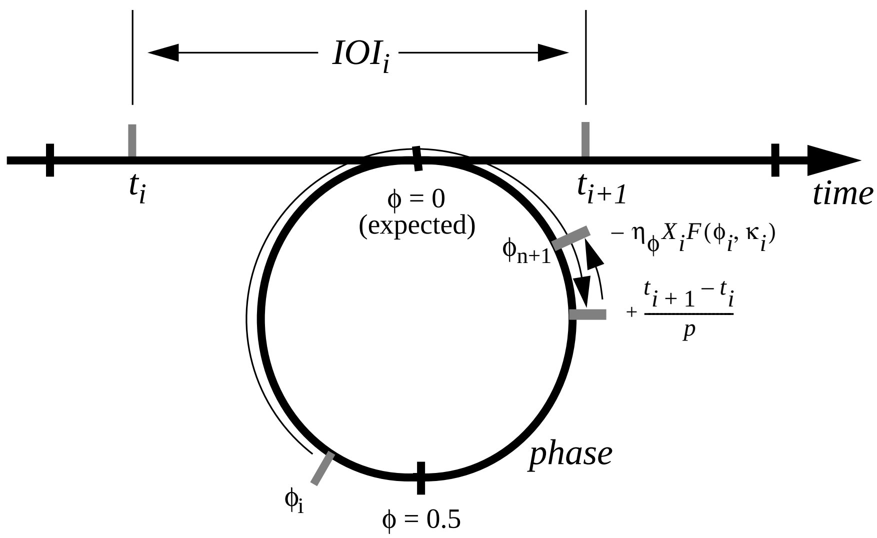

A.
Imag(z)
Real(z)
α1 = γ2,1α2
α1 = 0
α1 > γ2,1α2
α1
B.
Pattern
Stability
1.
Osc 1
Osc 2
α1 < 0, α2 < 0
2.
Osc 1
Osc 2
α1γ2, 1 > α2
3.
Osc 1
Osc 2
α2γ1, 2 > α1
not mutually exclusive
4.
Osc 1
Osc 2
α1 > α2γ1, 2
α2 > α1γ2, 1

Figure 3



IOI_i
t_i
t_i+1
time
ϕ = 0
(expected)
ϕ_n+1
− η_ϕ X_i F(ϕ_i, κ_i)
+ (t_i+1 − t_i) / p
phase
ϕ_i
ϕ = 0.5

Figure 4

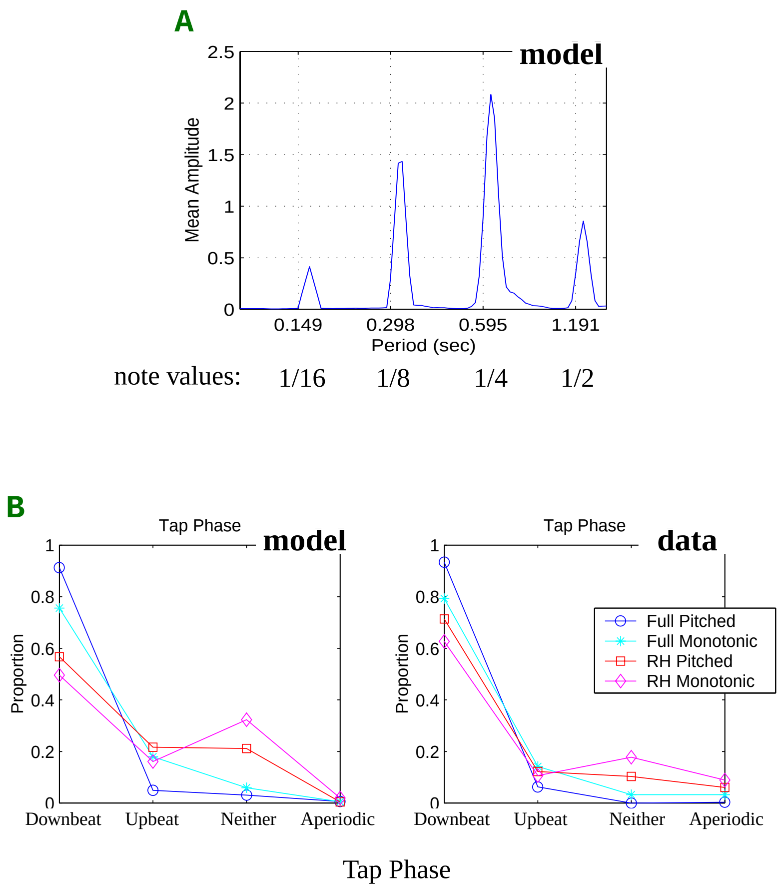

A
model
Mean Amplitude
Period (sec)
0.149
0.298
0.595
1.191
note values:
1/16
1/8
1/4
1/2
B
Tap Phase
model
data
Proportion
Downbeat
Upbeat
Neither
Aperiodic
Full Pitched
Full Monotonic
RH Pitched
RH Monotonic
Tap Phase

Figure 5

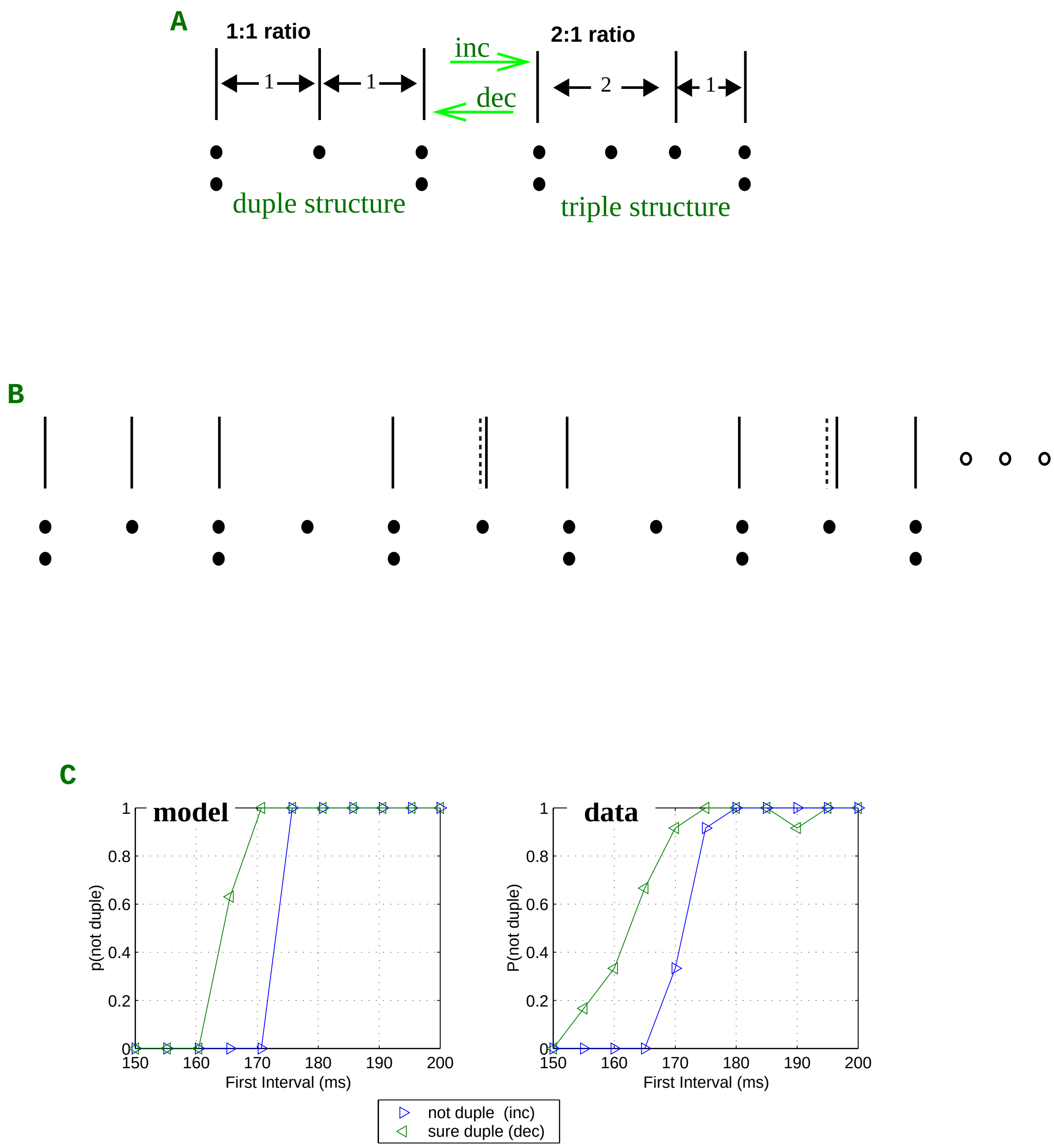
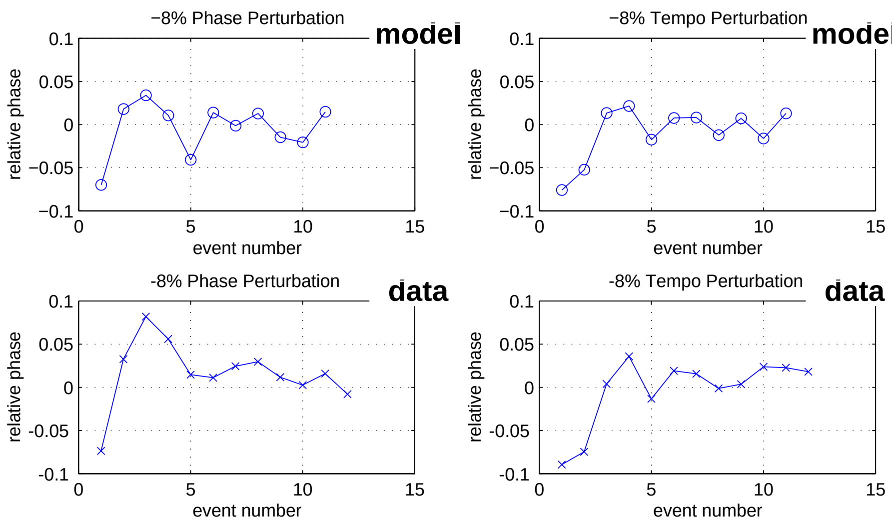
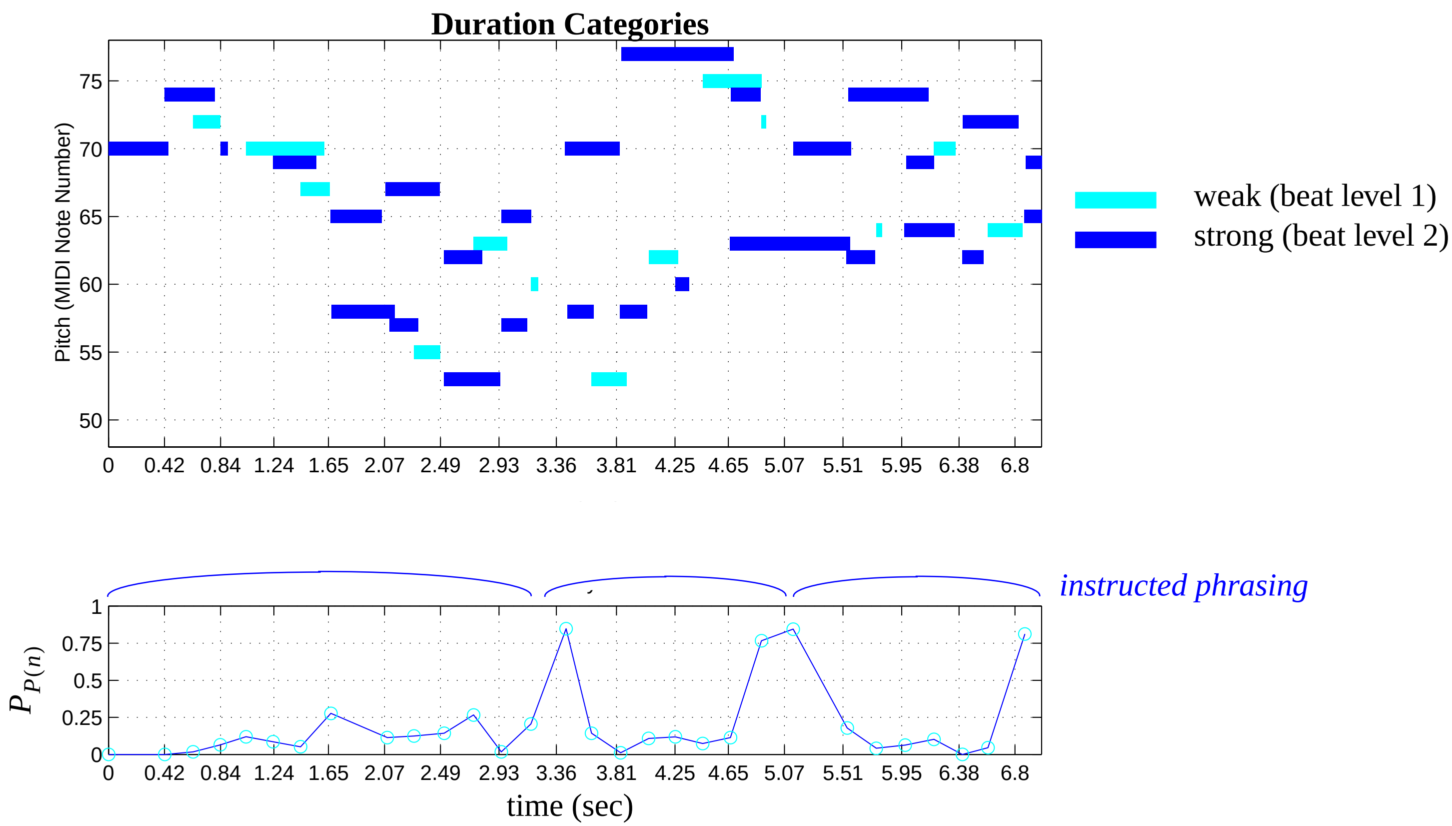

A
1:1 ratio
1
1
inc
dec
2:1 ratio
2
1
duple structure
triple structure
B
C
model
data
p(not duple)
P(not duple)
1
0.8
0.6
0.4
0.2
0
150
160
170
180
190
200
First Interval (ms)
not duple (inc)
sure duple (dec)

Figure 6



−8% Phase Perturbation
model
−8% Tempo Perturbation
model
-8% Phase Perturbation
data
-8% Tempo Perturbation
data
relative phase
event number
0.1
0.05
0
−0.05
−0.1
0
5
10
15

Figure 7



Duration Categories
Pitch (MIDI Note Number)
75
70
65
60
55
50
0
0.42
0.84
1.24
1.65
2.07
2.49
2.93
3.36
3.81
4.25
4.65
5.07
5.51
5.95
6.38
6.8
weak (beat level 1)
strong (beat level 2)
instructed phrasing
$P_{P(n)}$
1
0.75
0.5
0.25
0
time (sec)

Figure 8