

PERCEPTION OF TIME AS PHASE: TOWARD AN ADAPTIVE-OSCILLATOR
MODEL OF RHYTHMIC PATTERN PROCESSING¹

J. Devin McAuley

July, 1995

¹Thesis accepted by the Graduate Faculty of Indiana University for the degree of Doctor of Philosophy in Computer Science and Cognitive Science.

Abstract

Many human behaviors reflect the attunement of our perceptual systems to rhythmic patterns of stimulation. Examples include dancing to music, speech communication, and the performance of a symphony orchestra. However, developing a computational model of rhythm perception has proven to be difficult for two main reasons. First, rhythm is holistic, yet rhythmic patterns evolve over time. Second, periodicities in rhythmic patterns typically exhibit variability in their timing. Many previous approaches to rhythm perception have ignored these two problems by abstracting time to the level of musical notation, and thus failed to address the fundamental issue of the perception of time. The approach taken in this thesis is that the development of a model of rhythm perception must first address the perception of the time intervals which comprise rhythmic patterns.

I propose a class of *adaptive-oscillator* processing units which track periodicities in rhythmic patterns (beats). Modest random variations in the timing of rhythmic patterns do not reduce the adaptive oscillator's ability to attain synchrony, and can even improve it. An Entrainment Model of human time perception is then developed. The model is evaluated by comparing its performance on a series of tempo-discrimination simulations to data from analogous human listening experiments, investigating several rhythmic factors that influence listeners' ability to detect differences in the tempo of isochronous auditory sequences. Data obtained from the simulations agreed with the human data, providing support for the model. As an additional evaluation, two tempo-discrimination experiments were conducted to test model predictions regarding the perception of time as phase. The results of these two experiments also agreed with the model. Compared with other psychological models of time perception, the adaptive-oscillator-based Entrainment Model is the only model to provide a unified explanation for these tempo data. This thesis supports the adaptive-oscillator mechanism as a viable approach to modeling rhythm perception, addressing the holistic nature of rhythm, the problem of timing variability, and the perception of time. Furthermore, this thesis demonstrates how direct coupling of a computational system with the temporal structure of its environment is a potentially useful method for learning to interact with that environment.

Contents

1	Introduction	1
1.1	Rhythm in Music and Language	1
1.2	Rhythmic Pattern Processing	5
1.2.1	The Holism Problem	5
1.2.2	The Variability Problem	5
1.2.3	Two Models of Rhythmic Pattern Processing	6
1.2.4	Discussion	8
1.3	The Entrainment Hypothesis	9
1.4	The Role of Entrainment in Cognition	10
1.5	Rhythmic Pattern Processing via Entrainment	11
1.6	Thesis Overview	12
2	Time Psychophysics: Theory and Data	15
2.1	Chapter Overview	15
2.2	Psychological Dimensions	15
2.2.1	The Power Law	16
2.2.2	Weber's Law	17
2.3	Time as a Psychological Dimension	18
2.4	A Typology of Psychophysical Methods	19
2.5	Central Issues in Time Perception	21
2.5.1	The Nature of the Psychophysical Law	21
2.5.2	The Adequacy of Weber's Law for Time	25
2.5.3	Summary	37
2.6	Competing Theories of Time Perception	40
2.6.1	Clock-Counter Models	40
2.6.2	Quantal Models	42
2.6.3	Multiple-Look Models	42
2.6.4	Dynamic-Attending/Contrast Models	44
2.6.5	Discussion	45
3	Adaptive Oscillators	47
3.1	Overview	47

3.2	Mathematical Background	48
3.3	The Adaptive Oscillator	54
3.3.1	Phase Resetting	55
3.3.2	Phase Memory	56
3.3.3	Activation Sharpening	58
3.3.4	Period Coupling	58
3.4	The Dynamics of Adaptive Oscillation	62
3.4.1	Step 1: Phase Resetting	64
3.4.2	Step 2: Period Coupling	65
3.4.3	Step 3: Activation Function Modulation	66
3.4.4	Comparison with the Large & Kolen Oscillator	67
3.5	Evaluating the Adaptive Oscillator	69
3.5.1	Rationale	70
3.5.2	Method	72
3.5.3	Results	72
3.5.4	Discussion	74
4	The Entrainment Model	78
4.1	Model Specification	78
4.1.1	The Mapping Between External and Internal Periods	79
4.1.2	Time as Phase	79
4.1.3	The Just-Noticeable Phase Difference	82
4.1.4	Parameters	83
4.2	Model Predictions	84
4.2.1	Tempo Discrimination	85
4.2.2	Simulation 1: Duration & Number of Intervals	88
4.2.3	Simulation 2: Direction of Tempo Change	93
4.2.4	Simulation 3: Temporally-Directed Attending	96
4.3	Model Summary	99
5	Two Tempo Discrimination Experiments	103
5.1	Overview of the Listening Experiments	103
5.2	Experiment 1: Direction of Tempo Change	105
5.2.1	Rationale	105
5.2.2	Method	105
5.2.3	Results	107
5.3	Experiment 2: Dynamic Attending	112
5.3.1	Rationale	112
5.3.2	Method	112
5.3.3	Results	114
5.4	Discussion	117

6	Conclusions	121
6.1	Contributions of the Entrainment Model	121
6.2	Contributions of the Adaptive Oscillator	123
6.3	Evolving Adaptive Oscillators	125
6.4	The Role of Timing Variability	125
6.5	The Perception of Meter	126
6.6	Closing Thoughts	127

Chapter 1

Introduction

1.1 Rhythm in Music and Language

Rhythm permeates human experience, such as in listening and dancing to music, in speech communication, and in the intricate musical communication between members of a symphony orchestra. Although it is easy to agree that many human activities elicit a sense of rhythm, both in the perceiver and in the performer, it is much harder to agree on precisely what rhythm is. Confusion regarding a definition for rhythm stems mostly from distinguishing rhythm as a stimulus (e.g., the pattern of sounds played by a musician) and rhythm as a perception (e.g., the sense of periodic “beats” evoked by that pattern). In order to avoid some of this confusion, I will define rhythm in a way that combines both the “pattern” and “perception” uses of the term. For this thesis, it is also necessary to define the closely related terms of *beat*, *meter*, and *tempo*, as they are used in research on rhythm in music and language.

A *rhythm* is a temporal pattern which evokes a sense of periodicity, either in the form of periodic *beats* (pulses) or in the form of pattern repetition. Rhythmic patterns are inherently relative-time patterns in that the duration of events in that pattern are ideally determined in relation to a fundamental time unit, often called the beat duration (Jones, 1987). For example, the rhythmic pattern depicted in Figure 1.1 is based on a beat duration of t ms. This pattern repeats every $16t$ beats. *Meter* defines beats on a number of hierarchically related time scales, or levels. At each such *metrical level*, beats occur in a regular series such as the ticks of a metronome. The beat period at each metrical level (the time interval between the ticks of the metronome) is a multiple of the beat period of the metrical level below it. In the example shown in Figure 1.1, beats occur at the t , $2t$, $4t$, $8t$, and $16t$ levels.

Musical meter is used to constrain the set of possible “legal” rhythmic patterns, in a way similar (but not identical) to how formal grammars constrain the set of possible grammatical sentences (Lerdahl and Jackendoff, 1983). Each level of a metric hierarchy serves to constrain the set of rhythmic patterns that can be *best* described by that meter. For example, 2/4 time specifies a two-level metrical hierarchy in which

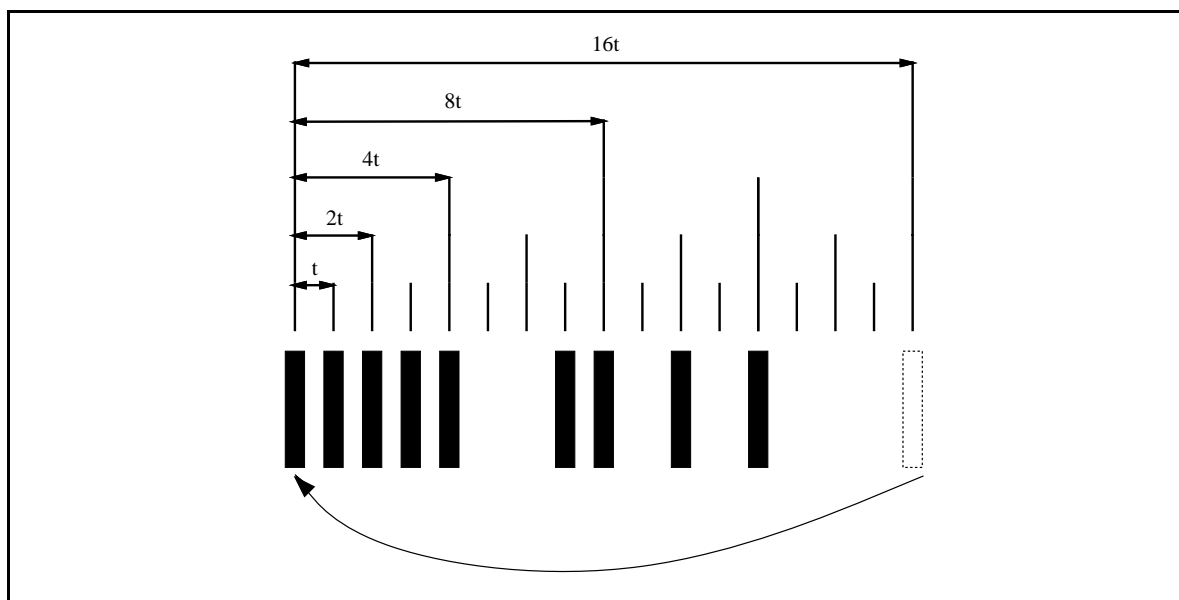


Figure 1.1: Illustration of the hierarchical structure of a simple rhythmic pattern based on multiples of two. The pattern repeats every $16t$ time units.

there are two quarter-note beats per measure. In terms of the metronome analogy, the quarter-note-level constraint on a rhythmic pattern is equivalent to having to adjust a metronome so that its ticks coincide with events in that pattern. The measure-level constraint (for 2/4 time) is equivalent to having to align the ticks of a second metronome (with a period twice the first) with events in the pattern, as well as with the ticks of the first metronome. How well both metronomes can be aligned with the rhythmic pattern determines how well it is described by that meter. For isochronous sequences, for which there is a pulse every t ms, the ticks of the second metronome can be aligned with every 2, 3, or 4 pulses, etc., and thus, the best meter is ambiguous. In contrast, the rhythmic pattern illustrated in Figure 1.1 is best described by a 2/4 or 4/4 meter (i.e., one based on multiples of 2).

Polyrhythms, common in African music, as well as some Western popular music, are not based on a metric hierarchy (Yetson, 1976). Instead, polyrhythms elicit a sense of rhythm by weaving together separate rhythmic lines, sharing only a common beat period (or micro-pulse). Polyrhythms can be described as the simultaneous presentation of two isochronous patterns that do not share a common denominator other than the micro-pulse itself (Deutsch, 1983; Handel, 1989). For example, the 3x4 polyrhythm shown in Figure 1.2 weaves together two isochronous sequences separated by 3 and 4 micro-pulses, respectively, forming a 12 micro-pulse repeating figure. Polyrhythms are problematic for theories of rhythm perception which attempt to parse rhythmic patterns strictly in terms of hierarchical relationships between time periods (Lerdahl and Jackendoff (1983); see Large (1994) for a discussion of some of the problems associated with representing polyrhythmic structure).

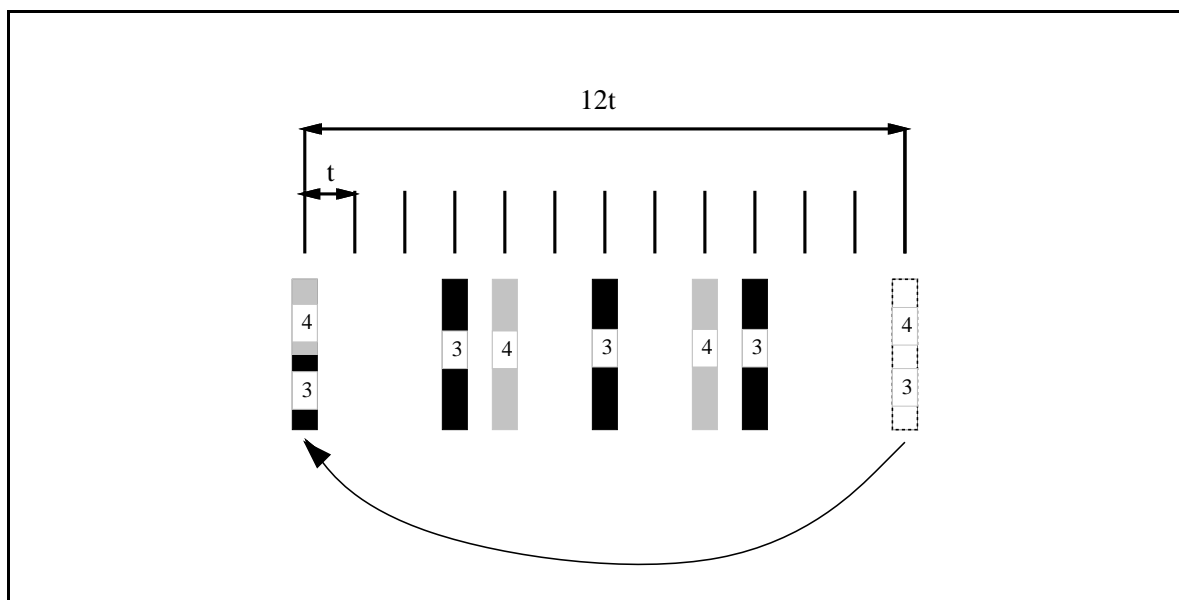


Figure 1.2: A 3x4 polyrhythm weaving together two isochronous sequences separated by 3 and 4 micro-pulses, respectively, resulting in a 12 micro-pulse repeating rhythmic figure.

Tempo refers to the rate (or speed) of a rhythmic pattern and is defined relative to the fundamental beat duration. For the example illustrated in Figure 1.1, the tempo of the pattern is determined by specifying what t is. Changing the tempo of a rhythmic pattern refers to changing t , and thus does not influence the *temporal structure* of that pattern. However, changing the tempo of a rhythmic pattern may in fact influence its “perceived” temporal structure (Gabrielsson, 1973; Handel, 1993). Thus, the same rhythmic pattern presented at various tempos may evoke different senses of rhythm.

With regard to the perception of rhythm, beats refer to a periodic series of *subjectively* equal time intervals (Fraisse, 1982; Handel, 1989). That is, even if successive beats t vary somewhat in duration (i.e., the pattern is not perfectly periodic) a listener’s perception of beats may still be *isochronous* (i.e., following at equal time intervals). The distinction between acoustic isochrony and perceived isochrony is especially evident in studies of speech rhythm. For example, for English it has been claimed that stressed syllables are separated by approximately equal time intervals (Pike, 1945). However, when the durations between stressed syllables in an utterance are measured, they are found to deviate substantially from isochrony. In spite of these deviations, listeners may still perceive that these stressed syllables are isochronous (Lehiste, 1977). Similarly, Japanese is a *mora*-timed language in that pedagogical explanations claim that all moras (a fundamental timing unit similar to the syllable) have equal duration (Bloch, 1942). However, when individual mora durations are

measured in spoken Japanese, they are not found to be equal, but can vary substantially. It is only when viewed at the word level that the mora can be salvaged as an *acoustic* isochronous timing unit, since words having the same number of moras tend to be the same duration (Port et al., 1987). At the mora level, isochronous beats are subjective.

With respect to such subjective beats in music and language, meter refers to a periodic pattern of subjectively stronger and weaker beats, such as the repeating strong-weak-weak pattern of a waltz rhythm. When we tap a foot in time with a favorite tune on the radio, we are tapping out the perceived beats of the meter at one or more metrical levels (i.e., we are tapping out some of the *metrical accents* (Lerdahl and Jackendoff, 1983)). A listener's placement of beats may coincide with the onset of events in a musical rhythm (such as with the onset of a tone), but may also occur at points in time which lack any physical marker. In addition, beats may still be "felt" after the pattern stops. If asked to "beat along" with the example shown in Figure 1.1, most listeners' would tap out beats at the $2t$ or $4t$ level, depending on the tempo.

The perception of rhythm also involves a sense of what events in a rhythmic pattern "go together." A number of stimulus factors have been shown to result in *phenomenal accents* which influence the formation of natural groupings of pattern events. In music, increasing the intensity of every other tone in an isochronous sequence of otherwise identical events tends to induce listeners to group tones in that pattern by twos, with the more intense (accented) tone beginning each group (Fraisse, 1956); following a large pitch change, tones tend to be perceived as accented and as beginning a group (Jones, 1981); and, lengthening every second or third tone of an initially isochronous sequence tends to induce listeners to group tones in the pattern by twos or threes, with the lengthened (accented) tone ending the group (Woodrow, 1951). Of course, many other stimulus factors influence the formation of groups, and no one such factor constitutes an immutable rule of rhythm perception (cf., Handel, 1989).

The pattern of these phenomenal accents influencing grouping may reinforce or conflict with the metrical accents, resulting in a more or less stable perception of rhythm. In the musical communication of rhythm, a musician's performance of a piece of music may result in phenomenal accents which coincide or conflict with the metrical accents implied by only the relative timing of the notes. Such expressive cues in performance, selectively enhance or mask the communication of rhythm to the audience. Each listener's perception of rhythm emerges through the interaction of these expressive cues provided by the performer with those structural cues provided by the sound pattern (Jones, 1987).

1.2 Rhythmic Pattern Processing

The development of a computational model of rhythm perception which can track the beat of a rhythmic pattern as well as humans can has proven to be difficult for two main reasons. First, rhythm is holistic perception, yet rhythmic patterns evolve over time. Second, periodicities in rhythmic patterns exhibit variability in their timing. In this section, I first outline these two problems, and then describe two models for extracting musical meter, one symbolic and the other connectionist, which highlight these two problems and thus serve to motivate this thesis.

1.2.1 The Holism Problem

Rhythmic pattern processing can be characterized as holistic because it involves an interaction between phenomenal accents and metrical accents that are distributed in time. Determining perceived rhythm requires examining the pattern as a whole, and not by simply combining the contributions of the individual accents, separate from the rest of the pattern. However, rhythmic patterns are not available all at once, since the pattern develops over time. Thus, a sense of a pattern's rhythm must emerge during the course of processing. Handel (1989) describes the phenomenon in the following way.

The rhythm of a game [of basketball] emerges from the rhythm of individuals, the rhythm among team members, and the rhythmic contrasts between opposing teams. In the same way, musical rhythm emerges from the lines of each instrument or instrumental section. Each line may be simple or complex, and yet, in a very real sense, the rhythm cannot be found at any one of those levels (Handel, 1989, p. 383).

The perceived pattern of phenomenal accents may reinforce or conflict with the expected pattern of metrical accents inferred by the evolving temporal structure of the rhythmic pattern. In addition, various factors influencing the perceived placement of phenomenal accents (discussed briefly in Section 1.1), may reinforce or compete with another.

1.2.2 The Variability Problem

The variability problem can be subdivided into two parts: intrinsic variability and intended (expressive) variability. I use *intrinsic* variability to refer to timing variability introduced because of human limitations in the perception and production of temporal intervals which comprise rhythmic patterns, and *intended* variability to refer to expressive changes in the timing of a rhythmic pattern that are introduced to enhance the communication of its rhythm.

With regard to the time intervals which comprise a rhythmic pattern, Fraisse (1963) identified the *psychological present* as the range of time intervals (between

approximately 25 and 2000 ms) within which we are able to maintain a “sense of pattern.” Adjacent events in a pattern that are separated by longer than about two seconds lose their cohesiveness and are perceived as isolated events. For example, consider the successive rings of a telephone which are separated by a second or so. When listening to the telephone ring, it requires effort to perceive successive rings as part of a pattern, to anticipate the next ring, or even to know whether the ringing has stopped. Events in a pattern that are separated by less than 25 ms or so, start to blend (or fuse) together into a single complex event. Thus, the perception of rhythmic patterns is necessarily constrained by these limitations on the time intervals which can form a pattern.

Within the psychological present, the perception of rhythmic patterns is influenced by human perception of the time intervals comprising those patterns. In general, the perception of time intervals within the psychological present is most accurate for intervals between approximately 500 and 600 ms (as measured by listeners’ ability to detect whether two intervals are the same or different), and progressively less accurate for time intervals outside this range (Fraisse, 1963; Woodrow, 1951). However, there is substantial variability among individuals in the location of this *preferred* range. Outside the preferred range, there is a systematic tendency for short intervals to be overestimated (i.e., perceived as having a longer duration than the actual duration) and for long intervals to be underestimated (i.e., perceived as having a shorter duration than the actual duration) (Fraisse, 1963; Woodrow, 1951). Consequently, these limitations on the perception of time intervals which comprise rhythmic patterns are sources of intrinsic variability.

With regard to intended variability, musicians create expressive patterns of phenomenal accentuation by slightly lengthening or shortening the specified duration of particular notes or rests (silences) in a pattern, by speeding up or slowing down specific portions of a pattern, etc., thus influencing listeners’ perception of that pattern’s rhythm (Clarke, 1989; Drake and Palmer, 1993). Such expressive deviations in timing may take advantage of human ability to perceive the time intervals comprising a pattern, so as to enhance the communication of its rhythm (Sternberg et al., 1982).

1.2.3 Two Models of Rhythmic Pattern Processing

Below, two complementary models of rhythmic pattern processing are described. It is not the purpose of this introductory discussion to provide a comprehensive review. Instead, these two models provide the reader with concrete examples of typical approaches to rhythm perception and serve to illustrate how such approaches address (or fail to address) the issues of holism and timing variability.

The Clock Model Povel and Essens (1985) propose a rule-based model of rhythmic pattern processing which they term the “Clock Model”. In the Clock Model, rhythmic patterns are represented as a sequence of time intervals measured relative

to a defined micro-pulse. For example, if we define a 200-ms micro-pulse, then the pattern of intervals 200 400 200 200 800 can be represented as 12114. Events in these patterns occur at the beginning of each time interval and are assumed to be identical. In the Clock Model, processing takes place in three steps. In the first step, phenomenal accents are determined according three preference rules suggested by Povel and Okkerman (1981) as follows: (1) isolated tones tend to be accented; (2) the second in a cluster of two tones tends to be accented; and (3) the initial and final tones in a cluster of three or more tones tend to be accented. In the second step of processing, all possible “clocks” (or metronomes) are generated up to a clock period equal to half the cycle duration of the rhythmic pattern. For example, for a rhythmic figure that repeats every 2 seconds, the slowest clock would have a 1-sec period. The concept of a clock is identical to the concept of the metronome introduced in Section 1.1. Each “clock” corresponds to a periodic sequence of ticks that can be aligned in different ways with the events in a rhythmic pattern. Thus, the generation of all possible clocks also includes the generation of all possible alignments of each clock. In the final step of processing, a rule is applied to each clock which scores the amount of dissonance between that clock and the natural patterning of accents determined in the first step. The best clock (i.e., the one with the least dissonance) indicates the model’s “perceived” meter.

One strength of the Clock Model is that it is a potentially useful tool for analyzing the structure of rhythmic patterns. In the Povel and Essens (1985) study, they used the range of possible clock dissonances to define a measure of *rhythmic complexity*. The dissonance of the best clock for a temporal pattern thus provided a measure of the rhythmic complexity of that pattern. Povel and Essens (1985) found that listeners’ ability to reproduce temporal patterns was correlated with this measure of rhythmic complexity; that is, the more rhythmically complex a pattern was determined to be, the less accurate listeners generally were at tapping this pattern, and the longer it took listeners to memorize the pattern prior to tapping. This result supports the hypothesis that listeners’ coding of temporal patterns is rhythm-based, and suggests that the Povel and Essens (1985) measure of rhythmic complexity is, to a first approximation, reasonable.

A main weakness of the Clock Model is that it determines the best clock by inspecting a static “snapshot” of the entire pattern. Thus, the holism problem is only addressed by ignoring the fact that rhythmic patterns evolve over time. Those rule-based approaches which do, to a degree, attempt to incorporate the intrinsic temporal constraint on rhythmic pattern processing (Longuet-Higgins and Lee, 1982; Miller et al., 1988) use hindsight to backtrack and reanalyze earlier portions of the pattern, in order to successfully extract the meter. That is, although these models process rhythmic patterns sequentially, the “left-to-rightness” of time is disregarded. A second weakness of the Clock Model is that by using a beat-based input representation, in essence a musical notation, the model only applies to perfectly periodic rhythmic patterns, and thus does not address the fundamental problem of timing

variability, inherent in music and speech.

The BeatNet Model In contrast to the rule-based Clock Model, Scarborough, Miller, and Jones (1990) propose a connectionist model of rhythmic pattern processing called BeatNet. The BeatNet Model consists of a one-dimensional array of oscillators, with periods defined by multiples of the shortest duration in the to-be-processed rhythmic pattern, and initial phases defined by the different possible alignments of the oscillators to that pattern. Thus, there are oscillators corresponding to periodic patterns of eighth-notes, quarter-notes, etc., for the different possible downbeats. This is the same as the generation of all possible “clocks” in the Clock Model except that now each clock is realized as an oscillatory processing unit. The BeatNet Model performs essentially the same task as the Clock Model, except that instead of using preference rules to determine meter, it uses a form of parallel constraint satisfaction. As the BeatNet Model makes a left-to-right pass through the score (without backtracking), it excites the various oscillators which have the same phase and period as the time intervals in the score. The oscillators are allowed to interact via excitatory and inhibitory connections in ways which implement constraints that are likely to lead to a pattern of activation corresponding to a coherent metrical structure.

To a limited degree, the BeatNet Model is successful at extracting coherent metrical structure without the use of explicit rules, without requiring that the whole pattern be available for inspection, and without the need for backtracking. In addition, since the oscillators only communicate with each other through local interactions, there is no executor which directs the processing of the system. However, in order for the system to successfully extract metrical structure, it needs to pre-inspect the to-be-processed rhythmic pattern in order to determine what the correct initial periods and phases of the oscillators should be. Furthermore, if the rhythmic pattern increased or decreased in tempo, the coherent metrical structure established by the activation pattern of these oscillators would be lost. Thus, one of the main weakness of the BeatNet Model is its inability to deal with timing variability. Like the Clock Model, the BeatNet Model essentially assumes musical notation as input, and thus breaks down when applied to rhythmic patterns which exhibit timing variability, as found with music and speech.

1.2.4 Discussion

Typical approaches to Artificial Intelligence (AI) problems, such as music and language processing, often factor out all aspects of perceptual and motor skills, reducing the problem to one of “identifying the right representation” (Brooks, 1991). Brooks (1991) observes that AI researchers tend to partition problems into two components: an AI component, which they attempt to solve, and a non-AI component (including the perceptual and motor components of the problem), which they don’t solve, but factor out using a process of abstraction. AI approaches to rhythmic pattern

processing are no exception to this practice.

In the symbolic Clock Model and the connectionist BeatNet Model, as well as a number of other related models (e.g., those of Longuet-Higgins and Lee (1982); Miller et al., 1988; and Stevens and Wiles, 1994), it is the perception of the time intervals which comprise rhythmic patterns that is abstracted to the level of musical notation, and thus implicitly designated the non-AI component of the problem. In some cases, the whole pattern-presentation process is also ignored. Consequently, in handling rhythm perception, these models do not address the fundamental issue of the perception of time, instead assuming that the time intervals between events in a pattern are obtained by an unspecified pre-processor (Port et al., 1995). Of course, for these cases in which identifying the right representation implies using musical notation as input, it must also be assumed that rhythmic patterns are perfectly periodic; as we’ve seen, this is not the case for music and speech, which can exhibit significant intrinsic and intended variability in the timing of their production. In addition, this type of approach assumes that the problem of determining the *relative timing* between events is already solved—a nontrivial part of the task itself (Port et al., 1995).

1.3 The Entrainment Hypothesis

An alternative approach to rhythmic pattern processing is based on the concept of a direct coupling between perceiver and environment, through which stimulus rhythms serve to entrain (or synchronize) a perceiver’s internal rhythms. In a seminal paper, Jones (1976) proposed that the temporal organization of perception, attention, and memory is inherently rhythmic. As part of this theory, it is assumed that rhythmic patterns such as music and speech potentially entrain (synchronize) a hierarchically nested set of attentional periodicities (oscillations), forming an attentional rhythm. In Jones’s view, the entrainment of attentional periodicities is the basis for the development of expectancies for *when* in time, events in a pattern are likely to occur. Moreover, according to this view, the entrainment of attentional rhythms guides the placement of attentional pulses (corresponding to points in time receiving attentional focus), thus influencing the overall perception of a stimulus and the form of its storage in memory, depending on the temporal structure of the stimulus context.

There is an accumulating body of psychological evidence that supports this *entrainment hypothesis*, demonstrating that listeners’ abilities to detect pitch changes, estimate time intervals, remember a pattern, and discriminate tempo is heightened when the to-be-detected change in the pattern, or the set of pattern defining features, coincide with perceived metrical accents in that pattern (Jones et al., 1981; Jones et al., 1982; Jones and Boltz, 1989; Kidd et al., 1984; Kidd, 1989; Kidd, 1993; Kidd, 1994). Jones et al. (1982) found that listeners were better able to detect pitch changes in a target tone (based on a melodic rule change) when the target tone occurred at a metrically accented location than when it occurred at a unaccented location. Similarly, in a task which required listeners to judge the melodic equivalence

of pairs of auditory patterns, Kidd et al. (1984) found that temporal uncertainty due to differences in rhythmic context (i.e., number and types of rhythmic patterns used in the experiment) reduced listeners' ability to discriminate melodic differences. In addition, Kidd (1993; 1994) found that listeners' ability to resolve the frequency of a target tone embedded within an auditory pattern depended on the magnitude of the temporal displacement of that tone from its "expected" location. In a study of time perception, Drake and Botte (1993) showed that listeners' ability to detect temporal deviations in a target interval or a set of target intervals comprising an auditory pattern is better when that pattern is metrically regular than when it is irregular. Finally, in a study of speech perception, Kidd (1989) found that listeners' judgments concerning the identity of a target syllable were systematically influenced by the temporal expectancies established by the tempo of a precursor phrase. The results from these diverse studies of auditory pattern perception provide broad support for the entrainment hypothesis.

1.4 The Role of Entrainment in Cognition

The ability of our perceptual systems to be entrained by rhythmic patterns, as suggested by Jones (1976), can be important for cognitive processing. First, entrainment is a form of pattern organization in time. Perceiving metrical structure through entrainment establishes the relative salience of the events comprising a rhythmic pattern. This is evident in speech processing, in which metrical accents (stress) often point to important content words in the spoken utterance. Thus, events (e.g., words) that coincide with subjectively stronger accents may be more important for the processing of a temporal pattern (e.g., understanding the intended meaning of a sentence) than those which coincide with weaker accents (Martin, 1972; Handel, 1989). Moreover, entrainment as a form of pattern organization offers a reduced memory representation for temporal patterns that preserves those features which coincide with metrical accents (Jones, 1976; Large et al., 1995; Martin, 1972; Povel, 1981).

Entrainment is also a basis for prediction in time. Entrainment by a rhythmic pattern enables the timing of events in that pattern to be predicted based on extension of the entrained periodicities. Thus, the timing of early events generates expectancies regarding the periodic occurrence of later events (Jones, 1976; Martin, 1972). Those events that coincide with their predicted locations in time serve to reinforce the entrainment-based predictions. Accurate prediction of the timing of an event is important because it facilitates the coordination of actions that must coincide with that event. Also, accurate timing prediction may also reduce the quantity of information that must be processed in order to identify or classify a pattern. For example, if the melody of a musical pattern can be identified from only those portions of the acoustic signal which coincide with beats of the meter, then the remainder of the acoustic signal can be discarded at an early stage in processing. Thus, entrainment as a form of prediction in time may help establish *what* information in a pattern should

be processed as well as *when* that processing should occur. Furthermore, entrainment-based timing prediction is independent of tempo—a necessary property for perceptual constancy. For example, for successful speech communication, it is critical that the perceived content of a speech utterance does not change with speaking rate.

1.5 Rhythmic Pattern Processing via Entrainment

There has been a recent surge of interest in the role of synchronized cortical oscillations in cognition. However, most of this interest has focused on how synchronized oscillations might be used for feature binding in visual processing (Gray et al., 1989), or for modeling processes related to auditory stream segregation (von der Malsberg and Schneider, 1986; Wang, 1995). Much less research has investigated the possible role of entrained cortical oscillations in the processing of rhythmic patterns, although there is significant physiological data to support such a possibility (John, 1967; John, 1972; Thatcher and John, 1977; Torras, 1985; Torras, 1986).

John (1967) describes a series of animal classical conditioning experiments resulting in cortical entrainment by rhythmic stimuli. In the avoidance paradigm, a flashing light followed by a brief electric shock is used to condition an animal to expect electric shock when the flashing light occurs. When the conditioned response is obtained (i.e., the animal moves to avoid the shock once it sees the flashing light), it is reported that cortical activity is observed at the frequency of the flashing light. Similarly, if an animal receives a reward for pressing a lever in response to a flashing light at a particular rate (e.g., a cat pressing a lever to receive milk in response to a 10 Hertz flashing light) and punishment for pressing the same lever in response to a flashing light at a different rate (e.g., no milk and a substantial wait until the next milk opportunity for a 6 Hertz frequency), then sustained cortical activity is observed at both frequencies and has been used to predict whether the animal will make a correct discrimination. Moreover, in these rhythmic conditioning studies, when the presentation of the conditioned rhythmic stimulus stops, sustained cortical activity continues to be observed at the conditioned frequency. Also, if pulses of the rhythmic stimulus are left out, such as in deleting light flashes, then the entrained neurons “fill in the missing beats” by firing when the deleted pulses would have occurred. With regard to these data, John (1967) observes that “it seems highly probable that such rhythms are not idiosyncratic to a particular conditioning situation nor to a particular species but represent a general capacity of the central nervous system to reflect the temporal pattern of prior stimulation for a period of time following the cessation of that external event (p., 296).”

Based on these and other data, Torras (1985) proposed that the entrainment of cortical activity is due to changes in intrinsic neuronal “firing” rate induced by the phase-resetting of rhythmic input. She then developed a detailed integrate-and-fire model of how oscillatory “pacemaker” neurons in certain invertebrates are entrained by rhythmic stimulation.

It was Torras’s idea of an oscillatory neuron with a plastic firing rate that inspired this thesis, as it seemed very applicable to the problem of modeling human rhythm perception (McAuley, 1993). Additional support for such an *adaptive oscillator* approach to rhythmic pattern processing comes from Baird, Troyer and Eeckman (1994a; 1994b) who have suggested links between the synchronization of cortical oscillations and entrainment theories of attention (Jones, 1976; Jones and Boltz, 1989; Jones and Yee, 1993). The concept of an adaptive oscillator (McAuley, 1993; McAuley, 1994a) lies between that of a single-neuron model and that of a psychological theory, providing a functional instantiation of Jones’s notion of an attentional periodicity and in that sense implements the entrainment hypothesis. However, I do *not* intend to suggest that an implementation of the entrainment hypothesis occurs at the single-neuron level or even that adaptive oscillators might correspond to single neurons. Instead, adaptive oscillators are intended to model the global dynamics of perceptual mechanisms involved in the processing of rhythmic patterns. The adaptive oscillator can be thought of as internalizing an expectancy for the occurrence of future inputs. In the case of “missing” inputs, as well as the case when the rhythmic pattern stops, the adaptive oscillator continues to predict future inputs. In essence, the adaptive oscillator internalizes a beat, retaining a memory of that beat after the pattern stops. In terms of the oscillator retaining a memory of “what the beat was”, the gradual return of the oscillator’s period to its resting value corresponds to memory decay. Several recent papers have proposed such adaptive-oscillator models for the perception of musical meter (Large, 1994; Large and Kolen, 1994; McAuley, 1994a).

1.6 Thesis Overview

The approach taken in this thesis is that the development of a successful computational model of human rhythm perception must first consider the perception of the *time intervals* which comprise rhythmic patterns. Therefore, modeling human perception of time is a necessary step towards the development of a comprehensive model of rhythmic pattern processing. With regard to Brooks’s discussion of the standard methodology in Artificial Intelligence research, my view is that the “non-AI” part of the rhythm perception problem (i.e., the time perception problem) *is* the problem. Thus, towards solving the rhythm perception problem, I propose a class of *adaptive oscillator* processing units which track periodicities in rhythmic patterns (beats), in spite of variability in the timing of those patterns. An Entrainment model of human time perception is then developed based on the proposed adaptive oscillator.

Chapter 2 reviews psychological data regarding listeners’ perception of time intervals which comprise rhythmic patterns, focusing on two central issues. First, what is the relationship between “clock” time and perceived time? Second, how do various stimulus factors influence listeners’ ability to detect timing changes. With regard to this second issue, the review focuses on listeners’ sensitivity to differences in the duration of isolated time intervals (bounded by two tones) and on listeners’ sensitivity to

differences in the tempo (rate) of isochronous tone sequences. The chapter concludes with a review of selected models of human time perception, focusing on those models which address tempo discrimination.

Chapter 3 begins with a brief introduction to the theory of coupled oscillators, including definitions for mathematical terminology used throughout this thesis. This introduction is followed by a detailed specification of a class of adaptive oscillator processing units which share five properties. The proposed oscillators extend prior research which formed the basis for this thesis (McAuley, 1993; McAuley, 1994a). The entrainment dynamics of the proposed class of oscillators are then examined by constructing Arnold maps (Arnold, 1983) for different parameter settings of the model. The dynamics of the proposed model are then compared with those of a similar oscillator model proposed by Large (1994) for the perception of musical meter. The chapter concludes with an evaluation of the proposed model to be entrained by temporal patterns which vary in rhythmic complexity on a scale correlated with listeners' ability to memorize and reproduce those patterns.

In Chapter 4, an Entrainment Model of human time perception is developed based on the adaptive oscillator proposed in Chapter 3. The model is then evaluated by comparing its performance in a series of three tempo-discrimination simulations to data from analogous human listening experiments. The first simulation investigates stimulus factors which influence the model's ability to detect differences in the tempo of isochronous tone sequences. The tempo data obtained in this simulation are then compared directly with tempo data from the listening experiments discussed in Chapter 2. The second simulation investigates the model's predictions regarding differential sensitivity to increases and decreases in tempo. Finally, the third simulation investigates the model's predictions concerning the influence of temporally-directed attending on the detection of tempo differences. The chapter concludes with a summary of the model's predictions derived from the three simulations.

Chapter 5 reports the results from two original human listening experiments designed to test predictions of the Entrainment Model derived from simulations 2 and 3, for which no human data was available for comparison. The first experiment investigates listeners' sensitivity to tempo differences for one- and three-interval isochronous sequences for four different base tempos. In order to be able to compare the data obtained in this experiment with the predictions of the model from Simulation 2, increases and decreases in tempo are not conflated. The second experiment, motivated by Simulation 3, investigates the effect of deviations from listeners' timing expectations for the onset of a comparison sequence on their ability to detect differences in the tempo of that comparison sequence. The chapter concludes by discussing the implications of these tempo data for the Entrainment Model, as well as for the competing models described in Chapter 2. Several additional listening experiments are then suggested which would tease apart several unresolved issues.

For those readers that are more interested in the adaptive-oscillator mechanism than the Entrainment Model of human time perception, it is safe to skip Chapter 2

and begin with Chapter 3, without loss of coherence. However, there is one caveat. It will be necessary to return to Chapter 2, before proceeding with Chapters 4 and 5, as Chapter 2 provides important background for comprehension of the modeling issues discussed in these chapters. For those readers without a background in time psychophysics, Chapter 2 will be particularly useful as an introduction to the issues addressed in this thesis.

Chapter 6 summarizes the main contributions of this thesis, and suggests directions for future research. Overall, this thesis will support the proposed adaptive oscillator as a viable approach to the problem of modeling human rhythm perception, addressing the holistic nature of rhythm perception, the problem of timing variability, and the perception of the time intervals which comprise rhythmic patterns. Furthermore, this thesis will demonstrate how direct coupling of a computational system with the temporal structure of its environment is a potentially useful method for learning to interact with that environment.

Chapter 2

Time Psychophysics: Theory and Data

2.1 Chapter Overview

The purpose of this chapter is to review time perception data and theories relevant to this thesis, providing the background and motivation for the modeling and experimental work discussed in Chapter's 4 and 5. Stevens (1975) makes an important distinction between quantitative and qualitative sensory continuum, which will be discussed in Section 2.2. With regard to this distinction, an obvious question for someone interested in time perception is what type of sensory continuum is time. This question is addressed in Section 2.3. Section 2.4 describes the different psychophysical methods used in time perception studies which are necessary for understanding the experimental data presented in the remainder of the chapter. Section 2.5 investigates two central issues in time psychophysics pertinent to this thesis: (1) the nature of the psychophysical law for time and (2) the constancy of the Weber fraction for time discrimination. Finally, Section 2.6 reviews relevant psychological models of time perception, including several different proposed time mechanisms.

2.2 Psychological Dimensions

Everyday, we are bombarded with an enormous variety of different sensations. In effect, sensations come in a variety of shapes, sizes, and colors. We experience many different continua: color, heat, pitch, loudness, heaviness, brightness, etc. Stevens (1975) distinguishes between two types of perceptual continua: *prothetic* and *metathetic*. The prothetic continuum applies to sensations that can be described on a quantitative scale according to their degree or magnitude. Loudness is a good example of a prothetic variable. We describe the loudness of a sound by its magnitude (or quantity). Prothetic variables are also additive. Thus, two soft sounds added together produce a louder sound. Pitch, on the other hand, is a metathetic variable.

Metathetic variables are described on a qualitative scale and are better characterized in terms of position rather than magnitude. Thus, we describe pitch as being low or high. Sound frequencies above approximately 20,000 Hertz or below approximately 20 Hertz, fall off the low and high ends of the human pitch scale and are not perceived by most listeners (Handel, 1989; Moore, 1989). To specific pitch positions in between an intermediate range of these frequencies, we can even assign labels (e.g., the notes of a diatonic scale). Pitch is not additive, like loudness. Adding two pitches together produces a chord, not a higher pitch. Questions we ask about quantitative prothetic continua are concerned with *how much* (e.g., how much brightness, how much loudness, how much electric shock, etc). In contrast, questions we ask about the qualitative metathetic continua are concerned with *what* (e.g., what pitch or what color).

Psychophysics (Fechner, 1966) is a branch of experimental psychology that studies the relationship between the physical properties of a stimulus and the mental properties of the evoked sensation. Three primary problems in psychophysics are: (1) determining the *absolute-limen*, or detection threshold for a stimulus (e.g., the smallest detectable sound intensity); (2) determining the functional form of the *psychophysical law* relating stimulus magnitude to subjective magnitude (e.g., relating the sound intensity of a tone to the perceived loudness of the tone); and (3) determining the *just-noticeable difference* (JND), the sensitivity of an observer to an increment in the stimulus (e.g., the smallest change in sound intensity that a listener is able to reliably detect) (Watson, 1973; Stevens, 1975).

2.2.1 The Power Law

For prothetic continua, the psychophysical law governing the relationship between stimulus magnitude and psychological magnitude is generally accepted to be a power law (although there is still considerable debate) (Stevens, 1975); that is, psychological magnitude Ψ grows as a power function of the stimulus magnitude Φ :

$$\Psi = k\Phi^x. \quad (1)$$

In the formula, the value of the constant k depends on the units of measurement and that of the exponent x differs from one sensory continuum to another. The value of the exponent can be thought of as characterizing a particular sensory continuum. Figure 2.1 shows the power-law relationship for loudness ($x = 0.67$), brightness ($x = 0.33$), and electric shock ($x = 3.5$) (Stevens, 1975). For loudness and brightness, psychological magnitude grows less and less rapidly with increasing stimulus intensity. In contrast, the sensation of electric shock grows more and more rapidly as the electric current is increased. An important property of power functions is that a constant ratio of stimulus magnitudes corresponds to a constant ratio of perceptual magnitudes. The constant ratio property is advantageous for perceptual stability.

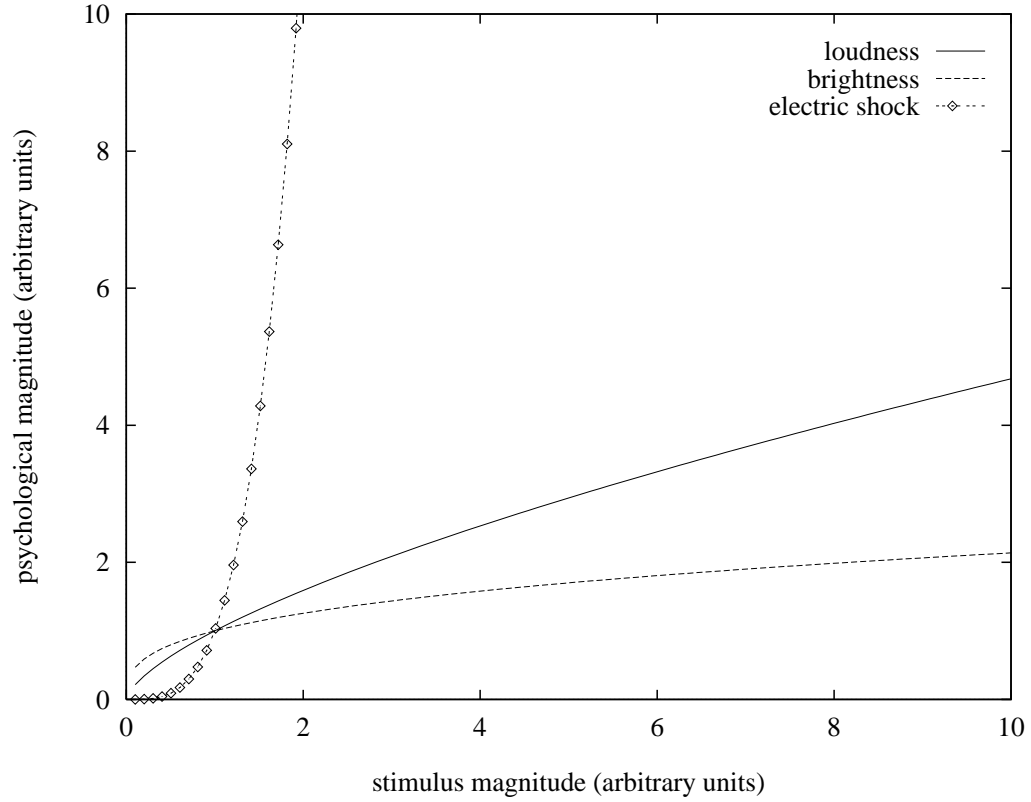


Figure 2.1: Psychophysical power law ($\Psi = k\Phi^x$) describing relationship between stimulus magnitude and psychological magnitude for loudness ($x = 0.67$), brightness ($x = 0.33$), and electric shock ($x = 3.5$).

The relative proportions of a prey being tracked by a predator should not be distorted as the predator moves in for the kill. The perceived relation among speech sounds should remain the same whether the speech is loud or soft. Power functions applied to metathetic continua have been somewhat less successful (e.g., for pitch, a power function only adequately describes an intermediate range of the audible sound frequencies (Moore, 1989)).

2.2.2 Weber's Law

An important issue in psychophysics is the relationship between the just-noticeable difference (JND) in sensation ($\Delta\Phi$) and the absolute magnitude of the stimulus (Φ), where JND is the stimulus difference, along the tested dimension, yielding 70%-correct performance of subjects detecting that difference. For many sensory dimensions, the relationship between $\Delta\Phi$ and Φ has been found to be a constant ratio (Weber's Law)

$$W = \frac{\Delta\Phi}{\Phi}. \quad (2)$$

Thus, data are often reported in terms of the Weber fraction (W) or in terms of the relative JND (where W is represented as a percentage). If Weber's Law holds, then relative JND is constant. Data supporting Weber's law are often reported as being described by the linear relationship

$$\Delta\Phi = W\Phi + \Phi_0. \quad (3)$$

where the y-intercept Φ_0 is zero. In this case, the Weber fraction W is the slope of the line.

2.3 Time as a Psychological Dimension

Is time a prothetic or metathetic variable? The answer to this question is not straightforward, partly because it is not obvious what it means for time to be a stimulus. Time can only be an indirect stimulus in that time does not stand by itself as a physical variable. That is, time can not be "detected" using a sensor, such as occurs with the brightness of light or the intensity of sound. To be perceived, time must be somehow tied to an event or series of events: tones have duration, as does the silent gap between two tones and the beat duration specifying the tempo of a series of events. When listening to a familiar tune on the radio, it is possible to make a number of different temporal judgments: (1) we can specify the order of the notes in the melody, measuring time in discrete steps; (2) we can estimate the time-interval between any two note onsets; (3) we can judge the duration of any note; and, (4) we can judge the tempo of the melody (i.e., how fast or slow it is) according to an estimate of the beat duration abstracted from the temporal pattern of all the notes. Notice that for all of these examples, the physical stimulus of the melody is the same. What distinguishes each type of time percept are the choice of the attended time markers (beginning and ending) and the attended time scale (ordinal or interval (Stevens, 1951)). For (1), time is measured on an ordinal scale (i.e., only the before-after ordering of events is important). In contrast, I propose that for (2), (3), and (4) time is measured on an interval scale. In (2), the onsets of the before-tone and the after-tone delineate the time interval. In (3), the onset and offset of a single tone specifies duration. In (4), multiple tone-to-tone onsets delineate a pattern of intervals which can be used to provide an abstract interval-based estimate of tempo: "the beat". Precisely how this beat is abstracted is pertinent to this thesis.

In this thesis, the measurement of perceived time is restricted to that of *empty intervals* (i.e., the time between the onsets of two adjacent tones) and isochronous patterns of empty intervals which convey tempo (fast or slow). For purposes of comparison, the perception of filled intervals (tone durations) will also be discussed, although only briefly. The range of time-intervals investigated will be between 25 ms and 3000 ms. Many temporal properties of importance to speech and music perception are conveyed by filled and empty time-intervals within this range (e.g., voice onset

time, syllable duration, speaking rate, and musical beat, meter, and tempo). This time scale encompasses what Fraisse (1963) refers to as the *psychological present*. The psychological present (or cognitive time scale according to Port et al., 1995) is the time frame within which successive events can be perceived as components of a single pattern. Events that are separated by longer than a couple seconds lose their cohesiveness and sound like isolated events (e.g., rings of the telephone). Events separated by less than about 25 ms, tend to blend (or fuse) together into a single complex sound.

It seems relatively straightforward to assume that the time perception of filled-intervals is a prothetic continuum. The duration of a tone can be thought of as having a physical magnitude since there is a correlation between the physical energy of a tone and its perceived duration (Moore, 1989). Moreover, the perception of tone duration is additive since concatenating two short tones of the same frequency produces a single longer tone. I would also propose that the time perception of empty intervals is a prothetic continuum, although this claim is a somewhat more difficult to support since it is impossible to attribute a physical magnitude to silence. Nonetheless, it seems reasonable to think about the duration of an empty interval (since it can be measured with a clock) and conclude that the perception of duration with empty intervals is additive. Combining two short empty intervals produces a longer empty interval, the magnitude of which can be “heard” when bounded by two tones and measured with a clock. By extension, these two arguments of filled and empty intervals perception as prothetic would seem to imply that tempo is also a prothetic variable, since tempo perception is at least partially based on the abstraction of a beat duration from a pattern of temporal intervals. However, it requires some imagination to think of tempo as a magnitude since “combining” two fast or slow sequences does not produce a faster or slower sequence. Thus, the prothetic distinction for tempo does not seem to hold true.

2.4 A Typology of Psychophysical Methods

There have been a number of past attempts to organize the various methodologies used in time perception research (Allan, 1979; Bindra and Waksburg, 1956; Woodrow, 1951). There are two basic research methodologies used: time scaling and time discrimination. The main property that distinguishes time scaling from time discrimination is the confusability of the stimulus set. In time-scaling experiments, time intervals in the stimulus set are obviously different; thus, time-scaling experiments do not address questions concerned with the just-noticeable difference for time. Instead, time-scaling experiments are used to obtain absolute measurements of subjective duration (perceived time). In time-discrimination experiments, the time intervals in the stimulus set are highly confusable; i.e., the stimuli are very similar and the listener may not be able to detect a difference between two stimuli. Time-discrimination experiments have been used to measure both JNDs and subjective duration.

There are at least five time-scaling tasks which have been used successfully: magnitude estimation, category rating, production, ratio setting, and synchronization. In a *magnitude-estimation* task, the listener provides duration estimates by assigning a magnitude (number) to each time interval. The scale of magnitudes is either arbitrary, in which case the experimenter may or may not provide an initial reference value (e.g., “base all of your judgments of duration relative to this interval x , which is assigned magnitude y ”), or is based on clock time. In a *category-rating* task, the listener assigns each time-interval stimulus to one of a set of n predefined categories, which are ordered by magnitude. In a *production* experiment, the listener receives verbal or written instructions specifying what time interval to produce, which might involve tapping a finger or foot, or adjusting the interval between two tones, via a knob. In a *ratio-setting* experiment, the listener “hears” specific time intervals and is asked to generate a specific proportion of that interval. Reproduction of the interval is specified by a proportion of 1.0. Specifying proportions less than or greater than one is called fractionization or multiplication, respectively. As with production, a ratio-setting task may involve any of a number of different methods for generating the specific proportion of the time-interval. In a *synchronization* task, either the listener hears a single time interval and attempts to respond in synchrony with its termination (i.e., predict the end of the interval) or a series of intervals and attempts to respond in synchrony with the onset of each time-interval (i.e., predict the next tone).

Types of duration-discrimination tasks include the method of comparison and the single-stimulus method. There are many different variants of the method of comparison. In general, two temporal patterns are sequentially presented on each trial. The first pattern, usually constant throughout a block of trials, is referred to as the *standard*, while the second pattern, which normally varies from trial to trial, is called the *comparison*. In a same-different task, the listener judges whether the standard and comparison are the same or different according to a criterion, such as differing only in duration, tempo, etc. In the roving-standard version of the same-different task, the standard varies from trial to trial. In a two-alternative forced-choice task (2AFC), the listener selects either the standard or comparison pattern according to a particular criterion (e.g., “which is faster”). In a standard, two-alternative forced-choice task (S/2AFC), there are two comparison patterns, one of which is different from the standard, and one of which is identical to the standard. The listener’s task, in this case, is to decide which comparison pattern is different from the standard.

In the simplest form of the *single-stimulus method*, one of two highly confusable patterns (i.e., patterns which have close to the same duration or tempo) is presented on each trial and the listener makes a binary judgment about which one it was (e.g., the short one or the long one, etc). In a many-to-few task, the binary response is maintained, but the number of confusable patterns is increased (i.e., given a single stimulus on each trial, the listener must separate the set of stimuli into two subsets). In an identification task, the number of response categories is also increased. The

listener's task is to assign each presented stimulus to one of the response categories. The primary difference between identification and category rating is that the stimuli are highly confusable.

2.5 Central Issues in Time Perception

Allan (1979) maps out four central issues in time perception that are investigated using the time-scaling and time-discrimination paradigms: the nature of the psychophysical law for time, the adequacy of Weber's law for describing time sensitivity, the source of the time-order error in time judgments, and the role of non-temporal information in time perception. This review focuses on the first two of these issues (the nature of the psychophysical law and the adequacy of Weber's law) which are pertinent for this thesis. The time-order error, which refers to the effects of sequential presentation order on time perception, is not addressed directly in this review. Although, this issue is intrinsically linked to the over- and underestimation of time intervals, and thus briefly touched on in the discussion of the nature of the psychophysical law for time. The role of non-temporal information in time perception, which refers to how changing physical properties of the stimulus, such as spectrum and intensity, can affect estimates of time-intervals, is also not addressed in this review.

2.5.1 The Nature of the Psychophysical Law

Many of the earliest studies in psychophysics investigating the psychophysical law for time, reported data in terms of the over- and underestimation of short and long time intervals relative to an intermediate *indifference interval* (cf. Fraisse 1978). As discussed by Fraisse, Horing (1864) reported that for time intervals between 0.3 and 1.4 seconds, intervals less than about 0.6 seconds are overestimated and intervals greater than about 0.6 seconds are underestimated. The terms underestimation, overestimation and indifference-interval have been used very loosely in the current literature and therefore need clarification. They refer to values calculated using either a comparison or ratio-setting (reproduction) task (Woodrow, 1951).

In the context of a reproduction task, in which the listener may be asked to tap their finger at a rate specified by the stimulus, overestimation indicates that the mean time-interval of the listener's reproductions is longer than the stimulus time interval and underestimation indicates that the mean time interval of the reproductions is shorter than the stimulus time interval. The indifference intervals are those for which the mean of the reproductions is approximately equal to the stimulus time interval, neither over- or underestimated.

In the context of a comparison task, the term indifference interval has been used in a way which is somewhat misleading. For example, in a 2AFC "which is longer" task, an indifference interval corresponds to the standard time interval for which the frequency of longer judgments is the same regardless of the *order* of the standard

and comparison time intervals. In this case, the indifference interval corresponds to a zero *time order* error. Positive and negative time order errors (in which case listeners exhibit a bias in the frequency of longer judgments on the comparison pattern) have been attributed to over- and underestimation of the standard time interval. Woodrow (1951) argues that associating over- and underestimation of the standard with the time-order error is misleading because the errors are due to the ordering of the standard and comparison patterns and are not *necessarily* due to biased subjective estimates of the standard.

For the purposes of consistency, the terms overestimation, underestimation, and indifference interval will be used in exactly the same way in this thesis for describing both reproduction data and comparison data. Overestimation will refer to mean subjective durations that are longer than the stimulus duration, underestimation will refer to mean subjective durations that are shorter than the stimulus duration, and indifference intervals will refer to mean subjective durations that are approximately equal to the stimulus duration. In the context of a comparison task, the terms overestimation, underestimation, and indifference interval will only be used when it is reasonable to suppose that discrimination errors are due to over- or underestimation of the standard time interval.

Many studies that have attempted to establish indifference intervals, have reported substantially varying results, with the reported indifference interval ranging from about 0.3 to 5.0 seconds (Fraisse, 1963)!. One explanation for the substantial variability observed in indifference interval estimates is due to Hollingworth (1910). Hollingworth proposed a *central tendency hypothesis* which states that estimates of a stimulus attribute, such as the size of a ball, have a tendency to reflect, or center on, the average of the stimulus values observed in the given setting. Thus, a ball is determined to be “large” or “small” based on ball sizes that we are familiar with. For time perception, the central tendency hypothesis implies that time estimates are over- or underestimated according to the mean of the range of time intervals that an individual has become familiar with. The often-reported indifference interval of 0.6 seconds may reflect the mean of the time-intervals that we experience in our day-to-day lives. However, prolonged exposure to a specific range of time-intervals with a mean that differs from 0.6 seconds, such as would occur during an experiment, may result in a shift in the indifference interval towards this mean. In support of a central-tendency effect, Fraisse (1948) found that for two ranges of time-intervals used with the same set of subjects, a significant difference was observed in the estimated indifference interval, which was correlated with the difference in the mean time intervals of the two stimulus sets.

Related to the central tendency hypothesis, (Fraisse, 1948) hypothesized that in time perception, small differences in duration are minimized (assimilated) and large differences in duration are maximized (exaggerated). Such *perceptual anchor* effects have been reported by a number of researchers (Fraisse, 1948; Goldstone et al., 1957; Goldstone et al., 1959; Turchioe, 1948; Woodrow, 1951). For example, Goldstone and

Boardman reported an experiment in which three groups of listeners estimated the duration of a 1.0 second tone. A preceding context tone with a 0.1 second duration was found to shrink the listeners' estimates of the 1.0 second duration, whereas a context tone with a 2.0 second duration was found to exaggerate the listeners' estimates of the 1.0 second duration. A context tone with a 1.0 second duration, included as a control condition, was found to have no significant biasing effect on time estimates. Very recently, Nakajima et al. (1992) and ten Hoopen et al. (1993), in claiming to have found a "new" time-illusion, reported similar anchor effects for very brief time-intervals (< 200 ms) using the method of comparison (see (Allan and Gibbon, 1994) for a discussion challenging the "newness" of this reported time illusion).

In an attempt to formalize the notions of short and long, (Fraisse, 1963) suggested a qualitative distinction among three time perception zones: short, long, and indifferent. According to his classification *short* intervals are less than 0.5 seconds, having the qualitative defining property that the beginning and ending sounds are what is perceived, rather than the time interval between them. *Indifferent* intervals between 0.5 and 1.0 seconds are neither short nor long and the events marking the interval form a perceptual unit. *Long* intervals are classified as lasting more than about 1.0 second and an effort is required to keep the events marking the time-interval within the same "psychological present".

In much contemporary research, the primary debate about the nature of the psychophysical law for time focuses on whether the relationship between clock time and subjective time is a power function (see Equation 1), as for most prothetic continua, or is instead a linear function of the form

$$\Psi = m\Phi + b. \quad (4)$$

Earlier research which presented data in terms of over- and underestimation of time intervals is, in principle, consistent with both the linear and power characterization of the psychophysical law for time. Thus, it is possible to determine parameter values for both linear and power functions which satisfy the following properties: (1) there is a single indifference interval $\bar{\Phi}$ for which $\Psi = \bar{\Phi}$; (2) intervals (Φ) shorter than the interval $\bar{\Phi}$ are overestimated ($\Psi > \Phi$); and (3) intervals (Φ) longer than $\bar{\Phi}$ are underestimated ($\Psi < \Phi$).

Figure 2.2 illustrates these three properties for a linear function and a power function. For both equations, 0.6 seconds is the indifference interval. For a linear function (Equation 4) to predict the overestimation of short intervals and the underestimation of long intervals, the slope m must be between 0.0 and 1.0 and the y -intercept b must be greater than zero. A y -intercept greater than zero implies that there is a minimum subjective duration, a concept initially proposed and supported by (Efron, 1973), and later contested by (Allan, 1979). For a power law

$$\Psi = k\Phi^r \quad (5)$$

to predict the over and underestimation of short and long time intervals, the constant k must be greater than 0.0 and the exponent r must be between 0.0 and 1.0 (i.e., the

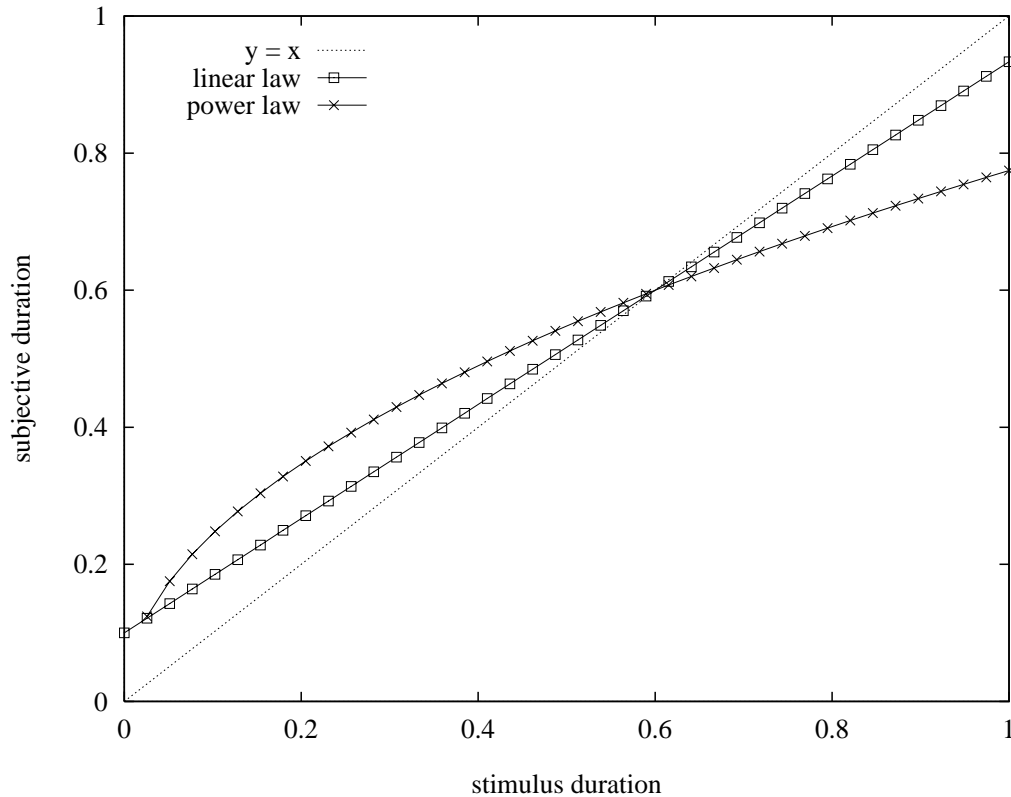


Figure 2.2: Linear and power functions consistent with the over- and underestimation of short and long intervals. In both equations, the indifference interval is 0.6 seconds.

power function is concave-down). As the exponent decreases, the magnitude of the over- and underestimation increases.

Conflicting data supporting both a linear and power law have been reported. The main support for a power law comes from ratio-setting data (Eisler, 1975) and from magnitude and verbal estimation data (Bobko et al., 1977; Eisler, 1975). However, there are three primary problems with interpreting these data as support for a power law: (1) there is enormous variability in the reported exponents; (2) there is low correlation between exponents reported in ratio-setting experiments and those reported in the estimation experiments; and (3) often the reported exponent is close to 1.0 and thus the data would be equally well fit with a linear function (Allan, 1979). The main *direct* support for a linear law comes from discrimination data (Creelman, 1962; Kinchla, 1972). Following a detailed comparison of linear and power function theories Allan (1979) concludes that, although the nature of the psychophysical law for time is still somewhat murky, empirical data seems to support a linear relationship. This issue is re-examined in Chapter 4 by evaluating the Entrainment Model's ability to discriminate tempo changes, under the assumption of a linear psychophysical law for single time intervals.

2.5.2 The Adequacy of Weber's Law for Time

Another debate centers around the determination of the just-noticeable time difference and whether or not it obeys Weber's law. If Weber's law holds across some range of intervals, the just-noticeable difference in duration (ΔT) is a constant proportion of the base interval T (i.e., $\frac{\Delta T}{T}$ is a constant). For example, if the just-noticeable difference is determined to be 10% of the base interval, then for a 1-second base interval, the just-noticeable duration change (i.e., that detected with 70% accuracy) is 100 ms. See Section 2.2.2 for a more detailed description of Weber's Law. The adequacy of Weber's law for describing human accuracy in the discrimination of time changes has been investigated for isolated intervals (empty and filled), for intervals embedded within a larger temporal context (i.e., a rhythmic pattern), and for tempo. These data will be reviewed in the next three sections in some detail, as time discrimination is the focus of the modeling efforts reported in Chapter 4.

Before reviewing these data, the clarification of a few specific terms is necessary. The term isolated-interval discrimination refers to the discrimination of duration changes to isolated intervals (empty and filled). The review of these data does not address how isolated-interval discrimination is influenced by non-temporal information such as the spectra of the sounds marking or filling the interval. The term embedded-interval discrimination will be used to refer to the discrimination of a duration change to a single interval that is embedded within a larger sequence of intervals (e.g., the discrimination of a duration change to the third interval of a ten interval sequence). The term tempo discrimination will be restricted to mean the detection of a tempo change to an isochronous sequence. The discrimination procedures used to establish just-noticeable time differences include the method of comparison and the method of reproduction. With the method of reproduction, the standard deviation of the reproduced intervals is often used as the measure of JND. Thus, if asked to reproduce a 1 sec time interval by repeatedly tapping a finger, the standard deviation of the between-tap time measurements can be used as a measure of JND.

Isolated-Interval Discrimination

By far, the largest number of time-discrimination studies have addressed the discrimination of duration changes to isolated intervals. Studies of isolated intervals can be divided into three groups: (1) those reporting data which support a simple or generalized version of Weber's law (Divenyi and Danner, 1977; Getty, 1975; Treisman, 1963), (2) those reporting data which support a non-linear increasing relationship between the ΔT and T (Abel, 1972; Chistovitch, 1959; Creelman, 1962; Drake and Botte, 1993; Small and Cambell, 1962) (Weber's law predicts a linear relationship between ΔT and T), and (3) those reporting data which support an invariant or decreasing ΔT (Allan and Kristofferson, 1974; Allan, 1979; Kristofferson, 1977; Kristofferson, 1980).

The most frequently cited articles supporting Weber's law for isolated interval

discrimination are Getty (1975) and Treisman (1963). However, both Getty's and Treisman's data do not in fact support the simple version of Weber's law: $\frac{\Delta T}{T} = k$. Using both the method of reproduction and the method of comparison, Treisman found that between 25 and 3000 ms, JNDs are not strictly proportional to T , but are consistent with a generalized version of Weber's law (Treisman, 1963):

$$\frac{\Delta T}{T + b} = k \quad (6)$$

where b is the constant "noise" present in the system. Like the simple version of Weber's law (Equation 2), Treisman's generalized rule proposes a linear relationship between ΔT and T . However, for this generalized version of Weber's law, the ratio $\frac{\Delta T}{T}$ is not constant, but instead varies as a function of T :

$$\frac{\Delta T}{T} = k \frac{(1 + b)}{T}. \quad (7)$$

Using the method of reproduction, Getty found that data from two well-trained listeners for intervals between 50 and 2000 ms were in agreement with a different generalized version of Weber's law (Getty, 1975):

$$\Delta T = (k^2 T^2 + b^2)^{\frac{1}{2}}. \quad (8)$$

With this rule, Getty hypothesized that time sensitivity was determined by two factors: the base interval (T) and the intrinsic timing variability (b^2) of a reproduction task. Like Treisman's generalized version of Weber's law, Getty's version does not predict a constant ratio $\frac{\Delta T}{T}$.

Using the method of comparison, Divenyi and Danner (1977) report data averaged over three subjects for which they calculate a power-law (see Equation 1) exponent of 0.93 for describing the relationship between ΔT and T for empty interval discrimination. Since this exponent is close to 1.0 (consistent with both a linear and power-law relationship) they argue that time discrimination obeys Weber's law. This conclusion is somewhat weak because the data used to calculate the power-law exponent included only three base values of T : 25, 80, and 320 ms.

A much larger body of work supports a non-linear relationship between ΔT and T . Divenyi and Danner (1977) calculate power-law exponents of 0.78, 0.74, and 0.77 for duration discrimination data from Abel (1972), Chistovitch (1959), and Creelman (1962) respectively. Getty's generalized version of Weber's law could also be included in this group since it actually specifies a non-linear relationship between ΔT and T .

Kristofferson and colleagues (Allan and Kristofferson, 1974; Allan, 1979; Kristofferson, 1977; Kristofferson, 1980) have conducted a series of experiments which suggest that at least across a range of time intervals, JND is constant. Since constant JND implies that ΔT is independent of the base interval T , this is a very significant departure from Weber's law. Allan and Kristofferson (1974) initially propose that

JND is related to a fundamental time quantum. For time intervals between 100 ms and 2000 ms, they compute that this time quantum is about 50 ms. However, JND data from subsequent discrimination studies (Kristofferson, 1977; Kristofferson, 1980) was more step-like, suggesting that the time quantum (initially reported as 50 ms) is halved and doubled as a function of the base interval T . In these subsequent studies, the time quantum is reported as 13 ms for a base interval of 100 ms, as 25 ms for a base interval of 200 ms, as 50 ms for a base interval of 400 ms, and as 100 ms for a base interval of 800 ms.

In summary, the isolated interval data shows at least that the simple form of Weber's law does not hold across the range of time intervals between 50 ms and 2000 ms. Instead, the relationship between ΔT and T in this range of time intervals is better described by a non-linear (increasing) function for a majority of the studies. Kristofferson and associates are the only researchers, as far as I'm aware, to report single interval data for which JND is constant or varies as a step function of a base interval. Furthermore, the variability in their estimates of the time quantum leads one to question the validity of their step-function description of the relationship between ΔT and T .

Embedded-Interval Discrimination

The issue of the adequacy of Weber's law for describing time sensitivity is significantly complicated when the time change to be detected is embedded within an auditory pattern. In such cases, a number of contextual factors have been shown to influence JNDs in the interval targeted for a time change. When the tones surrounding the target interval vary in frequency, listeners' abilities to detect time changes in the target interval are worse than without variations in frequency, especially when there is uncertainty on each trial as to the pattern of tone frequencies comprising the target context (Espinoza-Varas and Watson, 1986; Monahan and Hirsh, 1990; Sorkin et al., 1982). Similarly, listeners' abilities to detect time changes are worse when the pattern of intervals surrounding the target interval are irregular (or non-metrical) than when the context intervals are regular (or metrical), as well as worse when there is uncertainty in the temporal context of the target interval (Bharucha and Pryor, 1986; Drake and Botte, 1993).

This discussion of embedded-interval discrimination focuses on differences between listeners' abilities to detect time changes in isolated intervals and listeners' abilities to detect time changes to a target interval within an *isochronous* context. In such cases, listeners are typically asked to detect delays or advances in the onset of a single tone within an otherwise isochronous (fixed inter-onset-interval) pattern. For onset delays, the interval preceding the onset is lengthened and the interval following the onset is shortened to preserve the overall sequence duration, while for an onset advance, the preceding interval is shortened and the following interval is lengthened. Of course, when the onset delay or advance occurs to the last interval of the sequence it is not

possible to shorten or lengthen the following interval to preserve the duration of the sequence.

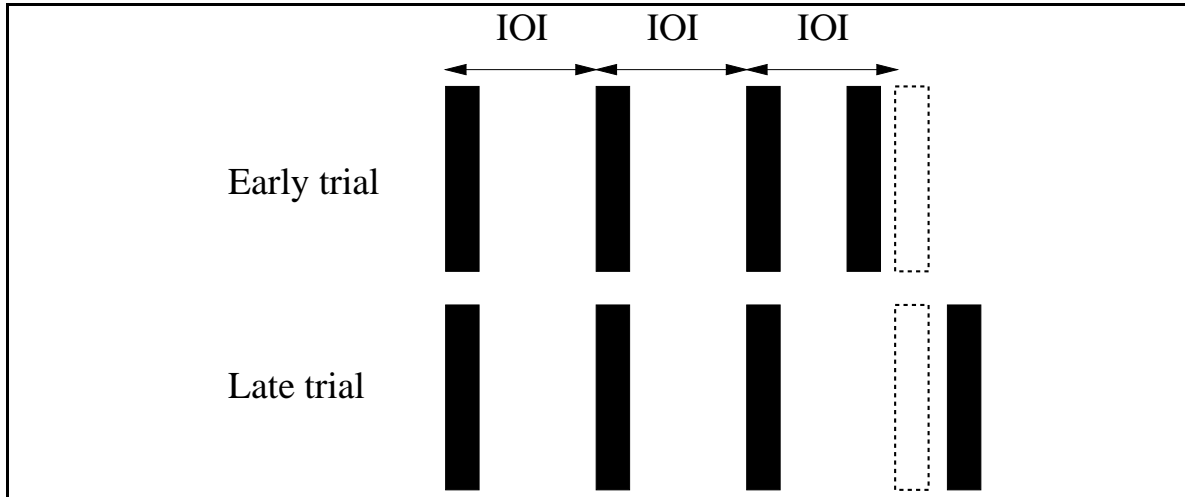


Figure 2.3: Example of “early” and “late” experimental trials from Halpern and Darwin (1982).

In one such embedded-interval study, Halpern and Darwin (1982) investigated the discrimination of a single interval embedded within a four-tone (three interval) isochronous pattern. The listener’s task in this case was to determine whether the last click (a very brief tone) in a four-click sequence was early or late with respect to the expected onset time established by the fixed inter-onset-interval (IOI) of the first two intervals (see Figure 2.3). The investigated range of IOIs was between 450 and 1450 ms. For the third interval of the sequence, the IOI was either 10% shorter or 10% longer than the IOI of the first three clicks (i.e., the last click of the sequence was either early or late). In this experiment, Halpern and Darwin found that relative JNDs did not vary significantly from a Weber’s law description of the data, with the mean constant relative JND reported as 5.4%. However, the overall trend of the data is in the direction of a U-shaped relative JND curve with a minimum relative JND of 4.7% at the 1000 ms IOI and maximum values of 6.4% and 5.5% at the 450 ms IOI and 1450 ms IOI respectively. In addition, listeners in the Halpern and Darwin experiment were found to exhibit a bias for responding “early” to sequences with IOIs less than 550 ms and for responding late to sequences with IOIs greater than 700 ms. These listener response biases as a function of the IOI are simply explained if it is assumed that listeners are overestimating “short” IOIs and underestimating “long” IOIs, as have been reported in a number of the single interval studies investigating the nature of the psychophysical law for time (see Section 2.5.1). If listeners overestimate the IOI of the first two intervals, then the expected onset of the fourth click is late with respect to the actual onset, and therefore the listeners should exhibit a bias for responding early. On the other hand, if listeners underestimate the IOI of the first

two intervals, then the expected onset of the fourth click is early with respect to the actual onset, and therefore the listeners should exhibit a bias for responding late.

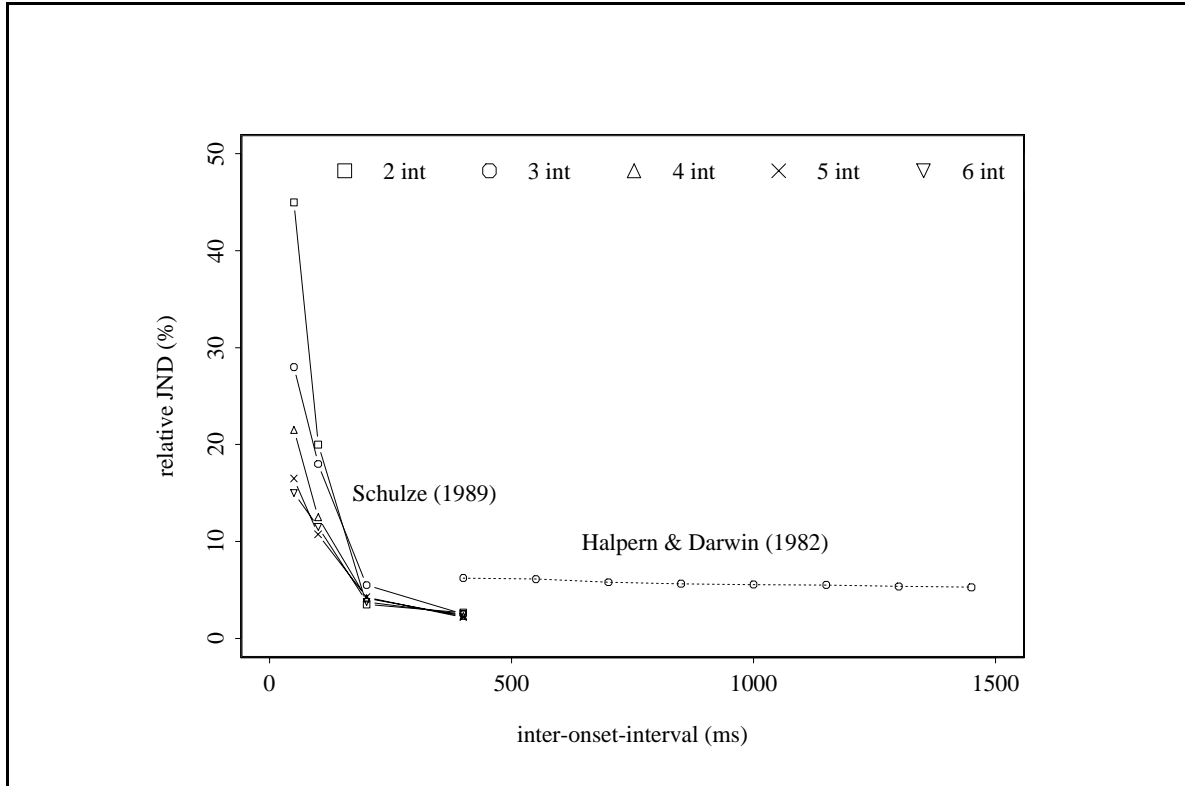


Figure 2.4: A comparison between the embedded-interval discrimination data of Schulze (1989) and that of Halpern and Darwin (1982).

Schulze (1989) designed an experiment similar to that of Halpern and Darwin (1982), but instead of making an early/late distinction, listeners determined whether the last interval of an auditory pattern was the same duration or longer than the preceding isochronous intervals. In the Halpern and Darwin study, the evaluation of the data focused on testing Weber's law. Schulze, on the other hand, was primarily interested in investigating how varying the number of isochronous intervals in an auditory pattern influences time sensitivity. Schulze hypothesized that listeners' abilities to detect a deviation in an interval should improve as a function of the number of preceding observations of the same interval. Thus, increasing the number of preceding intervals identical to the tested interval should decrease the relative JND in the tested interval. To test this multiple-look hypothesis, relative JNDs in the last interval of a sequence were determined for 2- to 6-interval sequences and for IOIs of 100, 200, 300, and 400 ms.

Figure 2.4 shows the data from both experiments. All data in the figure are plotted in terms of the relative JND as a function of the base IOI. Whereas Halpern and Darwin's data are consistent with Weber's law, Schulze reports data that are

not consistent with Weber's law. However, this inconsistency is minor considering the 4-tone 400 ms IOI is the only condition in the Schulze study overlapping with the Halpern and Darwin study. In the Schulze (1989) study, not only is relative JND a decreasing function of the base IOI, the absolute JND is found to be constant (or perhaps a decreasing function of the base IOI). Thus, Schulze's embedded-interval data provides some support for the constant JND claim of Kristofferson and colleagues (Kristofferson, 1980). And Schulze and Kristofferson argue that one reason for the reported departures from Weber's law may be due to their extensive training of the listeners. Differences supposedly due to "training effects" are difficult to evaluate, especially in retrospect. However, it is clear that the combined data from Schulze (1989) and Halpern and Darwin (1982) support the hypothesis that *relative* JNDs are a decreasing function of IOI for short IOIs (< 300 ms), are fairly constant for an intermediate range of IOIs, and perhaps increase as function of IOI for the longest IOIs.

With respect to the influence of the number of intervals, Schulze reports that, as predicted by the multiple-look hypothesis, increasing the number of intervals preceding the altered interval heightens time sensitivity. Furthermore, improvement is greater for the shorter IOIs (50 and 100 ms) than for the longer IOIs (200 and 400 ms), leading Schulze to suggest that at about the 100 ms IOI, a shift occurs in the type of processing listeners use to discriminate time changes.

In an expansion of the Schulze paradigm, Hirsh et al. (1990) investigated listeners' abilities to detect deviations in intervals embedded at a number of different positions within 6- and 10-tone isochronous sequences, for IOIs of 50, 100, and 200 ms. Consistent with most of the isolated interval data, they found that relative JND was not constant across the investigated range of IOIs, as Weber's law would predict, but increased as a function of the IOI. For the 50 ms IOI, the mean relative JND combining all interval positions was 15% – 30%, decreasing to 10% – 15% for the 100 ms IOI, and to a minimum of 5% – 6% for the 200 ms IOI.

Somewhat surprising with respect to the reported Schulze data that supported a multiple-look hypothesis, Hirsh et al. (1990), found that relative JNDs were independent of the position of the delay, except at the 50 ms IOI. However, for the 50 ms IOI relative JNDs were found to decrease significantly as a function of the position of the altered interval, as would be predicted by a multiple-look hypothesis; i.e., interval deviations early in the sequence were harder to detect than interval deviations later in the sequence. A similar, but non-significant, trend was observed for the 100 and 200 ms IOIs. This interaction between the observed improvement due to the later position of the altered interval and the specific IOI condition lead Hirsh et al. (1990) to conclude, like Schulze (1989), that short IOIs are processed differently than longer IOIs, placing the boundary for differential processing of time intervals at about 100 ms, the same as reported by Schulze (1989). This differential processing hypothesis is a common theme in theories of time perception.

In a generalization of the Hirsh et al. (1990) study, ten Hoopen et al. (1994) used

the method of comparison to investigate listeners' abilities to detect anisochrony for base IOIs between 60 and 720 ms. Thus, instead of delaying or advancing a single onset within the comparison pattern, every other onset of the comparison pattern was either delayed or advanced, resulting in an anisochronous "duple" pattern. The listener's task was to detect the presence or absence of anisochrony in the comparison pattern. For IOIs shorter than 300 ms, relative JNDs were found to be a decreasing function of IOI. However, for IOIs greater than 300 ms, relative JNDs were found to be fairly constant (in agreement with Weber's law). These data are consistent with the combined data of Schulze (1989) and Hirsh et al. (1990). Ten Hoopen et al. (1994) conclude that their data offer additional support for distinct processing of short and long intervals.

To summarize these embedded-interval data, relative JNDs for target intervals within an isochronous context are adequately described by Weber's law only above 300 ms. For IOIs, less than about 300 ms, relative JNDs are a decreasing function of IOI and for IOIs greater 300 ms, relative JNDs are fairly constant, with some data suggesting that relative JNDs gradually increase for the longer IOIs, resulting in a U-shaped relative JND curve. Although there is still considerable debate, placing the embedded interval at a later position in the sequence tends to improve discrimination thresholds, *especially* for the shortest IOIs. This has led a number of researchers to suggest that intervals less than about 100-300 ms are processed differently than intervals greater than about 300 ms.

Tempo Discrimination

Tempo discrimination is yet another type of time judgment task. Similar to embedded-interval discrimination, Weber's law seems to apply only under certain conditions. One important issue concerns the relationship between isolated-interval discrimination and tempo discrimination. A somewhat surprising fact is that up until a seminal report by (Michon, 1964), apparently no studies made a direct comparison between isolated-interval discrimination and tempo discrimination. In fact, in reviewing a recent (at the time) series of publications on tempo sensitivity (Pollack, 1952; Mowbray, 1956), Michon acerbically points out that:

This research [that on tempo] appears to be completely independent of the orthodox strain of time psychologists [those that studied isolated interval discrimination] because one can hardly find any cross reference in the publications from both sides.

Therefore, in order to be able to compare tempo data with that of isolated interval data, Michon defined the tempo of a pattern in terms of the fixed IOI of that pattern. This is in contrast to the convention of defining the tempo of a pattern as the number of events/beats that occur per unit time (see Section 1.1). Thus, using the IOI definition of tempo, short IOIs correspond to fast tempos and long IOIs correspond

to slow tempos. For the same reason of comparability, I will continue to define tempo as a function of IOI throughout this thesis.

Using the method of comparison, Michon measured listeners' abilities to detect changes in the tempo of isochronous sequences for IOIs between 67 and 2700 ms. Although it is not completely clear from Michon's description of the experiment, it appears as though Michon used fairly long sequences (i.e., each sequence was composed of a large number of events). Listeners' performance in the experiment improved significantly over the course of five experimental sessions. The average sensitivity of the listeners in each of the first four sessions, relative to their performance in the final session, was 1.9, 1.6, 1.4, and 1.0 respectively. Data were reported from only the final session and thus reflected very well-trained listening performance.

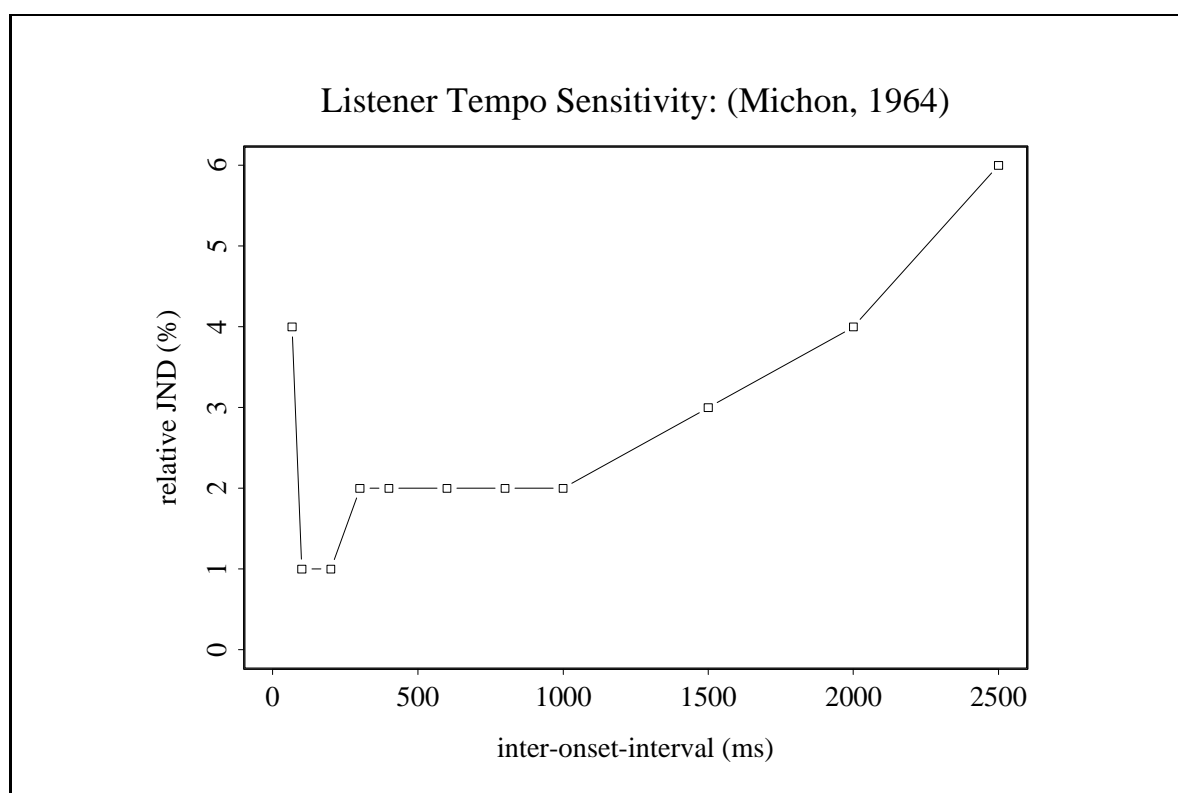


Figure 2.5: Tempo discrimination data reproduced from Michon (1964).

Michon reported a minimum relative JND of 1.0% for the 100 ms IOI and a secondary minimum region of about 2.0% for IOIs between about 300 and 1000 ms. Both minimum regions were significantly lower than the minimum relative JND of about 5% reported for isolated intervals. Michon's data (approximated from Michon (1964)) are shown in Figure 2.5. Consistent with data from many of the isolated interval studies, relative JNDs were fairly constant for IOIs between 300 and 1000 ms, gradually increasing for IOIs longer than about 1000 ms. However, unlike the isolated interval data, relative JNDs did not immediately increase for IOIs less than 300 ms.

Instead, for IOIs between 100 and 300 ms, relative JNDs were fairly constant and even lower than the relative JNDs for IOIs between 300 and 1000 ms. The existence of this double minimum was perplexing to Michon and suggesting to him that fast sequences (IOIs < 300 ms) might be processed differently than slow sequences (IOIs > 300 ms), perhaps by two relatively independent timing mechanisms, each most accurate at a different rate. Thus, the existence of the double minimum in the tempo data and not in the isolated interval data indicated to Michon that engaging the *fast* timing mechanism, which improved time sensitivity for the short IOIs, required the repetition of the time intervals. Although not mentioned by either Schulze (1989) or Hirsh et al. (1990) this conclusion of Michon's is consistent with their observation that increasing the number of intervals (i.e., repeating a time interval) especially improves time sensitivity for the shortest IOIs in their studies.

Expanding on the Michon (1964) study, Drake and Botte (1993) conducted a series of tempo discrimination experiments, the design of which clearly benefitted from the intervening thirty years of time perception research. Three main issues addressed in these experiments were related to the adequacy of Weber's law for time. First, what is the relationship between isolated-interval sensitivity and tempo sensitivity? Second, what is the effect of increasing the number of isochronous intervals on tempo sensitivity? And finally, what influence is there of musical training on tempo sensitivity? We will examine the data from this series of experiments and a followup study (Drake and Botte, 1994) in detail, as these data, which are the most comprehensive tempo-discrimination data for isochronous sequences to date, form a basis for evaluating the predictions of the Entrainment Model developed in Chapter 4.

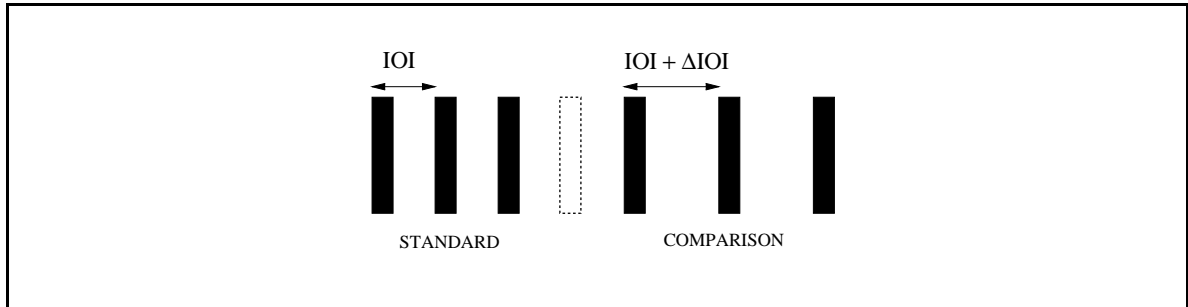


Figure 2.6: “Which is faster” tempo discrimination task: the subject listens to a standard isochronous sequence followed by a comparison sequence that is either faster or slower than the standard, and then indicates which sequence is faster.

The Drake and Botte (1993) study investigated listeners' abilities to detect changes in the tempo of isochronous sequences for tempos (IOIs) between 100 to 1500 ms, for one-, two-, four-, and six-interval sequences. The one-interval patterns permit a direct comparison with the isolated-interval data. Discrimination thresholds were determined using the method of comparison and an *adaptive-tracking* procedure (Levitt, 1971). Each experimental trial consisted of a fixed standard sequence followed by a

comparison sequence, illustrated in Figure 2.6. The listener's task was to determine which sequence was faster. Two successive correct responses resulted in a 1% decrease in the tempo difference between the standard and comparison sequences (measured as a percentage of the standard sequence's IOI). An incorrect response increased the tempo difference by 1%. This "2-down/1-up" adaptive procedure converges to the relative JND for tempo discrimination (or the tempo difference necessary for 70.7% correct discrimination judgments).

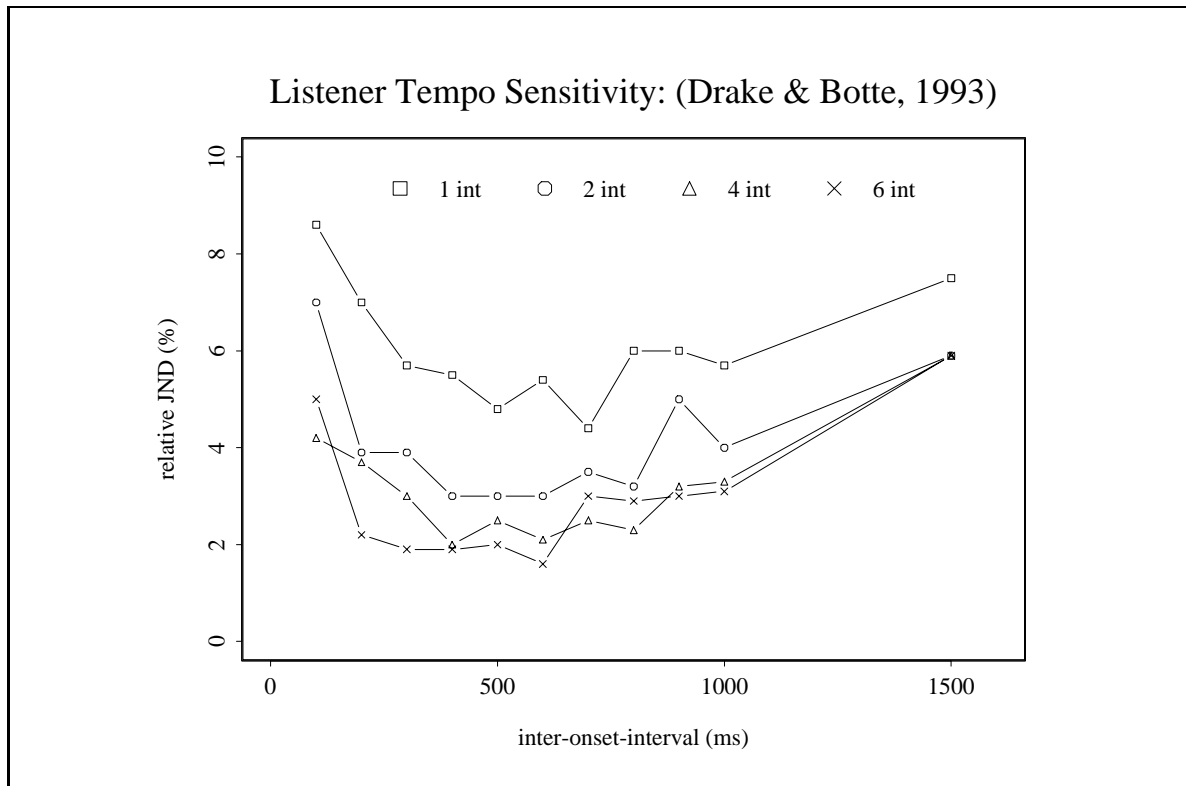


Figure 2.7: Tempo discrimination data reproduced from Experiment 1 of the Drake and Botte (1993) study. Mean relative JNDs are shown for 1-, 2-, 4-, and 6-interval sequences for IOIs between 100 and 1500 ms.

The discrimination data from the first of these experiments are shown in Figure 2.7. Relative JNDs for tempos between 100 and 1500 ms were not constant, as would be predicted by Weber's law, but instead were U-shaped as a function of the IOI. For IOIs less than about 300 ms, relative JNDs were a decreasing function of IOI. For IOIs in between 300 and 900 ms, IOIs were fairly constant and at a minimum value. For IOIs longer than about 900 ms, mean relative JNDs gradually increased again. As predicted by a multiple-look hypothesis, relative JNDs were smaller for multiple-interval sequences than for single-interval sequences and could be described as a decreasing function of the number of intervals.

The mean relative JND was 6% for the single-interval sequences, which is within the range of relative JNDs reported for isolated-interval discrimination. For multiple-interval sequences, the mean threshold improved to 3.4%. The best performance was for a 6-interval sequence in the 400-ms IOI condition, with the reported threshold below 2%. For this condition, the average listener was able to reliably detect a 8-ms change in the fixed 400-ms IOI! As shown in Figure 2.7, when increasing the number of intervals in the sequences, Drake and Botte observed more improvement in the relative JNDs for the short IOIs (fast tempos) than the long IOIs (slow tempos). For example, for the 1500-ms IOI condition, relative JNDs stopped decreasing after the addition of a second interval, but for IOIs less than 300 ms, relative JNDs improved through the addition of six intervals. This differential pattern of improved sensitivity is consistent with Michon's suggestion that multiple intervals are necessary to engage a "fast" timing mechanism which heightens time sensitivity for short IOIs only.

In a separate experiment, Drake and Botte (1993) examined the influence of musical training on tempo sensitivity. Tempo sensitivity was compared between musicians and non-musicians for single and multiple-interval sequences for IOIs of 300, 600, and 900 ms. In general, the musicians were able to detect smaller changes in tempo than non-musicians, for both single and multiple-interval sequences. For the single-interval sequences, the minimum relative JND occurred at the 600-ms IOI condition for both musicians and non-musicians. However, for the multiple-interval sequences, the minimum relative JND for the musicians extended to the 300-ms IOI condition. Drake and Botte interpret this result as providing additional evidence that distinct processing occurs for fast sequences (those with IOIs less than 300 ms), assuming that musical training influences one's ability to take advantage of this processing.

Drake and Botte (1993) suggest three zones of tempo sensitivity: (1) a zone of optimal sensitivity between 300 and 900 ms, for which relative JND is fairly constant and thus for which Weber's law holds, (2) a zone of lesser sensitivity for tempos slower than the 800-ms IOI condition, for which relative JNDs are higher than predicted by a multiple-look model (to be described in Section 2.6), and (3) a zone of potentially heightened sensitivity for tempos faster than the 300-ms IOI condition, for which increasing the number of intervals in the sequences improves thresholds more than would be predicted by a simple multiple-look model. This third zone of potentially heightened tempo sensitivity is evidence that supports distinct processing of IOIs less than about 300 ms.

Comparing the tempo discrimination data from Drake and Botte (1993) (see Figure 2.7) with that from Michon (1964) (see Figure 2.5), we find a significant inconsistency between the two: Drake and Botte report a single region of optimal sensitivity, for IOIs between 300 and 900 ms, whereas Michon reports two optimal regions, one at about 100 ms and one at about 600 ms. A possible explanation for this inconsistency is that in Michon's experiment, sequence duration was apparently fixed, although this is not completely clear from his description of the experiment. If so, sequences with different IOIs would be composed of different numbers of intervals; for example, for a

2-second sequence duration, sequences with an IOI of 100 ms would have consisted of 20 intervals, whereas sequences with an IOI of 500 ms would have consisted of only 4 intervals. Thus, the faster sequences would have had more intervals than the slower sequences for a given sequence duration. In contrast, Drake and Botte (1993) measured relative JNDs for different IOIs for a fixed number of intervals. This suggests that the qualitative difference between their tempo discrimination data and Michon's may be due to the differing number of intervals for each IOI condition.

In the hope of clarifying the differences between their tempo data and Michon's, Drake and Botte (1994) investigated the extent to which increasing the number of intervals in a sequence improves discrimination thresholds. They focused on three primary questions: (1) If relative JND decreases as a function of the number of intervals in the sequence, then at what point will the addition of an interval stop improving thresholds? (2) Does the number of intervals (n) in Question 1 vary as a function of the IOI? (3) Is n determined by a particular sequence duration d (the temporal window) as well as the IOI (i.e., is $n = \frac{d}{IOI}$)?

To address these questions, Drake and Botte (1994) measured relative JNDs for IOI conditions between 50 and 1500 ms for which the number of intervals in a sequence of a given IOI condition was increased, one interval at a time, with relative JND measured after each addition, until the addition of an interval no longer lowered the relative JND. Let n equal the number of intervals beyond which no reduction of the relative JND occurs. Then, define *optimal sequence duration* d as n multiplied by the IOI: $d = n[IOI]$. In which case, the relative JND at the optimal sequence duration d can be thought of as the optimal relative JND for that IOI condition. If Michon used fairly long sequences, then we can assume that Michon was measuring optimal relative JNDs for each IOI condition.

Figure 2.8 compares the optimal relative JNDs obtained by Drake and Botte (1994) and those obtained by Michon (1964) for each IOI condition. Notice that when Drake and Botte's data are plotted as optimal relative JND, there are two zones of optimal sensitivity, one at about 100 ms and one at about 600 ms, as reported by Michon, thus suggesting a resolution to the inconsistency. However, these two regions are not well defined and there are discrepancies between the Michon data and the Drake and Botte optimal data, with performance in the Michon experiment significantly better than in the Drake and Botte experiments.

Based on their results, Drake and Botte (1994) propose that optimal sequence duration represents listeners' inability to integrate timing evidence beyond a limited temporal window, which limits the number of useful sequence intervals (as shown in Figure 2.9). If window duration (i.e., optimal sequence duration) is plotted as a function of IOI, a discontinuity is observed in window duration between IOIs of 300 and 500 ms. For IOIs shorter than 300 ms, the window duration is approximately 1 sec, whereas for IOIs longer than 500 ms, the window duration is approximately 2.5 seconds. Drake and Botte suggest that this discontinuity in window duration provides strong additional evidence that short intervals (fast tempos) and long intervals

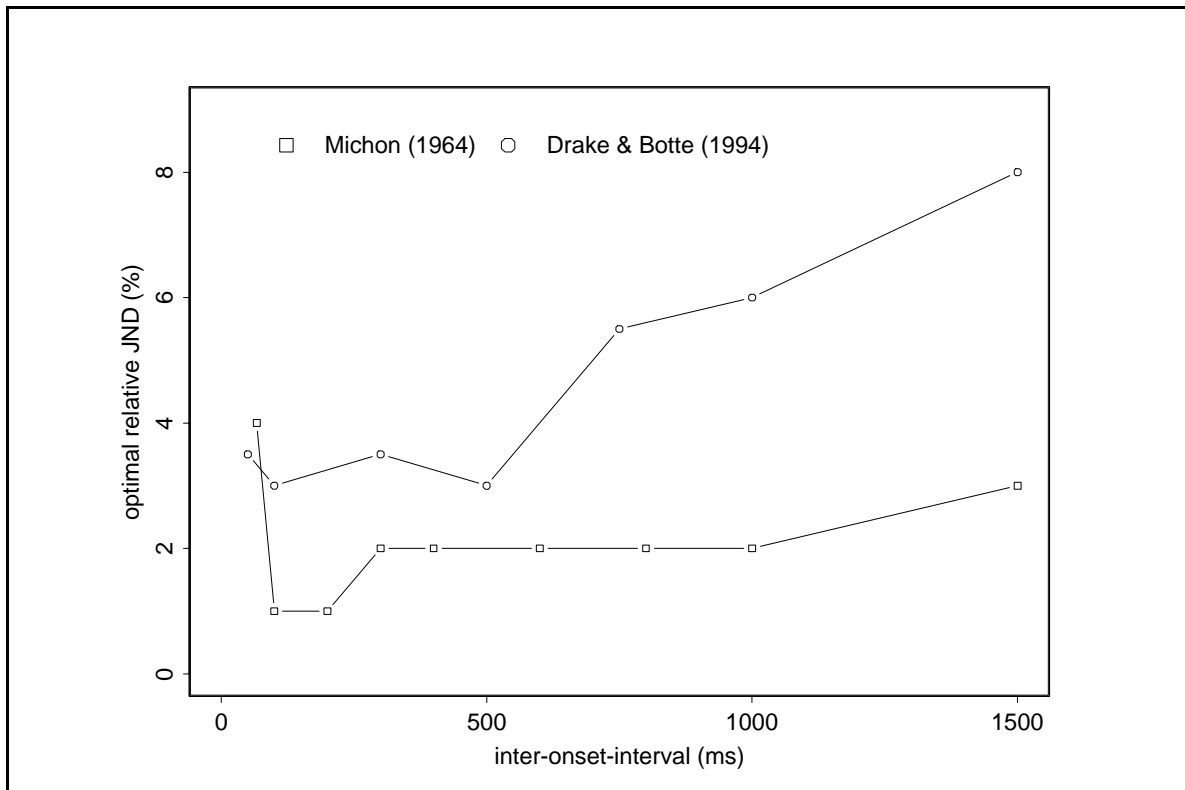


Figure 2.8: A comparison between the optimal relative JNDs obtained by Drake and Botte (1993) and those obtained by Michon (1964). The data shown are approximated from these studies.

(slow tempos) are processed by different timing mechanisms, and that the boundary between the type of processing occurs between 300 and 500 ms.

2.5.3 Summary

In this section, we defined four central issues in time psychophysics: (1) the nature of the psychophysical law for time, (2) the adequacy of Weber's law for describing time sensitivity, (3) the source of the time-order error in time judgments, and (4) the role of non-temporal information in time perception. Our discussion focused on the nature of the psychophysical law and the adequacy of Weber's law.

With regard to the first issue, early studies investigating the nature of the psychophysical law for time reported data in terms of the overestimation of short intervals and the underestimation of long intervals relative to an intermediate indifference interval. Although attempts to pinpoint the indifference interval created controversy because of the conflicting values obtained, it has been generally accepted that intervals shorter than approximately 500 ms are overestimated (exhibit positive constant error), and intervals longer than approximately 1000 ms are under-estimated (exhibit

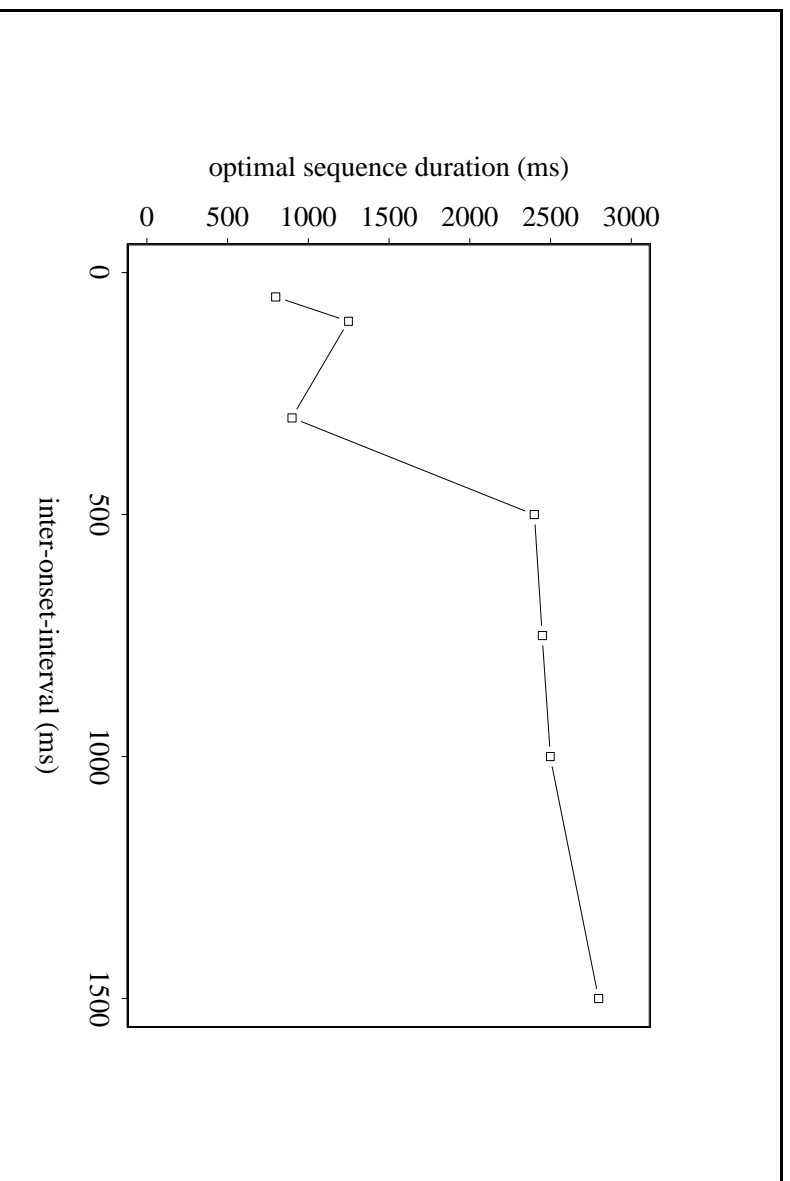


Figure 2.9: Optimal sequence duration as a function of IOI, reproduced from Drake and Botte (1994)

negative constant error), with the indifference interval (exhibiting zero constant error) falling somewhere in the range of 500 to 1000 ms (Fraisse, 1982). The main point to be gained from this research is that over- and underestimation does occur relative to some intermediate range of durations; it is not the precise location for the cutoff values for over- and underestimation that matters.

The debate about the nature of the psychophysical law for time eventually shifted away from over- and underestimation of time to evaluating whether subjective time estimates are better described by a linear or a power law. With regard to this issue, there is still considerable debate, given the existing data that supports both a linear and power psychophysical law. The over- and underestimation data provide little help in resolving this debate since, given the correct choice of parameters, both the linear and the power functions will predict over- and underestimation of time, in agreement with the earlier studies.

Regarding the second issue, just-noticeable time differences have been investigated for isolated intervals, for embedded intervals, and for tempos, all with the purpose of evaluating the adequacy of Weber's law for describing time sensitivity. For isolated intervals (empty and filled), it is generally accepted that Weber's law does not provide an adequate fit to JNDs obtained across the entire range of intervals between 50 and

2000 ms. For these data, time sensitivity can be divided into three zones. (1) A zone for IOI conditions between 300 ms and 1000 ms, for which relative JND is generally fairly constant (in agreement with Weber's law) at a minimum value. (2) A zone for IOI conditions less than about 300 ms, for which relative JND abruptly increases with shorter IOIs. (3) A zone for IOI conditions greater than 1000 ms, for which relative JND gradually increases with longer IOIs. Thus, the overall shape of the relative JND curve for isolated intervals is U-shaped, although there is considerable debate about whether or not relative JND increases for IOIs longer than about 1000 ms (see Halpern and Darwin, 1982).

U-shaped relative JNDs have also been reported for studies investigating the discrimination of a target interval embedded within an isochronous context. However, these data suggest that the repetition of an interval lowers relative JND as a function of the number of repeated intervals preceding the tested interval, thus possibly influencing the temporal location of the boundaries between the three zones of time sensitivity outlined above. Both Schulze (1989) and Hirsh et al. (1990) observed that decreases in the relative JND are greatest for IOI conditions less than about 100 ms, suggesting that the zone of constant minimum relative JNDs (zone 2) extends to a shorter IOI condition. This greater improvement in the relative JND for short IOIs has led a number of researchers to propose that listener use a different type of processing for short IOIs (less than 300 ms) than they do for longer IOIs (greater than 300 ms). (Schulze, 1989; Hirsh et al., 1990; ten Hoopen et al., 1994).

U-shaped relative JNDs have also been attributed to tempo discrimination for which the tempo of a sequence is defined in terms of the fixed IOI of the sequence. Drake and Botte (1993) proposed three zones of tempo sensitivity that are analogous to the three zones of time sensitivity for isolated intervals (stated above): (1) a zone of optimal tempo sensitivity for IOI conditions between 300 and 900 ms, for which relative JND is fairly constant (in agreement with Weber's law); (2) a zone of lesser (but potentially heightened) tempo sensitivity for IOI conditions less than 300 ms, for which relative JND abruptly increases for shorter IOIs when the sequence has only a few intervals, but for which relative JND remains fairly constant for shorter IOIs when the sequences has a sufficiently large number of intervals; (3) a zone of lesser sensitivity for tempos slower than the 900-ms IOI condition, for which relative JND is an increasing function of IOI and is not substantially lowered by increasing the number of isochronous intervals. Michon argued that the heightened tempo sensitivity observed for IOI conditions in what we are calling zone 2, is due to engaging a "fast" timing mechanism which requires the repetition of the time intervals. Drake and Botte (1994) provided additional evidence for distinct processing of fast sequences by showing that there is an abrupt discontinuity in the optimal sequence duration (temporal window size for Drake and Botte) that occurs between the 300-ms and 500-ms IOI conditions.

2.6 Competing Theories of Time Perception

Models which attempt to account for the perception of time intervals which comprise rhythmic patterns can be divided into four main classes: (1) clock-counter models, (2) quantal models, (3) multiple-look models, and (4) dynamic-attending/contrast models. The following review will address the general assumptions and predictions of each of these approaches to time perception. The emphasis of the discussion will be on a multiple-look model proposed by Drake and Botte (1993) (one of the few models which addresses the relationship between the discrimination of isolated intervals and the discrimination of tempo changes to isochronous sequences) and the dynamic-attending/contrast model proposed by Jones and Boltz (1989) (a kindred spirit of the Entrainment Model proposed in this thesis).

2.6.1 Clock-Counter Models

By far, most models of human time perception are based on a modular information-processing approach, in which there is a central timer (or clock), perceptual store, a reference memory, and a comparator (Church and Broadbent, 1990). Such clock-counter models (not to be confused with the clock model proposed for rhythm perception by Povel and Essens (1985)) propose that time is measured using a central timer (the clock) which *accumulates* (or counts) duration in a perceptual store. This measurement of duration is then transferred to a reference memory for possible later comparison with other durations (measured by the central timer). A flow diagram outlining this approach is illustrated in Figure 2.10. A large variety of clock-counter models have been proposed (Abel, 1972; Creelman, 1962; Church and Broadbent, 1990; Divenyi and Danner, 1977; Killeen and Weiss, 1987; Miall, 1989; Treisman, 1963); these differ mainly in the form of the central timer and perceptual store.

For the majority of the clock-counter models, the form of the central timer is a single pacemaker (oscillator) which generates internal pulses at an average rate (λ) and the form of the perceptual store is a “count” of the number of pulses that occur during the duration of the to-be-measured time interval (T) (Abel, 1972; Creelman, 1962; Treisman, 1963; Divenyi and Danner, 1977). A comparison is made between two time intervals (T and $T + \Delta T$) by comparing the count for the first interval (λT) (maintained in the reference memory) with the count of the second interval ($\lambda[T + \Delta T]$) (maintained in the perceptual store). This clock-counter approach requires a switch which starts the counting process at the beginning of the to-be-measured time interval, stops the counting process at the end of that time interval, and clears the counter when the measured duration is transferred from the perceptual store to the reference memory.

JNDs in clock-counter models are related to the variance in the counting process.

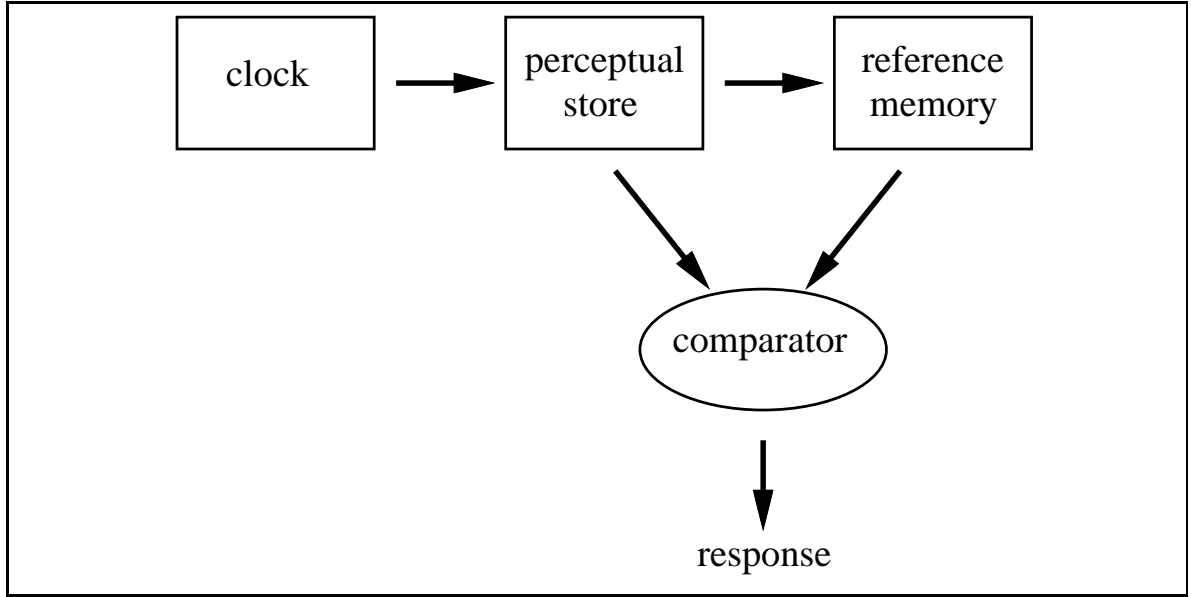


Figure 2.10: Flow diagram illustrating the clock-counter model—a modular information-processing approach to the perception of time.

Typically, the pacemaker is assumed to be a random Poisson source

$$P(n) = \frac{(\lambda T)^n}{n!} e^{-\lambda T} \quad (9)$$

in which case the variance is equal to the average number of counts λT (Creelman, 1962). Thus, for a constant mean rate λ , JNDs are proportional to T , in agreement with Weber's law. A number of studies based on this form of the clock-counter model have proposed that the mean rate (λ) is influenced by a number of stimulus factors, as well as is function of the base-rate T (Abel, 1972; Divenyi and Danner, 1977). The assumption that the mean rate λ varies as a function of the base interval T is necessary in order to explain the increase in relative JNDs for intervals less than about 300 ms, as well as to explain the over- and underestimation of short and long time intervals commonly reported in the literature.

Several connectionist clock models have also been proposed, differing from the clock-counter variety in three important ways (Church and Broadbent, 1990; Miall, 1989). (1) The single pacemaker of the clock-counter models is replaced with a set of oscillators with periods spanning a wide range of time intervals. (2) The accumulator of the clock-counter models is replaced by a distributed representation of time, based on the $+1/-1$ phase of each oscillator. Thus, a time interval T is converted into a binary $+1/-1$ vector. Notice that this oscillator-based encoding of time is identical to the process of converting 43590 seconds into the distributed representation of 12 hours, 30 minutes, and 5 minutes, except that a binary clock is used instead. (3) The memory for a time interval is encoded by a set of connection weights as opposed

to a single value (the pulse count), thus permitting more than a single time interval to be stored in memory at once, but not such that multiple stored time intervals can influence each other in ways related to the temporal structure of the pattern. In such connectionist models, the specific proposed methods for discriminating the duration of two sequentially presented time intervals vary. However, in general, duration comparisons are made by measuring the similarity of the representation of the first interval retrieved from memory and the input representation of the second interval. If this measure of similarity is less than a specified threshold, then the time change is detected. This implies that JNDs for single time intervals are determined by the selected threshold.

The main disadvantage of clock-counter models (both the connectionist version and accumulator version) is that they only concern the perception of isolated time intervals, making no predictions concerning how various contextual factors (in particular temporal context) may influence time sensitivity. By converting time into a static vector representation or a count in an accumulator, the intrinsically dynamic aspect of time perception is lost.

2.6.2 Quantal Models

Another approach to time perception has been proposed by Kristofferson (Kristofferson, 1977; Kristofferson, 1980). Kristofferson supposes that the main factor that influences time discrimination is the magnitude of a time quantum q . In the Quantal Model, the measurement of the interval T triggers the timing of an internal interval (the criterion). The variability of the criterion-estimate of T is given by a triangular probability distribution with a width equal to twice the time quantum. The standard deviation of this distribution, which Kristofferson equates with the just-noticeable difference is

$$JND = \sqrt{\left(\frac{q^2}{6}\right)}. \quad (10)$$

Thus, unlike clock-counter models, the Quantal Model predicts that JND is constant (with constant q) and independent of the the base-interval. Kristofferson (1980) reports data which suggests the doubling of q for T equal to 200, 400, and 800 ms, resulting in a step-like JND function. In terms of relative JND, each doubling of the time quantum q for multiples of a base interval T would introduce an abrupt increase in the relative JND at each multiple of T ; otherwise, relative JND is a decreasing function of T . Like clock-counter models, the Quantal Model is limited to predictions concerning isolated-interval discrimination and thus does not address the influence of temporal context on time sensitivity.

2.6.3 Multiple-Look Models

Now we turn to models which attempt to account for contextual aspects of time perception. Based on the hypothesis that listeners use the same timing mechanism

for both duration and tempo perception, a number of multiple-look models have been proposed (Drake and Botte, 1993; Schulze, 1989). In multiple-look models, each interval in a sequence provides an independent statistical observation (or “look”) of the tempo, and thus with multiple observations, the listener is able to improve the estimate of the tempo by a process of averaging the multiple-looks. For these models, the only difference between single-interval perception and tempo perception is that a single-interval task affords the listener only one “look”, whereas estimating the tempo of a sequence usually affords multiple looks. To detect a difference between the tempo of two sequence, the multiple-look model assumes that listeners use multiple-looks to abstract a tempo estimate of the standard sequence which is then stored in memory for comparison with the tempo of the comparison sequence. The multiple-look model does not, however, specify whether the tempo comparison is made directly with each interval in the comparison sequence or with a second multiple-look estimate of the comparison sequence’s tempo.

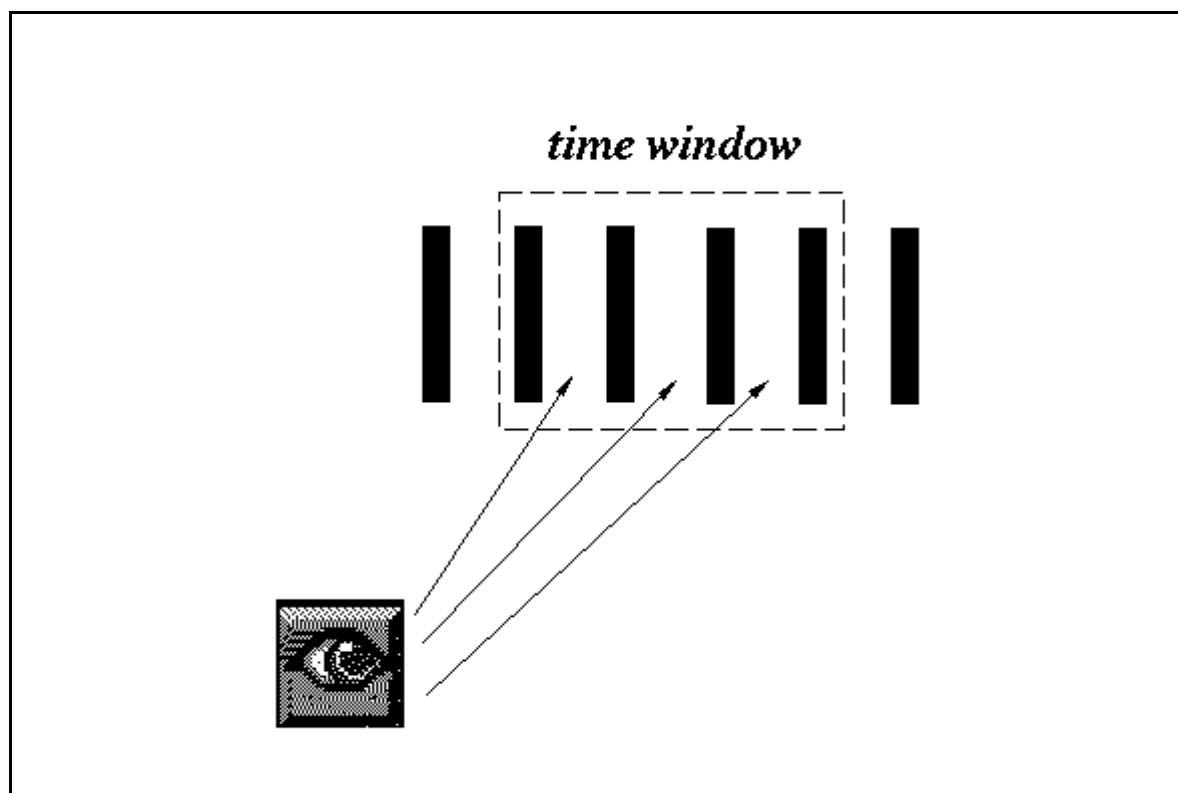


Figure 2.11: Schematic of the Drake and Botte (1993) multiple-look model.

In the Drake and Botte Multiple-Look Model (Drake and Botte, 1993), the just-noticeable difference in the tempo of a pattern (JND_n) decreases, relative to the just-noticeable difference in the duration of a single interval (JND_1), as a function

of the square-root of the number of pattern intervals, n (looks):

$$JND_n = JND_1(1/\sqrt{n}). \quad (11)$$

The above equation carries with it a couple assumptions: First, the values of JND_1 are assumed. Thus, by assuming a U-shaped relative JND curve for isolated intervals, Drake and Botte's Multiple-Look Model predicts a U-shaped relative JND curve for tempo. Second, relative JNDs for increases and decreases in tempo are assumed to be the same.

In Drake and Botte's model, the number of useful "looks" n_{\max} is limited by the duration of a temporal window d according to

$$n_{\max} = \frac{d}{IOI} \quad (12)$$

as illustrated in Figure 2.11. Thus, the number of useful looks for each tempo condition varies as a function of the IOI, with listeners being able to make use of more intervals with the faster sequences. For example, with a 1.0 second temporal window, 10 independent observations of a 100 ms IOI are possible ($n_{\max} = 10$), whereas for a 500 ms IOI, only 2 observations are possible ($n_{\max} = 2$). Consequently, Drake and Botte's Multiple-Look Model predicts that as the number of intervals in an isochronous sequence are increased, the relative JND should eventually decrease *more* for shorter IOIs than for longer IOIs, which is consistent with the qualitative distinction made between the processing of intervals < 300 ms and the processing of intervals > 300 ms, as discussed in Section 2.5.2.

In an experiment designed to evaluate the window hypothesis, Drake and Botte (1994) found evidence which suggests that window duration depends on tempo. For IOI conditions less than 300 ms, the window duration that best fit the data was found to be 1.0 second, whereas for IOI conditions greater than 300 ms, the window duration that best fit the data was found to be 2.5 seconds. This observed discontinuity in "window-size" is somewhat surprising, providing either strong additional evidence that the processing of intervals less than about 300 ms is qualitatively different from the processing of intervals greater than about 300 ms, or evidence that multiple-look models are misguided.

2.6.4 Dynamic-Attending/Contrast Models

A second class of model which addresses contextual aspects of time perception is the Contrast Model proposed by Jones and Boltz (1989). The Contrast Model is essentially one application of dynamic-attending theory (Jones, 1976) to time perception. Recall that as part of her entrainment hypothesis, Jones (1976) supposed that rhythmic patterns such as music and speech potentially entrain a hierarchically nested set of attentional periodicities, forming an attentional rhythm. And that it is the entrainment of attentional periodicities that forms the basis for the development

of expectancies for when in time, events in a pattern are likely to occur. Moreover, she suggested that entrained attentional rhythms guide the placement of attentional pulses, thus influencing the overall perception of a stimulus pattern, including the perception of the time intervals comprising that pattern.

In Jones and Boltz's Contrast Model, it is assumed that attentional periodicities (oscillators) are entrained by similar periodicities in the environment. The period of each oscillator, entrained by a series of time intervals (T_i), generates dynamic expectancies for "when" events marking future time intervals (e.g., T_{i+1}), will occur. Thus, the period of each oscillator Ω provides a continuously updated estimate of similar past and future time intervals (T). Time intervals (T) which violate the oscillator's period-based expectancies result in a temporal contrast: $(\Omega - T)$. The magnitude of such temporal contrasts are influenced by the temporal structure of the entraining pattern.

The predictions of the Contrast Model for time perception are based on the assumption that JNDs are related to temporal contrast. That is, the smaller the temporal contrast, the more sensitive listeners will be to a time change. Thus, similar to the Multiple-Look Model, the Contrast Model predicts that increasing the number of isochronous intervals in a pattern should improve listeners' abilities to detect tempo changes, since more isochronous intervals in a pattern improves the entrainment of the tracking oscillator, reducing temporal contrast. Moreover, the Contrast Model predicts that listeners' abilities to detect time and tempo changes should be better for metrical (regularly timed) sequences (of which isochronous sequences are one example) than for nonmetrical (irregularly timed) sequences, since temporal contrasts should be smaller for regular sequences than for irregular sequences.

2.6.5 Discussion

These above two *general* predictions of the Contrast Model are consistent with much of the embedded-interval and tempo data reviewed in Section 2.5.2. However, some *specific* aspects of these data are problematic for the Contrast Model, as well as for Drake and Botte's Multiple-Look model (discussed in Section 2.6.3). To close this chapter, I will summarize these problematic data, as well as some of the associated limitations of the contrast and multiple-look models. These data, which are a primary focus of the modeling efforts reported in this thesis, will be addressed again in Chapter's 4 and 5.

First, it was discussed earlier that listeners' sensitivity to the tempo of isochronous sequences improves more, with increasing number of intervals, for IOIs less than about 300 ms than for IOIs greater than about 300 ms. This differential improvement is substantial, resulting in an extension of the optimal zone of sensitivity to a shorter IOI (Drake and Botte, 1993; Michon, 1964). The Contrast Model only predicts that sensitivity should be heightened with reduced temporal contrast, not that it should be heightened differentially as a function of the base IOI. Many researchers

have suggested that this potentially heightened sensitivity implies that the processing of intervals less than about 300 ms is qualitatively different from the processing of intervals greater than 300 ms, possibly involving distinct mechanisms (Drake and Botte, 1993; Michon, 1964; Hirsh et al., 1990; ten Hoopen et al., 1994).

Second, in studies further investigating the influence of the number of isochronous intervals on tempo sensitivity (Drake and Botte, 1994), it was reported that optimal sequence duration, determined by the maximum number of intervals which lower the relative JND multiplied by the base IOI, is constant at about 1.0 second for IOIs less than about 300 ms, but shifts abruptly to about 2.5 seconds for IOIs greater than about 300 ms. This abrupt shift is not explained by the Contrast Model. In order for the Multiple-Look Model to explain this shift, it must assume two distinct temporal-window sizes within which listeners integrate timing information. Drake and Botte (1994) suggest that this result provides strong additional evidence that short IOIs (< 300 ms) are processed differently than longer IOIs (> 300 ms).

Third, a more fundamental issue is that both the Contrast Model and the Multiple-Look model implicitly assume a U-shaped relative JND curve for isolated-interval discrimination, in order to be able to predict a similar pattern of relative JNDs for embedded-interval and tempo discrimination. Thus, both the Contrast Model and the Multiple-Look model address the *relationship* between isolated-interval discrimination data and the embedded-interval and tempo discrimination data, but not the *process* of isolated-interval discrimination. Conversely, both the clock-counter models and the quantal models addressed the discrimination of isolated intervals, but not the discrimination of tempo or embedded intervals.

Finally, the oscillators in the Contrast Model are abstract oscillators. Consequently, the predictions of the Contrast Model are not yet linked to the dynamics of a specific *functioning* oscillator and are thus necessarily limited in specificity. In comparison, the predictions of the Entrainment Model of time perception developed in Chapter 4 are tied to the functioning of a specific oscillator. Before presenting the Entrainment Model, it is necessary to first investigate the oscillator on which the Entrainment Model is based: the Adaptive Oscillator.

Chapter 3

Adaptive Oscillators

3.1 Overview

Strogatz and Stewart (1993) relate an interesting anecdote concerning the origin of the concept of the “coupling” of oscillators. In 1665, the Dutch physicist Christiaan Huygens, observed from his bedside during an illness that two clocks hanging side by side were synchronized in the motion of their pendulums. For hours, he watched the motion of the two pendulums, yet they remained in perfect synchrony. When either or both of the pendulums were disturbed, synchrony was regained within a half hour. Huygens hypothesized that the two clocks must be influencing each other, perhaps through vibrations in their common support. To test this, he moved them to opposite walls of the room, and sure enough the pendulums fell out of lock-step, one clock losing several seconds a day relative to the other (Strogatz and Stewart, 1993). Huygens’ chance observation, launched an entire new branch of mathematics: the theory of coupled oscillators. The term *coupled* means that the oscillators are able to perturb each other in some way, as did the pendulum clocks on Huygens’ bedroom wall.

In this chapter, I introduce a class of adaptive-oscillator mechanisms, based on the coupled oscillator paradigm, for establishing and maintaining a single metrical level of a rhythmic pattern in spite of variability in the timing of that pattern. Section 3.2 provides a rudimentary introduction to the theory of coupled oscillators. Next, in Section 3.3, a class of adaptive oscillators is specified. There is an important distinction between the coupled clocks on Huygen’s bedroom wall and adaptive oscillators. Coupled clocks become entrained only by perturbing each others’ phases. Coupled adaptive oscillators are entrained by rhythmic input patterns by adjusting both phase and period. The additional adjustment of the oscillator’s period can be thought of as the internalization of an expectancy for the occurrence of future inputs with close to the same periodicity. In the case of “missing” inputs, as well as when the rhythmic pattern stops, the adaptive oscillator continues to predict future inputs. In essence, the adaptive oscillator internalizes a beat, retaining a memory of that beat after the

pattern stops. Section 3.4 examines the dynamics of this class of adaptive oscillators. The proposed oscillators are then compared with an oscillator model for meter perception recently proposed by Large and Kolen (1995). This chapter concludes with an evaluation of the proposed class of adaptive oscillators to be entrained by the Povel and Essens (1985) set of temporal patterns, varying in rhythmic complexity. The adaptive oscillator's ability to be entrained by these patterns is compared with listeners' abilities to reproduce these patterns by tapping.

3.2 Mathematical Background

To understand how coupled oscillators interact, it is first important to understand the mechanics of a single isolated oscillator. An oscillator is a system that generates periodic behavior. Formally, a function $f(t)$ is periodic if and only if there exists a real number Ω such that $f(t+n\Omega) = f(t)$ for all integers n . Ω is the period of the function $f(t)$ and $1/\Omega$ is the frequency or rate of oscillation. This definition of a periodic function, obtained from a standard calculus book, states that each value of a periodic function must precisely repeat every Ω time units. Unlike the textbook definition of periodic, biological oscillations such as the gait of a running animal, the firing of a pacemaker neuron in the heart, and the beat established by a musician exhibit variability on each cycle. In terms of a formal description of biological oscillations, we can view biological oscillations as being driven by an oscillator with a period Ω (in the strict sense) that changes over time (i.e., biological oscillations are driven by a non-stationary oscillator).

Instead of representing an oscillator's periodic motion as a time series, it can be represented as a phase-portrait which combines position and velocity to show the entire range of possible states; that is, the phase-space of the system (see Figure 3.1). Phase ϕ jointly describes oscillator position (within its cycle) and velocity, as a fraction of the oscillator's cycle: $t/\Omega \bmod 1$ (Winfree, 1980). All systems that generate periodic behavior, no matter how complex, will eventually traverse a closed curve in phase-space.

In phase-space, the motion of an ideal pendulum smoothly follows the circle's circumference in a clock-wise fashion. Zero-phase is arbitrarily located at 3 o'clock (zero velocity and maximum positive displacement). Likewise, there are many different conventions for describing an oscillator's motion in phase-space. Phase is often measured in degrees ($[0, 360^\circ]$), radians ($[0, 2\pi]$ or $[-\pi, \pi]$), and, informally, as the hours, minutes, or seconds of a clock.

In this thesis, two different representations will be useful for describing oscillator motion in phase-space. Both will reflect clock-wise movement on a unit circle with zero-phase (the beginning of the oscillator's cycle) positioned at 3 o'clock. For the first representation, ϕ will vary from 0.0 to 1.0 (see Figure 3.2A). This representation will be used for most of the mathematical analyses. For the second representation, ϕ will vary from -0.5 to 0.5 , making it possible to describe negative (early) and positive

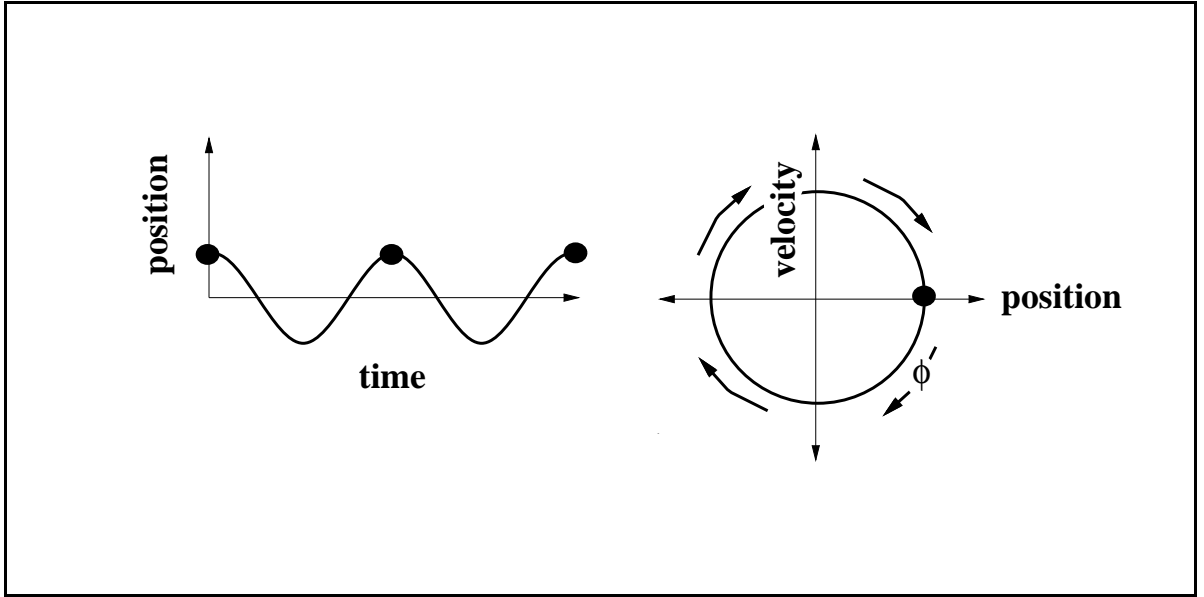


Figure 3.1: An oscillator's periodic motion can be represented as a time-series (shown on the left), or as a phase-portrait which combines both position and velocity to show the entire range of possible states (shown on the right). Each phase ϕ specifies a fraction of the oscillator's cycle.

(late) phase relative to the onset of the oscillator's cycle (see Figure 3.2B). The second representation will be useful for describing the learning equations of the model and for discussion the formulation of proposed the Entrainment Model of human time perception discussed in Chapter 4.

Although a single oscillator traverses a simple loop in phase-space (as in Figure 3.2), the motion of two or more coupled oscillators is much more complex, and in some cases the equations describing the system are intractable. In a system of coupled oscillators, the oscillators may interact with only their immediate neighbors or with all others; coupling may be unidirectional or bidirectional; and interactions may occur only at discrete points in time (pulse-coupling) or be continuous (Glass and Mackey, 1988).

The simplest case of coupled oscillation is a single driving (input) oscillator unidirectionally coupled to a driven (output) oscillator. Geometrically, the combined motion of the two oscillators in phase-space can be described as the trajectory followed on the surface of a torus (Abraham and Shaw, 1992) (see Figure 3.3). In the example shown, the torus represents the Cartesian product of two circles C_1 and C_2 . The set of all points on the surface of the torus represents all possible states of the system. Each point with coordinates (ϕ, θ) specifies the phases of the two oscillators. Suppose that θ is the phase of the input oscillator and that ϕ is the phase of the output oscillator. To investigate the long-term effect of coupling, we take snapshots of the input phase θ , every time the output oscillator reaches zero-phase, beginning

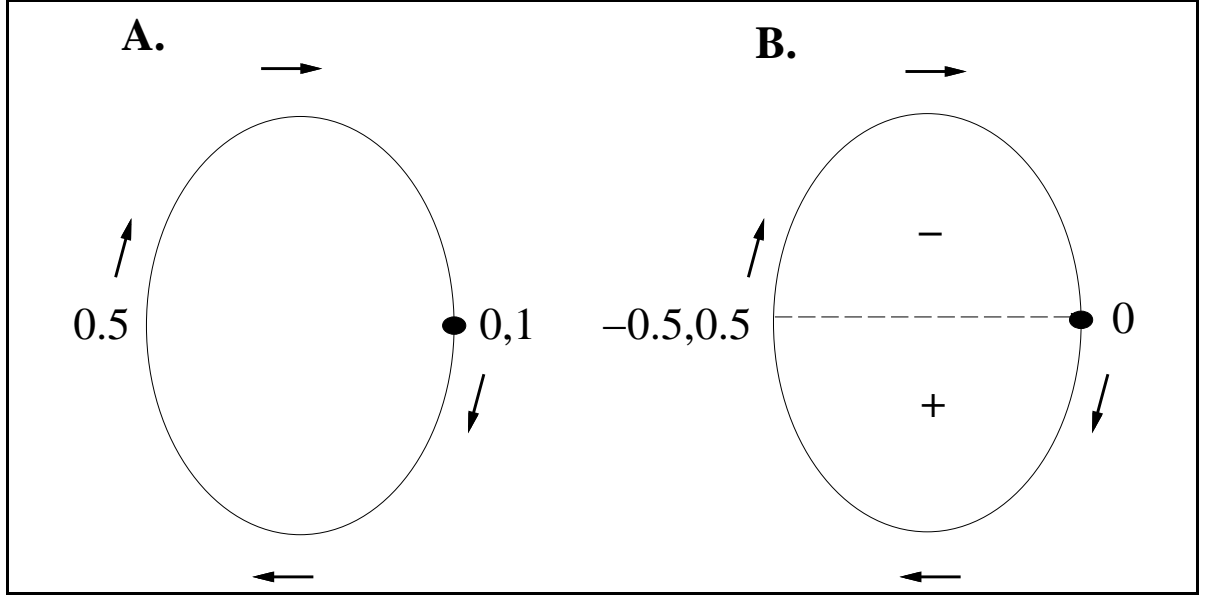


Figure 3.2: Phase is represented in two different ways in this thesis, either as $\phi \in [0, 1]$, or as $\phi \in [-0.5, 0.5]$.

its cycle:

$$\theta_{i+1} = (\theta_i + \frac{\Omega}{T} + g(b, \theta_i)) \bmod 1. \quad (13)$$

In this iterative equation, called a Poincaré map, i is the i th firing of the output oscillator with period Ω , θ_i is the phase of the input oscillator with period T relative to the i th strobed output cycle, and the function $g(b, \theta_i)$ specifies the influence of the driving (input) oscillator on the driven (output) oscillator. The parameter b is the coupling strength.

The attractor dynamics of the output oscillator are described by the limiting behavior of the iterated Poincaré map. The rotation (or winding) number (ρ)

$$\rho = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \Delta \theta_i \quad (14)$$

specifies the number of times the output oscillator winds a path around the torus for every cycle of the input (Winfree, 1980; Glass and Mackey, 1988). For rational ρ , p/q , the system is *mode-locked* and rests in a periodic attractor that reflects $p : q$ entrainment: p input cycles for every q output cycles. In this case, the trajectory of the coupled system in phase-space forms a closed path on the surface of the torus; that is, the two oscillators are synchronized. The process by which they become synchronized is called entrainment.

Consider a system in which $T = 2.0$, $\Omega = 1.0$ and the coupling strength b is zero. We can look at the behavior of the iterated Poincaré map by slicing the torus along

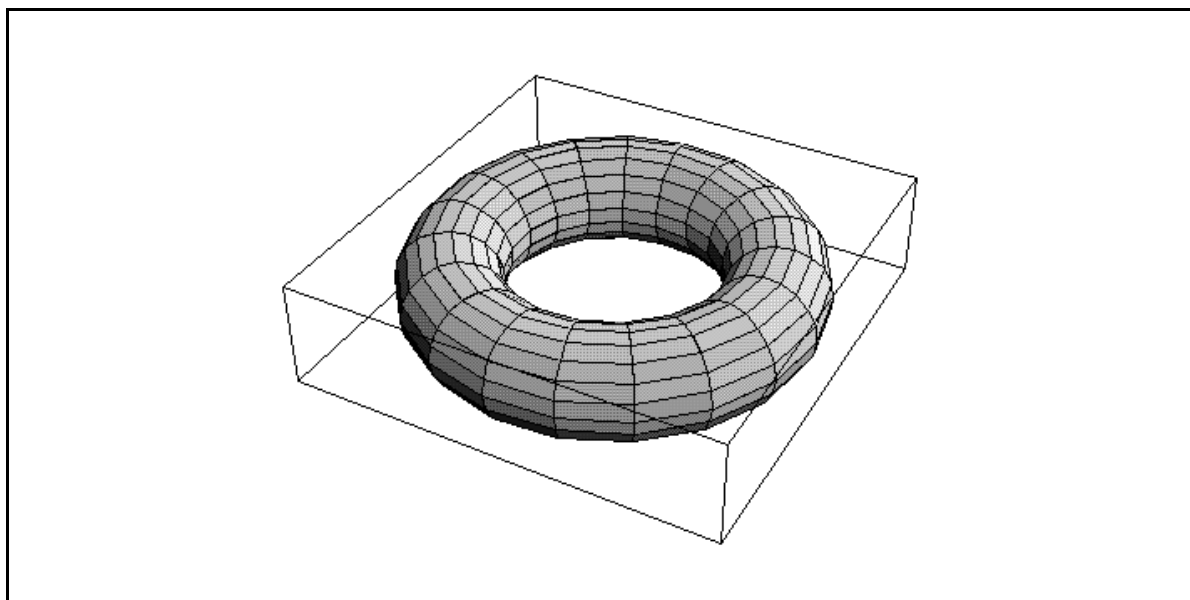


Figure 3.3: Geometric representation of the phase-space of two oscillators. The motion of two oscillators is described by a trajectory on the surface of the torus. Each point with coordinates (ϕ, θ) specifies the phases of the two oscillators.

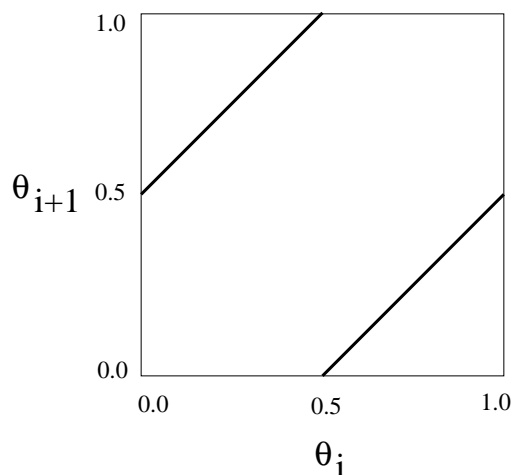


Figure 3.4: Two-dimensional projection of the surface of the torus (phase-return map) showing 1 : 2 entrainment between two oscillators.

both its horizontal and vertical dimensions, and then unfolding it into a sheet (as shown in Figure 3.4). This two-dimensional projection of the system's trajectory (or *phase-return map* (Winfree, 1980; Glass and Mackey, 1988)) shows the relationship between the phase of the input oscillator at the beginning of the i th output cycle and the phase of the input oscillator at the beginning of the $(i + 1)$ th output cycle (i.e.,

the change in the phase of the input during a complete output cycle). For example, if the input phase is 0.0 at the beginning of output cycle i , then it will be 0.5 at the beginning of output cycle $i + 1$. For all initial input phases θ_i , the input phase at the beginning of the next output cycle θ_{i+1} is a (mod1) phase change of 0.5. Thus, the input oscillator cycles once for every two cycles of the output oscillator (a 1:2 entrainment ratio or a winding number of $\rho = 0.5$).

For irrational ρ , the phase-space trajectory of the coupled oscillator system does not form a closed loop. Instead, the phases of the two oscillators drift slightly on each cycle, never precisely returning to their initial values. This type of behavior is called *quasi-periodicity*. Whether a system of two coupled oscillators exhibits quasiperiodic behavior or mode-locking depends on the initial ratio of output to input periods $\frac{\Omega}{T}$ (also called the bare winding number ρ_0), the coupling strength b , as well as the specific form of the Poincaré map. As the coupling strength is increased, ρ_0 regions which initially produced quasiperiodic behavior can be replaced by regions of overlapping mode-locking. The presence of overlapping mode-locking regions indicates that the system is sensitive to the initial phase conditions; that is, the specific mode-lock (or entrainment ratio) that the system attains depends on the initial phase settings of the oscillators. The mode-locking behavior of coupled oscillators is often highly structured. One way to examine this structure is to construct what is called an *Arnold map* (Arnold, 1983). The Arnold map shows specifically how mode-locking behavior varies as a function of the coupling strength and the bare winding number ρ_0 . Consider the specific Poincaré map (or circle map)

$$\theta_{i+1} = \theta_i + \rho_0 + b \sin(2\pi\theta_i) \quad (15)$$

which describes unidirectionally coupling between a single (driving) input oscillator and a single (driven) output oscillator. This particular circle map is called the sine circle map because of the sinusoidal coupling term: $b \sin(2\pi\theta_i)$. The parameter b specifies the coupling strength of the system.

The Arnold map for the sine circle map is shown in Figure 3.5. In this figure, the bare-winding number is represented on the x -axis and coupling strength is represented on the y -axis. Each point indicates that the system obtains a stable entrainment ratio for the specified combination of bare-winding number and coupling strength. Whether or not the system enters a stable entrainment ratio is determined for each point by using a large n to approximate the limit in Equation 14. If the obtained value is within a predetermined tolerance of any of a selected set of the entrainment ratios, the point is included in the graph. Each “tongue” shaped region in the Arnold map corresponds to a different entrainment ratio (or periodic attractor). The width of each tongue reflects the stability of the attractor (i.e, how sensitive the system is to perturbations in that region of parameter space). Only mode-locking regions corresponding to the selected set of entrainment ratios are shown, although mode-locking regions exist for all rational entrainment ratios.

The structure of the Arnold diagram is not arbitrary, but is determined by an

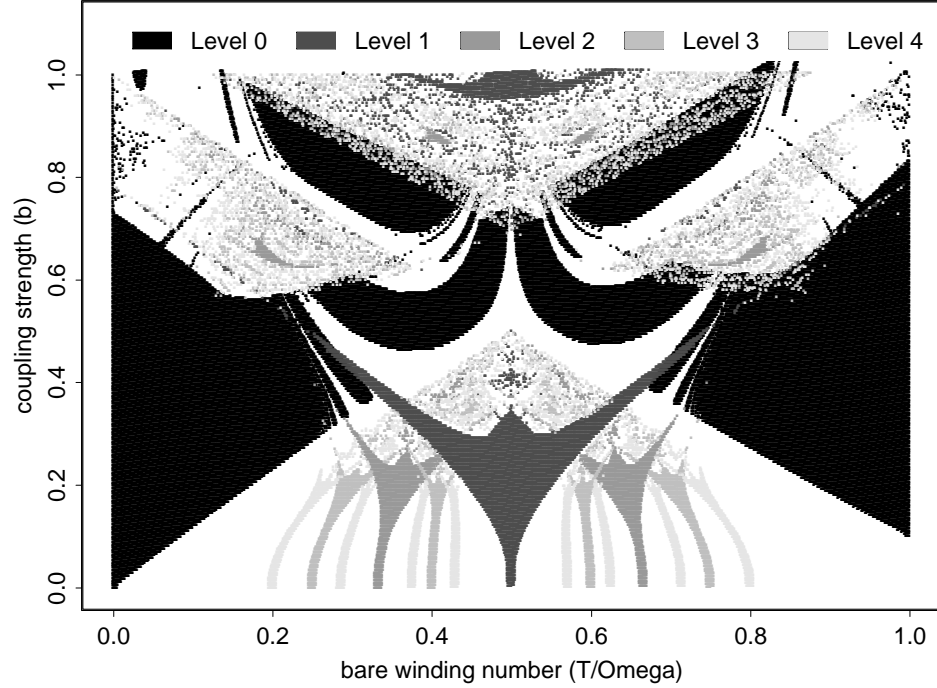


Figure 3.5: Arnold map specifying mode-locking for iterated sine circle map for first four levels of the Farey Tree. Each tongue-shaped region corresponds to a different entrainment ratio as specified by the level.

elegant construction called the *Farey series* (Shroeder, 1991), which can be used to enumerate all entrainment ratios in order of their stability. Given two entrainment ratios $p_1 : q_1$ and $p_2 : q_2$ as parents, the next less stable “offspring” ratio that lies between the two tongues is specified by

$$p_1 + p_2 : q_1 + q_2. \quad (16)$$

The tree structure of the Farey series is shown in Figure 3.6. On each level, all entrainment ratios share the same stability. As one branches deeper in the tree, the ratios become less stable. For example, $1 : 1$ entrainment is more stable than $1 : 2$ entrainment, and so on. In addition, any two adjacent ratios $p_1 : q_1$ and $p_2 : q_2$ satisfy the relation

$$|p_1 q_2 + q_1 p_2| = 1. \quad (17)$$

This property, called unimodularity, reflects the series of bifurcations observed with coupled oscillators and has been used to model/predict patterns of bifurcations between stable entrainment ratios in both biology and psychology (Winfree, 1980; Glass and Mackey, 1988; Treffner and Turvey, 1993).

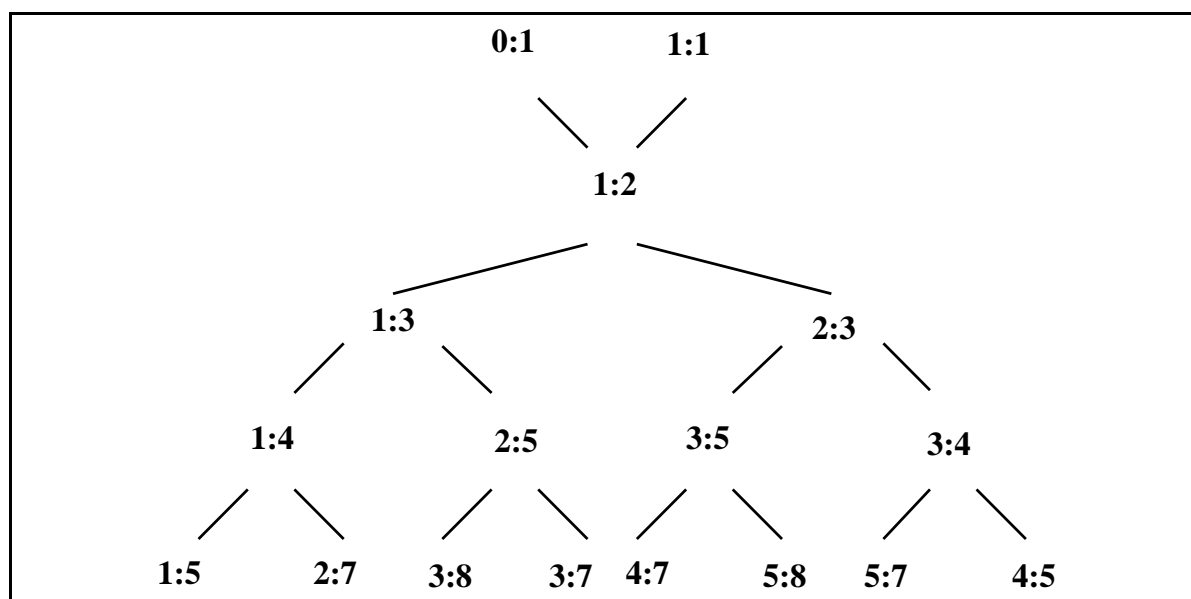


Figure 3.6: Levels 0 through 4 of the Farey tree structure that describes the relative stability of different entrainment regions for coupled oscillators. Entrainment ratios 0:1 and 1:1 comprise Level 0.

Summary This section introduced the mathematics of coupled oscillators. We examined the behavior of a simple coupled system, the sine circle map, in which a single driving (input) oscillator is uni-directionally coupled to a driven (output) oscillator. The dynamics of this system were specified by constructing an Arnold map showing parameter regions corresponding to stable periodic attractors.

Simple systems of coupled oscillators can be shown to exhibit very complex dynamical behavior, including mode-locking, quasi-periodicity and chaos, by varying both coupling strength and the ratio of the oscillator's periods. Stable periodic attractors exist for all rational entrainment ratios and their relative stability is described by an elegant construction called the Farey series.

3.3 The Adaptive Oscillator

The class of adaptive oscillators that I specify in this section have been developed over several years (McAuley, 1993; McAuley, 1994a; McAuley, 1994b) and share five primary properties. (1) Each oscillator has a resting value for its period that it gradually returns to in the absence of input. (2) Each oscillator has a periodic activation function. (3) Each oscillator is phase-coupled with the input by phase-resetting. (4) Each oscillator retains a memory of the phases at which previous phase-resets occurred, which is used as the output of the oscillator, as well as a measure of how well entrained it is. (5) Each oscillator's output is used as feedback to modify its period,

which in turn serves to align the beginning of each oscillator's cycle with future inputs. The output of each oscillator can also be used to modify the shape of the activation function, so that, as the oscillator becomes entrained by a rhythmic pattern, the time window for expected future inputs narrows.

This phase-resetting adaptive oscillator is based on a simple sinusoidal activation function given by

$$a(t) = (1 + \cos(\frac{2\pi t}{\Omega(t)}))/2 \quad (18)$$

where $\Omega(t)$ is the oscillator's period, initially equal to its resting value:

$$\Omega(0) = \bar{\Omega}. \quad (19)$$

As outlined above, each oscillator *adapts* by adjusting the phase and period of its activation function in response to perturbations from discrete inputs. Inputs on $[0, 1]$ represent event onsets with variable intensity values. This assumption is consistent with the view that rhythmic organization is primarily determined by the temporal pattern of event-onsets (Handel, 1993). Thus, rhythmic patterns are represented as patterns (streams) of pulses. Figure 3.7A shows the input representation for a four-tone isochronous pattern with a 300 ms IOI. Formally, this pattern is represented as

$$i(t) = \begin{cases} I & \text{if } t = nT \\ 0.0 & \text{otherwise} \end{cases} \quad (20)$$

where T is the 300 ms IOI and I is the intensity of the n th input pulse.

3.3.1 Phase Resetting

The adaptive oscillator uses a discrete form of phase coupling called phase-resetting. In the phase-resetting model, each weighted input $w_i i(t)$ (the coupling strength) is added to the activation of the oscillator, providing a measure of total activation. The total activation of the system is then compared with a threshold of 1.0. If the total activation exceeds this threshold, the oscillator resets its phase to zero, beginning its cycle again. Formally, phase-resetting is defined by the following piecewise function

$$\phi(t) = \begin{cases} 0 & \text{if } a(t) + w_i i(t) > 1.0 \\ \phi(t) & \text{otherwise.} \end{cases} \quad (21)$$

Figure 3.7 shows the behavior of a phase-resetting model. In the figure, an oscillator with a 500 ms period resets its phase in response to an isochronous pattern with a 300 ms IOI. Since the input pattern sequence is isochronous all of the phase resets occur at the same phase. In this example, the phase resets occur at $\phi = -0.4$.

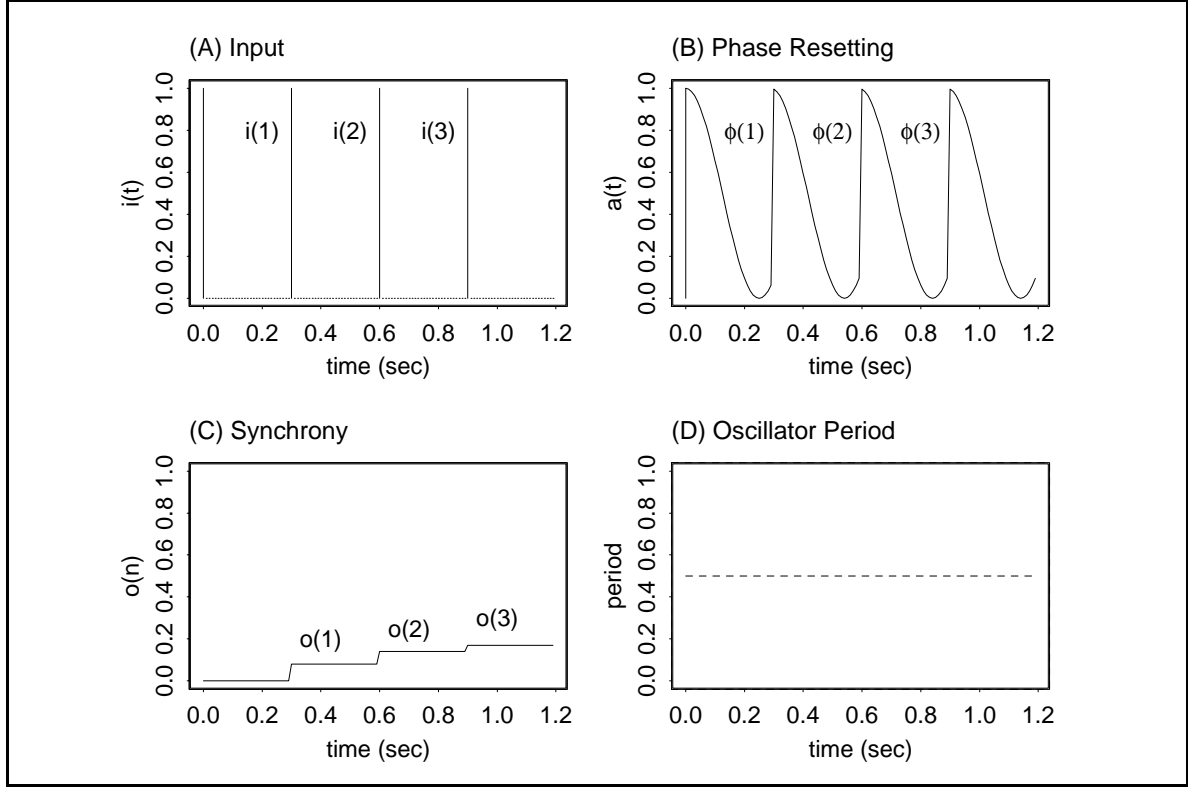


Figure 3.7: Phase-resetting of an oscillator with a 500 ms period in response to a rhythmic pattern with a fixed 300 ms inter-onset interval.

3.3.2 Phase Memory

The degree of synchronization with a rhythmic pattern can be measured by maintaining a memory of the phase at which the input events force the oscillator to reset (i.e., when the total activation exceeds threshold). The symbol $\phi^r(n)$ is used to indicate the phase of the n th input pulse $i^r(n)$, that forces the total activation above threshold. The superscript r indicates that the phase and associated input correspond to a phase-reset. Thus, the degree of synchronization is measured by maintaining a smoothed memory of $\phi^r(n)$, defined here as the output of the oscillator $o(n)$:

$$o(n) = (1 - w_o)o(n - 1) + w_o(1 - 2|\phi^r(n)|). \quad (22)$$

The parameter w_o establishes a weighting between the current reset phase $\phi^r(n)$ and the memory of the previous reset phases $o(n - 1)$. The output varies between 0 for poor synchronization and 1 for perfect synchronization.

Figure 3.8 shows the output sequence for ten inputs coinciding with beginning of the oscillator's cycle ($\phi^r(n) = 0.0$) followed by ten inputs that are 180 degrees out-of-phase with the oscillator (for five different values of w_o). For all w_o except $w_o = 0.0$, the output sequence converges towards 1.0 for the first ten synchronous inputs. For

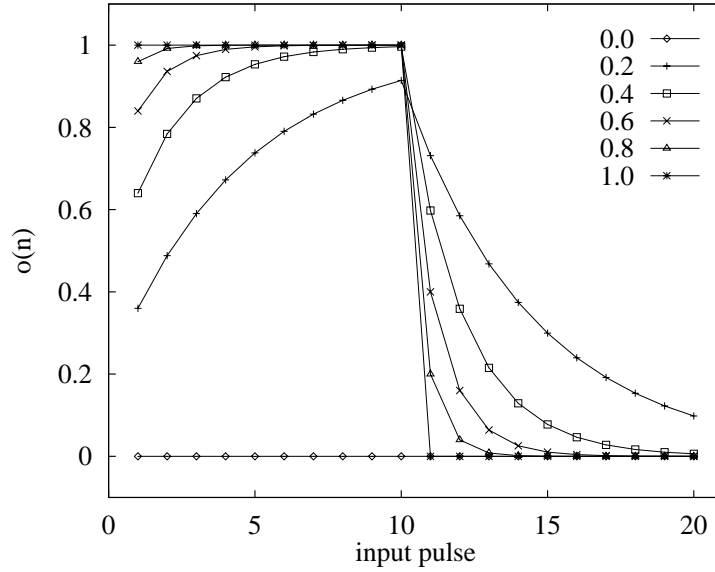


Figure 3.8: Oscillator output $o(n)$ measures synchronization using a memory of phases, for which input pulses force the total activation above threshold. The smoothing parameter w_o weights the n th such phase $\phi^r(n)$ relative to the memory $o(n-1)$ of previous $n-1$ reset phases. This figure shows the output for 10 successive synchronized pulses followed by 10 pulses 180 degrees out of phase, for a range of w_o values.

the second ten “out-of-phase” inputs, the output sequence decays towards 0.0. The effect of decreasing w_o is to smooth the output response. For $w_o = 1$, each output $o(n)$ depends only on the current reset phase $\phi^r(n)$ (i.e., the synchronization of the oscillator is measured by only a single input). In the figure, the $w_o = 1.0$ output sequence attains a value of 1.0 after the first synchronous input. Unless the following inputs are isochronous, as they are for the next nine inputs in the example shown in the figure, a single input is not enough information to provide an adequate measure of synchronization. For example, suppose that a spurious input in an otherwise periodic sequence, just by chance, coincides with the beginning of the oscillator’s cycle, in which case ($\phi^r = 0$). Although the output is 1.0, the oscillator is not in synchrony with the input sequence. Thus, unless the oscillator’s input environment is perfectly regular, $w_o = 1$ does not produce outputs which accurately reflect entrainment.

The opposite extreme is $w_o = 0$, in which case $\phi^r(n)$ is ignored and the model relies on its phase memory to measure synchronization. Since the output equation is recursive, $w_o = 0$ implies that the output $o(n)$ is fixed at its initial value $o(0)$ or at the attained output value when w_o was set to 0. Obviously, any measure of synchronization that ignores the phase of each new input, and relies instead on an initial value, is not adequate.

What is an appropriate choice for w_o ? If the input events are isochronous as in the

example (Figure 3.8) then a single input *does* provide sufficient information to measure synchronization, and $w_o = 1.0$ is appropriate. However, if the timing of the inputs is extremely variable, then the best strategy is to rely on the oscillator's phase-memory, as each new input provides very little useful information about synchronization. As an extreme example, suppose that the stream of inputs changes from a rhythmic pattern of time-intervals to a random pattern of time-intervals, then even switching to $w_o = 0.0$ might be appropriate, if the input pattern eventually returned to its prior rhythmicity. Thus, the appropriate choice of w_o depends on the temporal structure of the input pattern.

3.3.3 Activation Sharpening

The output of the adaptive oscillator can be used as feedback to modulate the shape of the activation function according to

$$a(t) = [(1 + \cos(\frac{2\pi t}{\Omega(t)}))/2]^{(1-o(n))\gamma_{\min} + o(n)\gamma_{\max}}, \quad (23)$$

in which output scales the exponent of the activation function between γ_{\min} and γ_{\max} . Figure 3.9 graphs this modified activation equation for outputs of 0.0, 0.3, 0.7, and 1.0, for $\gamma_{\min} = 1.0$ and $\gamma_{\max} = 10.0$. Modulating the “sharpness” of the activation function alters the temporal window within which inputs are able to perturb the phase and period of the oscillator. As the oscillator is entrained by the input pattern the output increases, this temporal window narrows. Thus, the temporal expectancy for the next input is more selective.

3.3.4 Period Coupling

Two properties characterize period coupling in an adaptive oscillator: (1) the adaptive oscillator uses its output $o(n)$ (a measure of synchrony) as a “teaching” signal to determine how much to adjust its period; and (2) the sign of the phase $\phi^r(n)$ (the reset phase of the n th input event) is used to determine the direction of the period change (increase or decrease). Both properties are expressed by the period coupling term

$$\mathcal{P} = \phi^r(n)(1 - o(n)). \quad (24)$$

This choice of \mathcal{P} guarantees that the oscillator does *not* adjust its period when either $o(n) = 1.0$ (i.e., it attains a perfect measure of synchrony) or $\phi^r(n) = 0.0$ (i.e., the current input coincides with the beginning of the oscillator's cycle). For negative reset phases, the input is assumed to be “early” with respect to the beginning of the oscillator's cycle and the oscillator's intrinsic period is shortened, speeding the oscillator up. Conversely, for positive reset phases, the input is assumed to be “late” with respect to the beginning of the oscillator's cycle and the oscillator's intrinsic period is lengthened, slowing the oscillator down. The amount of period adjustment

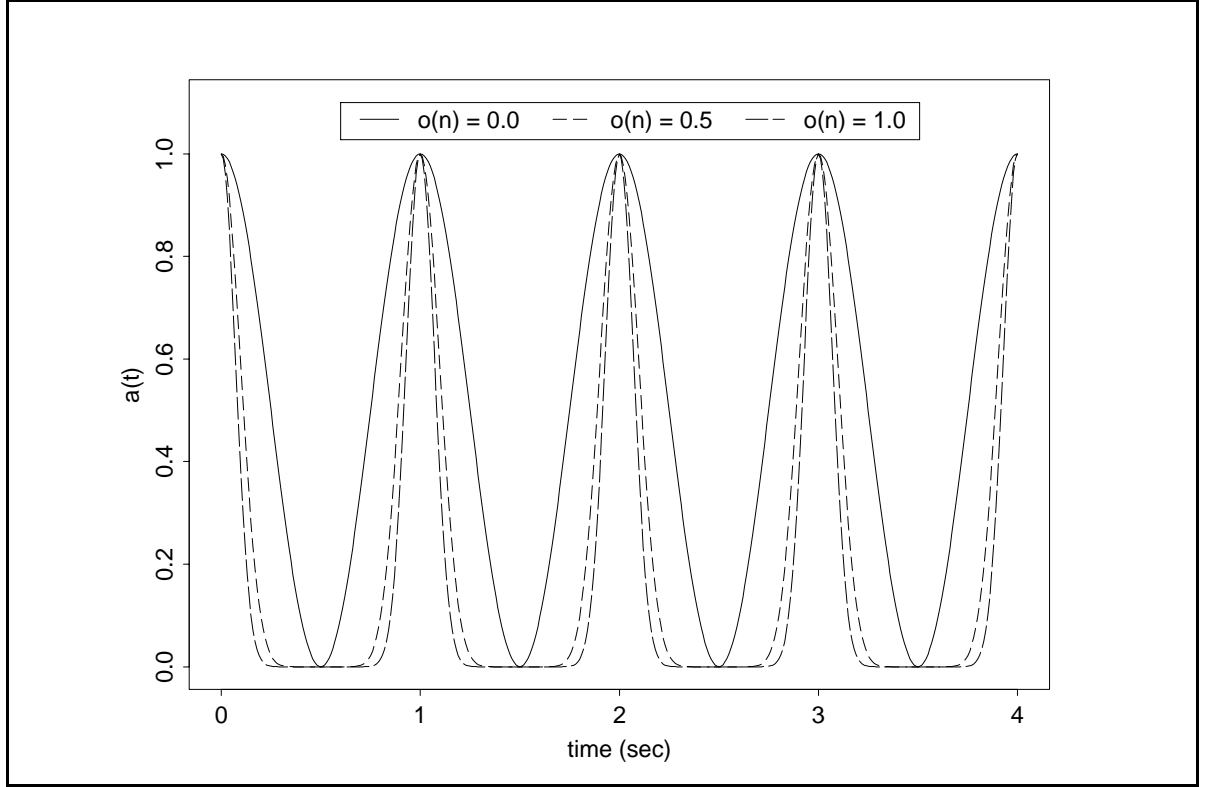


Figure 3.9: Feedback control of the “sharpness” of the activation function for outputs of 0.0, 0.3, 0.7, and 1.0 and for $\gamma_{\min} = 1.0$ and $\gamma_{\max} = 10.0$.

is inversely related to the output. If the output is large, the oscillator only adjusts its period by a small amount. If the output is small, the oscillator makes a much larger change in its period, enabling it to search the space of possible entrainment possibilities.

A complete description of period coupling in the adaptive oscillator is given by

$$\frac{\delta\Omega}{\delta t} = \alpha \frac{\Omega}{2} \mathcal{P}[\phi^r(n), o(n)] \mathcal{M}[i^r(n)] - \beta(\Omega - \bar{\Omega})(1 - \mathcal{M}[i^r(n)]). \quad (25)$$

In this differential equation, the period coupling term is scaled by $\frac{\Omega}{2}$ (to ensure that the period changes by at most one half of the oscillator’s cycle), by an input impulse response function \mathcal{M} (in order to spread the change in the oscillator’s period over the entire cycle), and by the entrainment rate α . The impulse response function is given by

$$\mathcal{M} = \frac{1}{1 + e^{-\Gamma(i^r(n)e^{-\Theta t} - 0.5)}}. \quad (26)$$

where Θ and Γ are the impulse response bias and gain respectively. A decay term

$$\Omega - \bar{\Omega}, \quad (27)$$

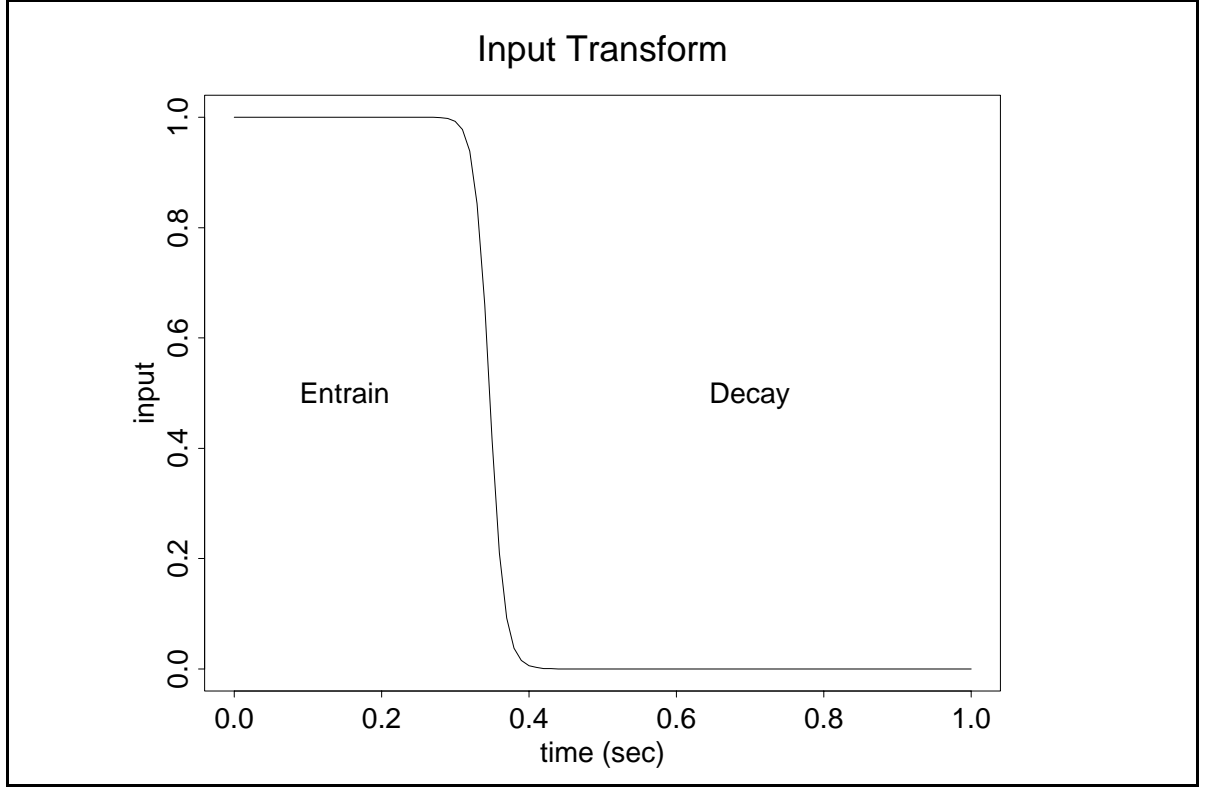


Figure 3.10: Input impulse-response function for $\gamma = 1000.0$ and $\Theta = 2.0$.

scaled by

$$(1 - \mathcal{M}[i^r(n)]) \quad (28)$$

and a decay rate β is included in Equation 25, so that in the absence of input, a penalty is incurred for large differences between the adapted and intrinsic periods, and the oscillator will gradually return to its resting rate.

Figure 3.10 shows the input impulse response for an input of 1.0, for $\gamma = 1000$ and $\Theta = 2.0$. The impulse response function extends the change in the oscillator's period in time following a phase-reset. The value of the impulse response function determines the weighting of the entrainment and decay processes: for an impulse response of 1.0, the oscillator is entrained by the input, for an impulse response of 0.0, the oscillator decays back to its resting period, for intermediate impulse response values, there is a mixture of entrainment and decay. For the specific parameterization of the impulse response function shown in the figure, the response function maintains a value close to 1.0 for approximately 0.4 seconds and then drops off sharply towards 0.0. Thus, following the phase-reset, the oscillator incrementally adapts its period in response to the input for approximately 0.4 seconds; after which, the period of the oscillator gradually returns to its resting rate, until this decay process is preempted by the next input. If the next input occurs, before the impulse response function

decreases towards 0.0 (before 0.4 seconds in the example), then there is no effectively no period decay.

If the impulse response gain $(,)$ is assumed to be a very large number, then the impulse response function is essentially a descending step function. In which case, Equation 26 reduces to

$$\mathcal{M} = \begin{cases} 1 & \text{if } e^{-\Theta t} < 0.5 \\ 0 & \text{if } e^{-\Theta t} > 0.5. \end{cases} \quad (29)$$

Thus, solving $e^{-\Theta t} = 0.5$ for t specifies the processing boundary between entrainment and decay, which is given by

$$IOI_{\Theta} = \left\lceil \frac{\ln 0.5}{\Theta} \right\rceil. \quad (30)$$

I term this processing boundary the inter-onset-interval threshold (IOI_{Θ}).

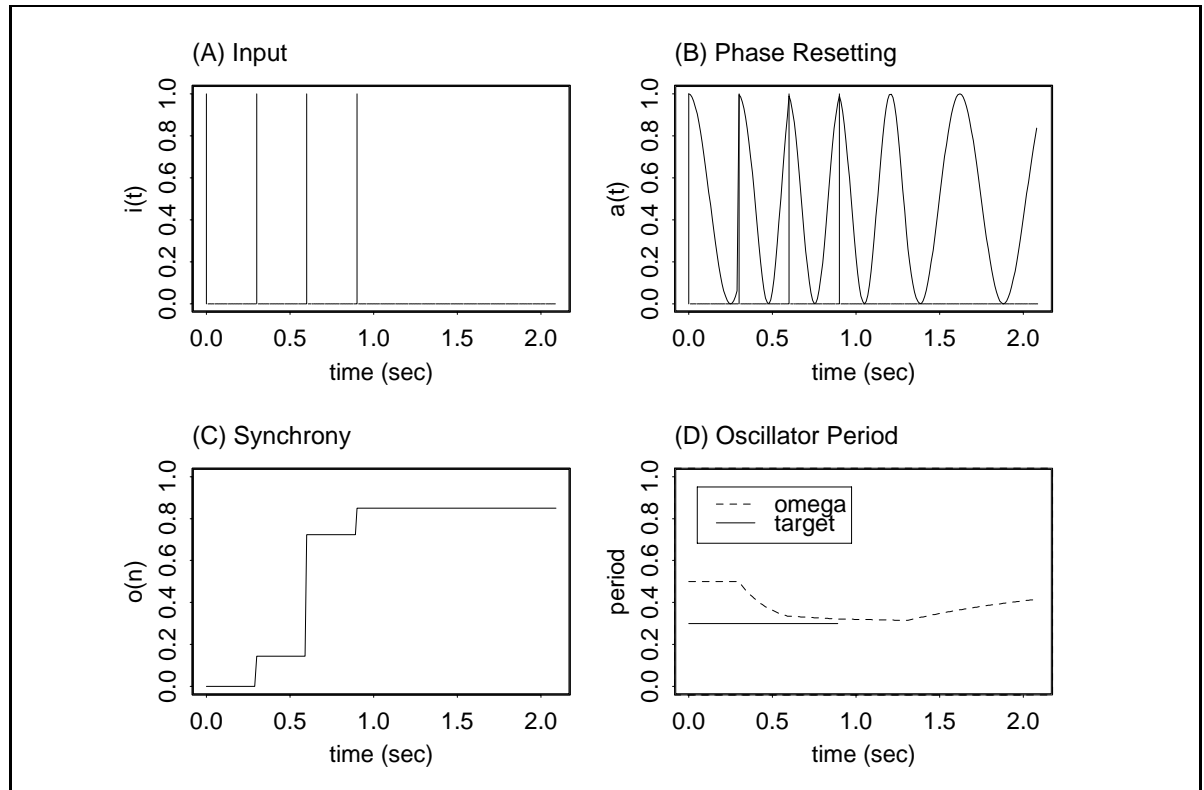


Figure 3.11: Phase-resetting and period adaptation oscillator with 500 ms intrinsic period, in response to isochronous input pulses (300 ms inter-onset-interval). (A) Input signal. (B) Phase-resetting with period adaptation. (C) Output signal. (D) Period entrainment and decay.

Figure 3.11 shows the response of an adaptive oscillator with both phase resetting and period coupling to an isochronous input pattern with a 300 ms IOI. The resting period of the oscillator is 500 ms. Panel **A** of this figure shows the isochronous input

pattern. Panel **B** shows the adaptive oscillator as it is entrained by this pattern. Period coupling forces successive inputs $i^r(n)$ to be closer to being in synchrony with the beginning of the oscillator's cycle. The precise change in the oscillator's period ($\Delta\Omega$) for each time step (ΔT), in the implementation of the adaptive oscillator, is determined using a discrete approximation to Equation 25:

$$\Delta\Omega = \alpha\Delta T \frac{\Omega}{2} \mathcal{P}[\phi^r(n), o(n)] \mathcal{M}[i^r(n)] - \beta\Delta T (\Omega - \bar{\Omega})(1 - \mathcal{M}[i^r(n)]). \quad (31)$$

As the oscillator is entrained by the input pattern, its period approaches 300 ms (the IOI of the input pattern), as shown in Panel **D**. Panel **C** shows that the output concurrently approaches 1.0, reflecting an accurate measure of synchronization. After the last input, the period gradually returns to its resting value of 500 ms.

3.4 The Dynamics of Adaptive Oscillation

In this section, empirical Arnold maps are constructed to investigate the mode-locking behavior of the proposed class of adaptive oscillators. This investigation is incremental, focusing on changes to four specific parameters: the phase resetting (or coupling) strength ($w_i I$), the entrainment rate (α), the output weight (w_o), and the exponent (γ_{\max}) specifying the maximum “sharpness” of the activation function. This section will conclude with a comparison of the proposed class of adaptive oscillators to a similar oscillator mechanism recently proposed by Large and Kolen (1995).

The construction of Arnold maps for the adaptive oscillator requires that the adaptive oscillator be specified as a circle map. This requirement necessitates three simplifying assumptions in the functioning of the adaptive oscillator mechanism. First, in terms of a circle map, iterative snapshots (ϕ_i) of the adaptive oscillator's phase are taken at each successive input pulse, which are assumed to be isochronous. Thus, the circle map for the proposed adaptive oscillator is given by

$$\phi_{i+1} = \begin{cases} 0 & \text{if } [w_i I + a(\phi_i + \frac{T}{\Omega_i})] > 1.0 \\ (\phi_i + \frac{T}{\Omega_i}) \bmod 1 & \text{otherwise} \end{cases} \quad (32)$$

where the activation of the oscillator is specified as a function of phase, T is the fixed period of the input pattern and Ω_i is the period of the oscillator updated after the i th input. The second assumption is that period adjustment occurs all-at-once after each phase-resetting input, and is thus not spread out in time by the input impulse-response function (\mathcal{M}). The circle map version of the equation for adjusting the period of the adaptive oscillator is given by

$$\Omega_{i+1} = \begin{cases} \Omega_i + \alpha(\phi_i + \frac{T}{\Omega_i})(1 - o_i) & \text{if } [w_i I + a(\phi_i + \frac{T}{\Omega_i})] > 1.0 \\ \Omega_i & \text{otherwise} \end{cases} \quad (33)$$

This assumption indicates that the oscillator's period does not “decay” in the absence of input to its resting value, since the oscillator's period is only adjusted after each

input-forced phase reset. The importance of including the decay process and the input impulse-response function will be addressed in Chapter 4, in the discussion of the Entrainment Model.

Parameter	Symbol	Step 1	Step 2	Step 3
Resting Period (seconds)	$\bar{\Omega}$	1.0	1.0	1.0
Decay Rate	β	0.0	0.0	0.0
Output Weight	w_o	1.0	1.0	1.0
Coupling Strength	$w_i I$	X	X	X
Entrainment Rate	α	0.0	X	X
Maximum Sensitivity	γ_{\max}	1.0	1.0	X

Table 1: Incremental investigation of stable entrainment of adaptive oscillators. The X 's indicate which parameters will be varied, and thus of are interest, in each step of the analysis.

The investigation of the adaptive oscillator's dynamics is in three steps. Step 1 examines stable entrainment (mode locking) of the adaptive oscillator, as a function of the bare winding number ($T/\bar{\Omega}$) and the coupling strength ($w_i I$), for phase-resetting without period adaptation (i.e., $\alpha = 0.0$). Step 2 varies the entrainment rate α to examine how the addition of period coupling influences stable entrainment. Finally, Step 3 describes the effect of modulating the maximum sharpness of the activation function (given by γ_{\max}) on stable entrainment, in comparison with the results from Step's 1 and 2.

Table 1 summarizes the three incremental steps in this investigation of the adaptive oscillators dynamics. Each column specifies the parameter settings of the adaptive oscillator for each step of the analysis; the X 's indicate which parameters vary

in each step and thus are of interest.

3.4.1 Step 1: Phase Resetting

Figure 3.12 shows the Arnold map constructed for an adaptive oscillator with phase-resetting, but without period adaptation ($\alpha = 0.0$). Stable entrainment of this oscillator was examined for bare winding numbers ($T/\bar{\Omega}$) (representing the uncoupled system) ranging from 0.0 to 1.0 in steps of 0.001, with the resting period of the oscillator ($\bar{\Omega}$) fixed at 1.0 second; and for coupling strengths ($w_i I$) also ranging from 0.0 to 1.0 in steps of 0.001. For each such initial condition (out of 1,000,000 total conditions), the circle map for the adaptive oscillator (Equation 32) was iterated for 500 steps, and the resulting phase differences ($\phi_{i+1} - \phi_i$) were accumulated to approximate the limit in Equation 14, which specifies the number of cycles of the adaptive oscillator per input cycle for the coupled system. Only those obtained values which corresponded to stable entrainment ratios in the first four levels of the Farey tree, within a tolerance of 0.001, were included as points in Figure 3.12.

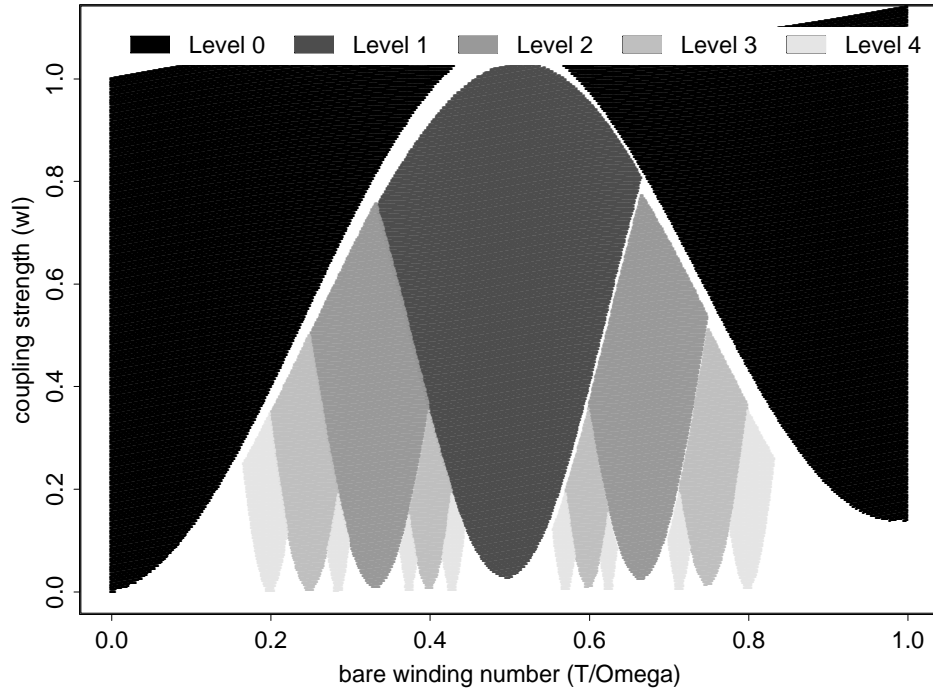


Figure 3.12: Arnold map for the phase-resetting adaptive oscillator with $\alpha = 0.0$. Parameter regions corresponding to stable entrainment ratios are coded according to their level in the Farey Tree.

This figure is coded by Farey-Tree level to show which initial conditions result in

stable entrainment to ratios at each of the first four levels of the Farey tree. From this coding, the tree-like structure of the Arnold map should be apparent; level 0 indicates the regions of 0:1 and 1:1 entrainment; level 1 indicates the single region of 1:2 entrainment; level 2 indicates the regions of 1:3 and 2:3 entrainment; level 3 indicates regions of 1:4, 2:5, 3:5, and 3:4 entrainment; and level 4 consists of regions of 1:5, 2:7, 3:8, 3:7, 4:7, 5:8, 5:7, and 4:5 entrainment. All parameter regions at the same level of the Farey tree have the same width, with regions at successive levels becoming narrower than preceding levels (corresponding to less stable entrainment). Because of the phase-resetting property of the adaptive oscillator, the sinusoidal shape of the oscillator's activation function is visible in the shape of each parameter region. As is apparent by comparing this figure with Figure 3.5, the entrainment of oscillators based on this phase-resetting circle map is more stable than the entrainment of oscillators based on the sine circle map.

3.4.2 Step 2: Period Coupling

Step 2 varied the entrainment rate (α) set to zero in Step 1, to examine how the addition of period coupling influences the entrainment of oscillators based on the phase-resetting circle map. In Step 2, the same procedure was used to construct Arnold maps as in Step 1, in which the bare winding number ($T/\overline{\Omega}$) and the coupling strength $w_i I$ ranged from 0.0 to 1.0 in steps of 0.001, with the resting period of the oscillator fixed at 1.0 seconds. Arnold maps were constructed for entrainment rates (α 's) of 0.1, 0.2, 0.3, 0.4, and 0.5. The Arnold maps determined for these five choices of α were found to be identical, thus it suffices to illustrate only a single case.

Figure 3.13 shows the Arnold map determined for $\alpha = 0.5$. It should be apparent to the reader that this Arnold map is identical to the Arnold map determined in Step 1 for phase-resetting only ($\alpha = 0.0$). Thus, the addition of a period coupling term does not alter the specific entrainment ratios achieved by the adaptive oscillator in response to isochronous input. That is, given the same initial conditions as the phase-resetting oscillator (coupling strength and bare winding number), the adaptive oscillator successfully adapts its period to match the “target” pattern of phase-resets imposed by the isochronous input. This target can only be measured, globally, by iterating the circle map to obtain its winding number $\frac{T}{\overline{\Omega}}$, yet the adaptive oscillator attains a period Ω_i that achieves this winding number using only local phase information.

For example, suppose the period of the input is 0.6 seconds and the resting period of the oscillator is 1.0 seconds, resulting in a bare winding number ($T/\overline{\Omega}$) of 0.6 seconds. From Step 1 (and Figure 3.12), we know that for a range of coupling strengths, the resultant pattern of phase resets will reflect 1 : 2 entrainment. From Step 2, we know that with the addition of period coupling, the adaptive oscillator will adjust its period (in this case to 1.2 seconds) to achieve 1:2 entrainment, and then to maintain this ratio in the absence of input forced phase-resetting.

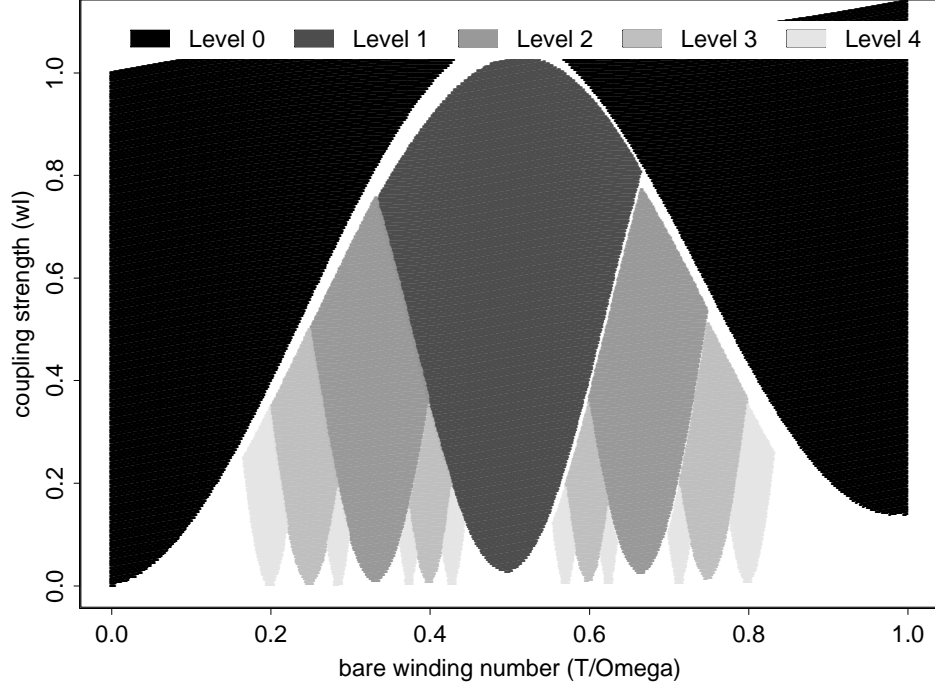


Figure 3.13: Arnold map for the phase-resetting adaptive oscillator with $\alpha = 0.5$.

In terms of the entrainment dynamics, there are two important implications of adding period coupling to the phase-resetting oscillator. First, the specific entrainment ratio attained by an adaptive oscillator (congruous with the pattern of phase resets) is maintained in the event of an occasional missing input, as well as when the pattern stops. Second, period coupling improves the stability of entrainment in the presence of noise in the timing of the input pattern, effectively placing each initial condition (specified by $T/\bar{\Omega}$) at the center of its corresponding parameter region.

3.4.3 Step 3: Activation Function Modulation

Step 3 varied the maximum exponent of the activation function γ_{\max} to examine the influence of modulating the shape of the activation function on stable entrainment. Using the same procedure as in Step's 1 and 2, Arnold maps were constructed for $\gamma_{\max} = 5$ and $\gamma_{\max} = 10$, with the entrainment rate (α) fixed at 0.5.

Figure 3.14 shows the Arnold map determined for $\gamma_{\max} = 5.0$ and Figure 3.15 shows the Arnold map determined for $\gamma_{\max} = 10.0$. As in Step's 1 and 2, parameter regions with stable entrainment are coded according to their level in the Farey tree. In comparison with the previous figures, one main effect is observed. As the maximum exponent (γ_{\max}) increases, regions corresponding to complex entrainment (i.e.,

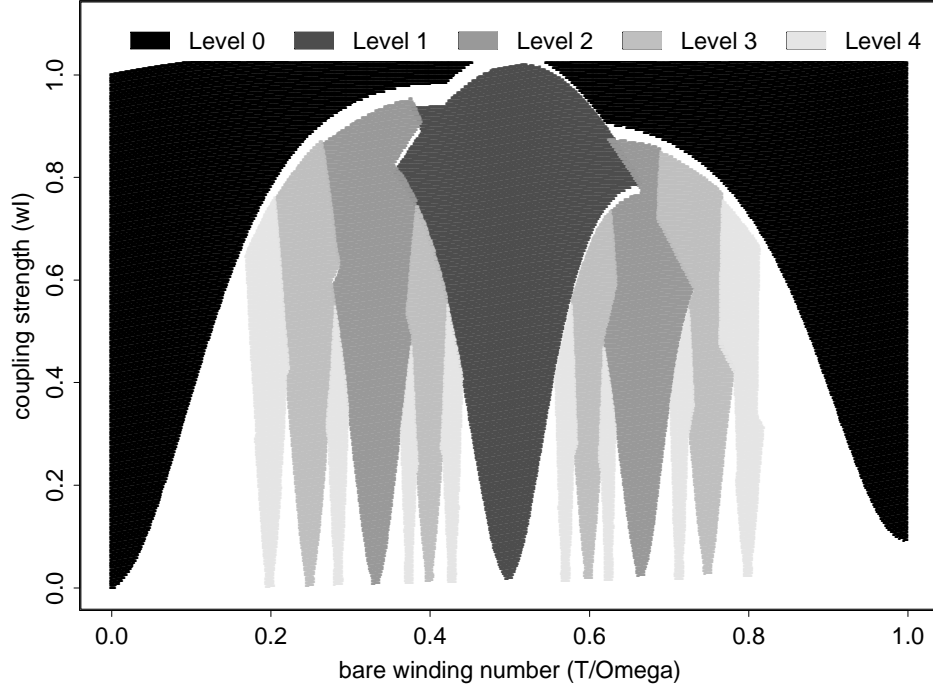


Figure 3.14: Arnold map for the phase-resetting adaptive oscillator for $\alpha = 0.5$ and $\gamma_{\max} = 5$

deeper levels in the tree structure) begin to widen, “eating away” regions of simpler entrainment, although not necessarily symmetrically. Thus, level 4 regions eat away portions of level 3, 2, 1, and 0 regions, level 3 regions eat away portions of level 2, 1 and 0 regions, and so on. These figures seem to be accumulating “white-space” because levels 5 and deeper, which were not tested for in the construction of the Arnold map, are becoming more and more prominent. Thus, sharpening the activation function of the adaptive oscillator in response to an input pattern tends to push it towards complex entrainment ratios. This is especially beneficial in the case where 0:1 entrainment (which corresponds to the period of the oscillator adapting out of control towards infinity) is replaced by a “more complex” ratio such as 1:3 or 1:4.

3.4.4 Comparison with the Large & Kolen Oscillator

In two recent reports, Large (1994), and Large and Kolen (1995), have proposed an independently developed adaptive-oscillator mechanism for the perception of musical meter, similar to the class of adaptive oscillators I have proposed in this thesis. In light of the present discussion concerning the dynamics of adaptive oscillation, I will outline their model’s main properties, as well as highlighting important similarities

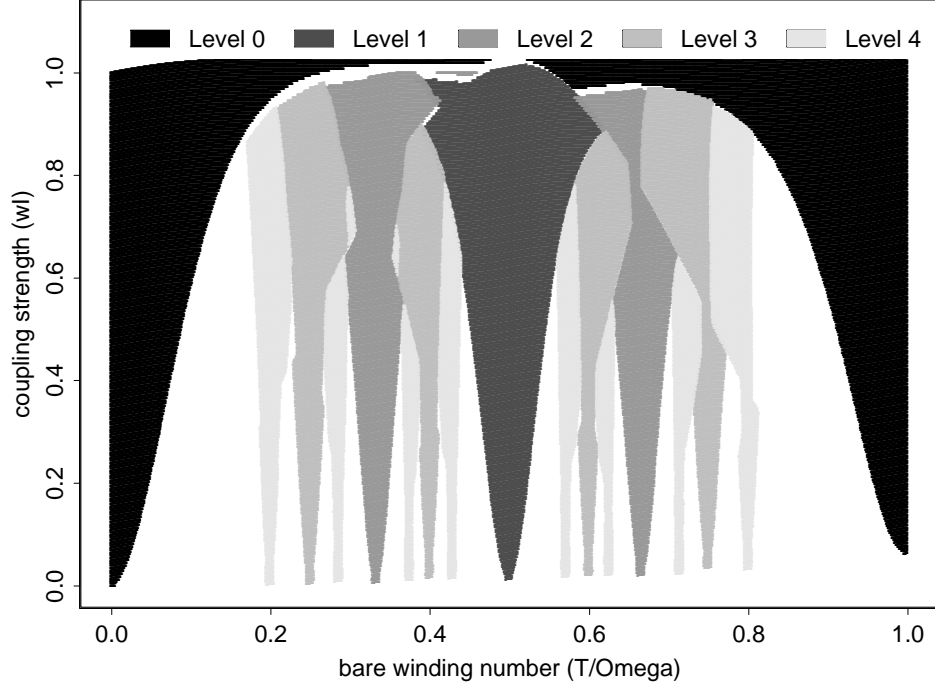


Figure 3.15: Arnold map for the phase-resetting adaptive oscillator for $\alpha = 0.5$ and $\gamma_{\max} = 10$

and differences between our two approaches. When possible, I will apply the same notation used in this thesis to the description of the Large and Kolen oscillator.

Like the present thesis model, the Large and Kolen oscillator has a periodic activation function with an intrinsic resting period. However, the shape of this activation function is not a simple sinusoid, but instead equal to

$$a(t) = 1 + \tanh\gamma\left(\cos\left(\frac{2\pi t}{\Omega}\right) - 1\right). \quad (34)$$

The hyperbolic tangent in this equation serves to sharpen the activation of the oscillator (reducing its “temporal receptive field”) with γ controlling how much the temporal receptive field of the oscillator is narrowed. In the present model, the output ($o(n)$) serves the same function as γ , with γ_{\max} in the present model controlling the maximum amount of activation sharpening.

Large and Kolen derive a series of *delta rules* for incrementally adjusting phase, period, and the γ controlling the temporal receptive field, following each input pulse. In terms of a circle map description, the rule for adjusting the oscillator’s phase is

$$\phi_{i+1} = \phi_i + b \operatorname{sech}^2(\gamma \cos 2\pi\phi_i - \gamma) \sin 2\pi\phi_i \quad (35)$$

and the rule for adjusting the oscillator's period is

$$\Omega_{i+1} = \Omega_i + \frac{T}{\Omega_i} - \frac{\alpha}{2\pi} \text{sech}^2(\gamma \cos 2\pi\phi_i - \gamma) \sin 2\pi\phi_i \quad (36)$$

where b is the coupling strength and α is the entrainment rate. The corresponding circle map rules of the proposed model are shown in Equation 32 and Equation 33, respectively. There are several important differences in our approaches, partially illustrated by these respective sets of equations.

First, in the model I propose, the oscillator's phase is reset in response to a measure of total activation that exceeds a fixed threshold, whereas in the Large and Kolen model, the oscillator's phase is adjusted incrementally based on the first derivative of the activation function. Large and Kolen observe that their model reduces to the sine circle map for $\gamma = 0.0$ (see Figure 3.5 for an Arnold map description of the sine circle map). In Section 3.4.1, we observed that the mode-locking of the phase-resetting model is more stable than the sine circle map. Thus, unless there is a principled reason to use the sine circle map instead of the phase-resetting map (and I will argue there is not) the phase-resetting model is the more appropriate basis from which to construct an adaptive-oscillator mechanism.

Second, in model proposed here, the oscillator's period is adjusted on the basis of the magnitude of the output (a measure of synchronization based on a memory of past phase resets), whereas in the Large and Kolen model, the oscillator's period is adjusted on the basis of the first derivative of the activation function. Large and Kolen observe that one feature of their model is that period-adjustment enhances the stability of the model by widening of the parameter regions corresponding to stable entrainment ratios in the constructed Arnold map. This partially corrects for the instability introduced by using the sine circle map as the basis for the model in the first place.

Finally, my model includes (1) a process by which, in the absence of input, the oscillator's period gradually returns to its resting value, and (2) an input impulse response transform which extends the adjustment of the oscillator's period in time, as well as specifies the interaction between the entrainment and decay processes. The Large and Kolen oscillator has neither of these two properties. In Chapter 4, I show that these two properties of the adaptive oscillator are critical to the parsimonious explanation of a range of tempo perception data using the adaptive-oscillator-based Entrainment Model.

3.5 Evaluating the Adaptive Oscillator

This chapter concludes with an evaluation of the proposed adaptive oscillator to-be-entrained by a stimulus set of 35 temporal patterns of varying rhythmic complexity, used by Povel and Essens (1985) to test listeners abilities to memorize and reproduce rhythmic patterns. Each of the patterns in this stimulus set consisted of a repeating

sequence of intervals 16 time units long, where a time unit was 200 ms—in the range of musical beats. All patterns were distinct permutations of a single set of intervals based on this 200 ms beat. Table 2 specifies the patterns, where the intervals comprising the patterns are described in terms of the number of beats. All patterns consisted of four 1-beat intervals (i.e., four 200-ms intervals), two 2-beat intervals, one 3-beat interval, and one 4-beat interval for a total cycled duration of 3200 ms (see Figure 3.16 for a graphical illustration of the patterns). Some of these patterns were easy for listeners in the Povel and Essens study and others were quite difficult—as indicated by how long it took listeners to learn a pattern and by how accurately they reproduced it. In general, patterns in this stimulus set are ordered according to their rhythmic complexity. All of these patterns evokes a sense of periodic beats for the perceiver. If asked to “beat along” with these patterns, most listeners tap out beats approximately every 400 or 800 ms, consistent with a 2/4 musical meter.

1.	111131224	11.	112131214	21.	111121234	31.	111211324
2.	112211314	12.	121112314	22.	111231124	32.	111312124
3.	211211314	13.	121211134	23.	113121124	33.	121113124
4.	221111314	14.	131212114	24.	211321114	34.	123111124
5.	312211114	15.	311211214	25.	231112114	35.	231121114
6.	112112134	16.	121112134	26.	111223114		
7.	211121314	17.	121121134	27.	121123114		
8.	131111224	18.	121141214	28.	123112114		
9.	132112114	19.	131211124	29.	211123114		
10.	211211134	20.	131211214	30.	311112124		

Table 2: Interval-based description of the temporal patterns from Experiment 1 of Povel and Essens (1985). For this pattern coding, 1 = 200-ms IOI, 2 = 400-ms IOI, 3 = 600-ms IOI, and 4 = 800-ms IOI.

3.5.1 Rationale

One can ask several questions with regard to the entrainment of the adaptive oscillator in comparison with listeners’ abilities to reproduce these patterns. First, can the adaptive oscillator achieve and maintain stable entrainment for temporal patterns of varying rhythmic complexity. Earlier, I showed stable entrainment of the adaptive oscillator by isochronous inputs. Second, if it can entrain to more complex rhythms, does the adaptive oscillator lock onto a beat period consistent with listeners’ perception of beats? Third, does the adaptive oscillator align its beats appropriately (i.e., are the down beats in the right place?). The ability of the adaptive oscillator to display appropriate behavior, as determined by answering the above three questions,

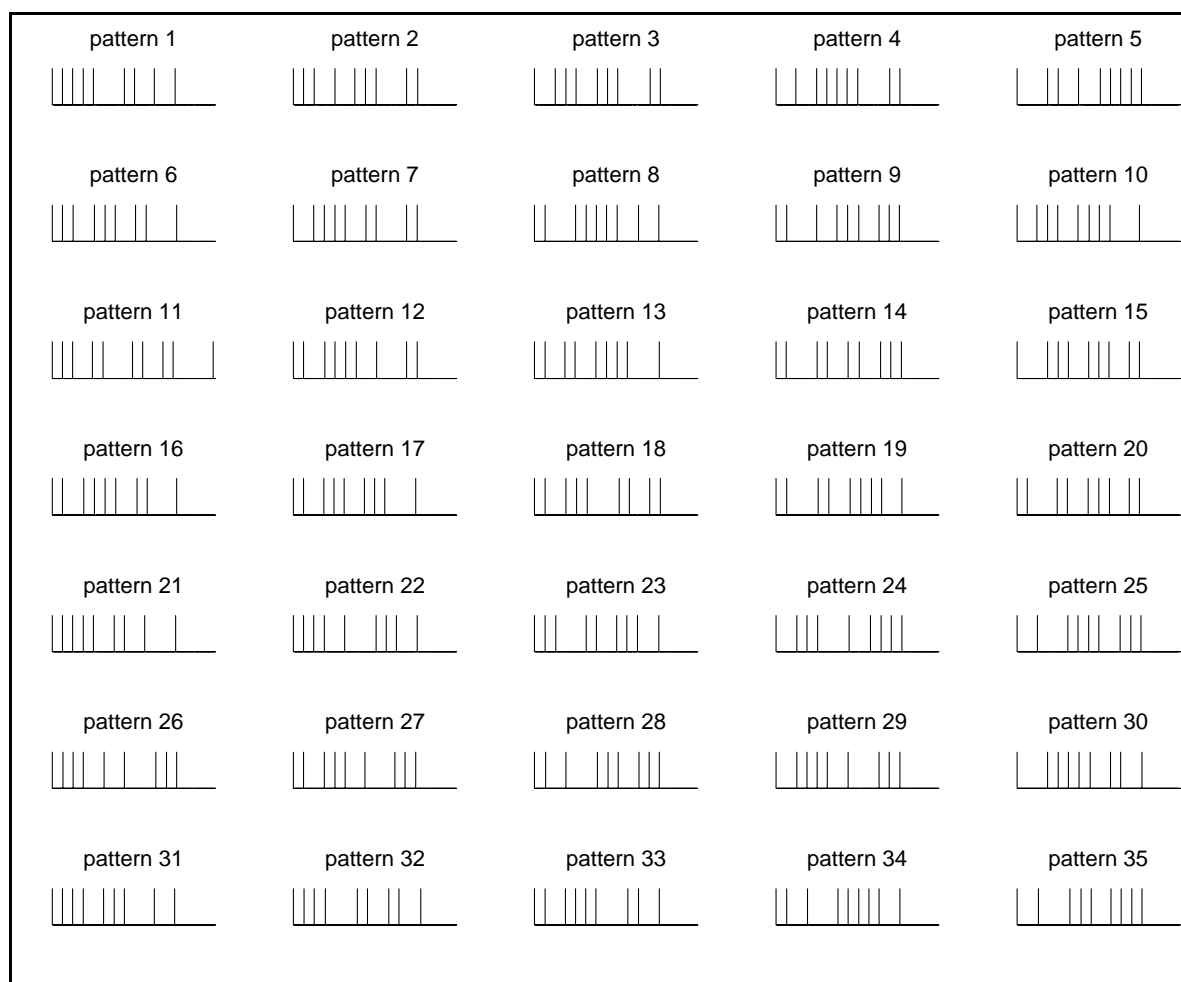


Figure 3.16: Input representation of the temporal patterns used in Experiment 1 of Povel and Essens (1985).

will demonstrate a successful *local* method for tracking beats of rhythmic patterns in spite of timing variability; that is, one that does not determine beats by applying rules to the entire pattern. In addition, this simulation addresses recent criticism of my proposed oscillator model by Large (1994) and Large and Kolen (1994) who suggest that it requires “strong assumptions about phenomenal accentuation in order to display appropriate behavior.” By strong assumptions about phenomenal accentuation, they are referring to accents introduced by modulating the amplitude of the input signal, such as an accent introduced by increasing the intensity of an input pulse relative to another. This criticism is addressed directly here, since the tested patterns are comprised of *identical inputs*.

3.5.2 Method

In order to answer these questions, a single adaptive oscillator was exposed to multiple cycles of each of the 35 rhythmic patterns in the Povel and Essens stimulus set. In all cases, the adaptive oscillator either achieved stable entrainment after four cycles or it was clear that it would not. For this simulation, the resting period of the oscillator was 500 ms, halfway between the maximum and minimum intervals comprising these input patterns. Before the cyclic presentation of an input pattern, the period of the oscillator was initialized to its resting value. The other parameter values of the oscillator were as follows: $\alpha = 4.0$, $\beta = 0.0$, $w_i = 0.5$, $w_o = 1.0$, $\gamma_{\max} = 1.0$, $\tau = 1000$, and $\Theta = 0.57$. These parameter settings were chosen to test the ability of a minimal version of the adaptive oscillator to be entrained by patterns in the set. For this minimal version, the adaptive oscillator did not alter the size of its temporal receptive field or decay its period towards its resting value. The adaptive oscillator was tested for four noise conditions: 0%, 5%, 7%, or 10% temporal variability added to each interval, as determined from a uniform random distribution.

3.5.3 Results

The results from all four noise conditions are summarized in Figure 3.17. For the no-noise condition, the adaptive oscillator locked onto a 400-ms beat period for 80% of the tested patterns, and for all but three of those patterns (Patterns 12, 33, and 34) beats coincided with natural accents in the patterns (cf., Handel 1989). For the remaining 20% of the tested patterns, its period oscillated between approximately 400 and 600 ms. With the addition of 5% “noise” to the timing of the intervals, the adaptive oscillator was able to be entrained by the remaining 20% of the tested patterns; thus, performance improved with noise. However, additional variability beyond 5% reduced the ability of the adaptive oscillator to be entrained by the tested patterns, with performance in the 10%-noise condition slightly below that in the no-noise condition.

Two examples (Patterns 5 and 12) illustrate the adaptive oscillator’s performance in the no-noise and 5%-added-noise conditions, in which entrainment by the Povel and Essens pattern set improved from 80% to 100%. The responses of the model to Patterns 5 and 12 in the no-noise condition are shown in Figures 3.18 and 3.19, and those for the 5%-noise condition are shown in Figures 3.20 and 3.21, respectively. In these figures, the last cycle of the tested pattern is displayed in Panel (A), the entrained response of the oscillator for the last cycle is shown in Panel (B), the output (its measure of synchrony) over the course of all four cycles is shown in Panel (C) and the corresponding changes in the oscillators’ period are shown in Panel (D).

For Pattern 5 in the no-noise condition, the oscillator failed to be entrained; instead its period oscillated between approximately 400 and 600 ms every two cycles as shown. However, in the 5%-noise condition for the same pattern, the period did converge to 400 ms after about a cycle and a half. For pattern 12 in the no-noise

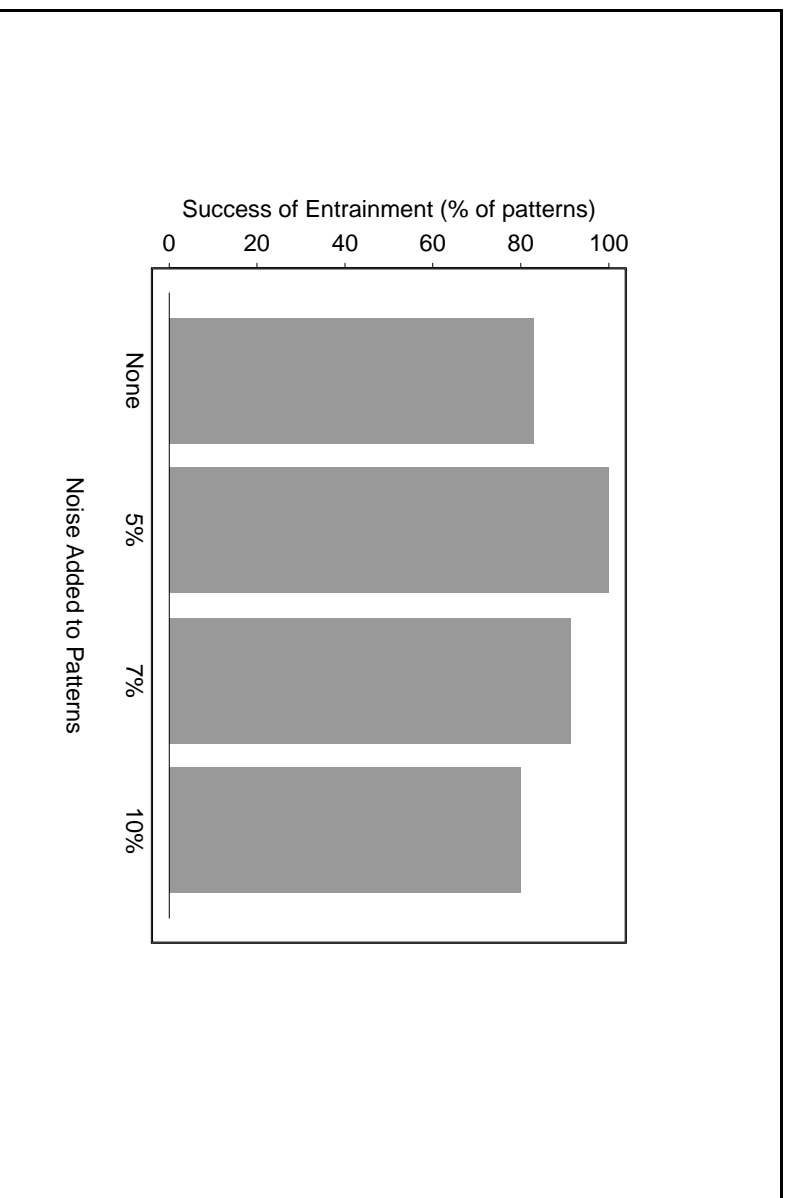


Figure 3.17: Performance of adaptive oscillator on Povel and Essens (1985) data set for the four noise conditions.

condition, the oscillator successfully achieved and maintained a 400-ms beat period, but its placement of beats did not coincide with natural accents in the pattern. For this pattern, most listeners would hear the second of the first two inputs as accented, not the first as indicated by the model. With the addition of 5% noise to Pattern 12, the oscillator still achieved a 400-ms beat period, but its placement of beats was shifted to correspond with a more natural pattern of accents than in the no-noise condition.

In summary, stochastic variability in the timing of intervals in a rhythmic pattern (or stochastic variability in the adaptive oscillator's period) can improve entrainment in the following two ways: (1) by forcing the adaptive oscillator out of a cyclic pattern of period changes and into a stable attractor (as evidence in Figure 3.20), and (2) by changing the way the oscillator phase locks to the input pattern, which may improve the placement of beats in some cases (as evidenced in Figure 3.21). Moreover, the results from this simulation clearly show that the phase-resetting adaptive oscillators proposed in this thesis, as well as earlier (McAnley 1994a; 1994b) do not require strong assumptions concerning phenomenal accentuation to display appropriate behavior, as asserted by Large (1994) and Large and Kolen (1994), who favor incremental adjustments in phase and dynamic modulation of the “sharpness” of the oscillator's

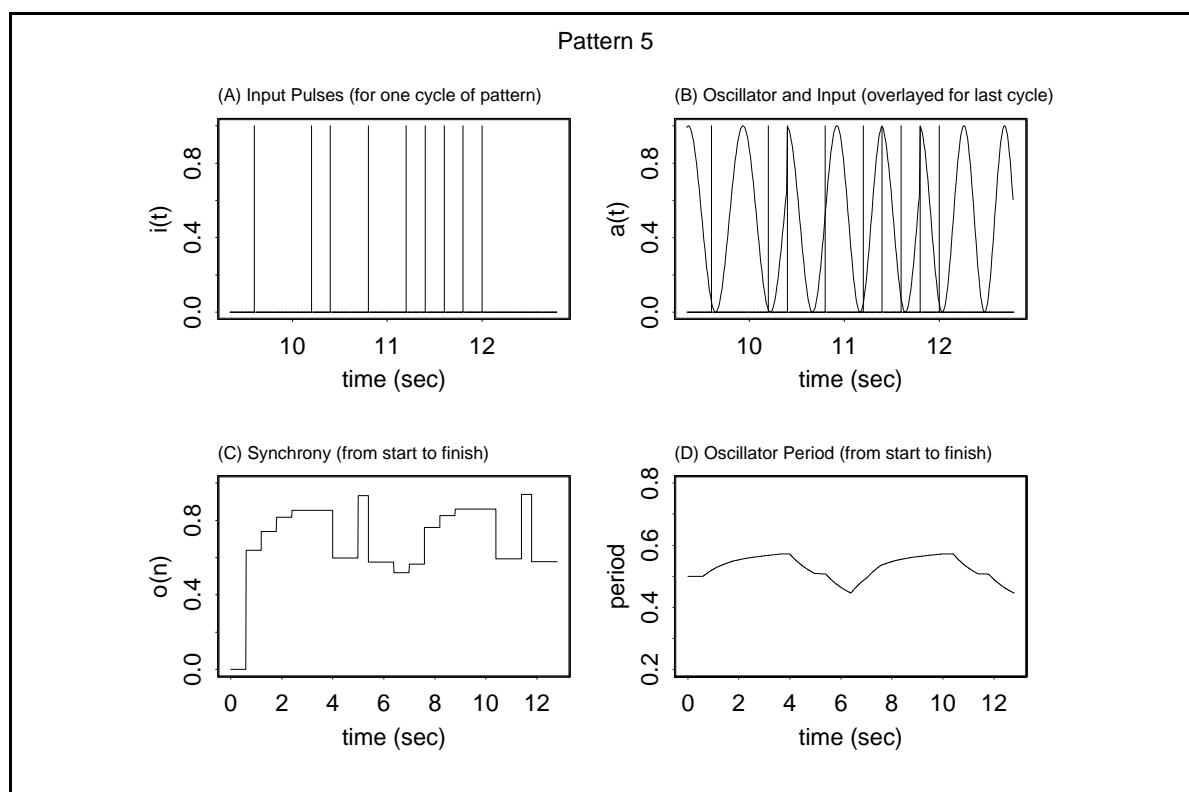


Figure 3.18: Performance of the adaptive oscillator on Pattern 5 in the no-noise condition.

activation function (its temporal receptive field). Instead, even a minimal version of the phase-resetting adaptive oscillator is able to appropriately entrain to 100% of the Povel and Essens' pattern set in the 5%-noise condition. However, there *are* some cases for which modulating the size of the temporal receptive field is advantageous, in particular to eliminate regions of 0 : 1 entrainment, in which the period of the oscillator grows unbounded towards infinity, or to enable the oscillator to establish more complex entrainment ratios (see Section 3.4.3 which examined the effect of modulating the size of the temporal receptive field on the stable entrainment of the adaptive oscillator, via sharpening the activation function).

3.5.4 Discussion

An interesting issue concerns comparing the time course of the adaptive oscillator's entrainment to rhythmic patterns with listeners' abilities to reproduce those patterns, assuming that in order for listeners to reproduce a rhythmic pattern, they must first be able to-be-entrained by it, in a sense engaging in a "cognitive oscillation." One way to obtain a measure of pattern difficulty for the model is to average the output (a measure of synchronization) over the course of the pattern's presentation. Using

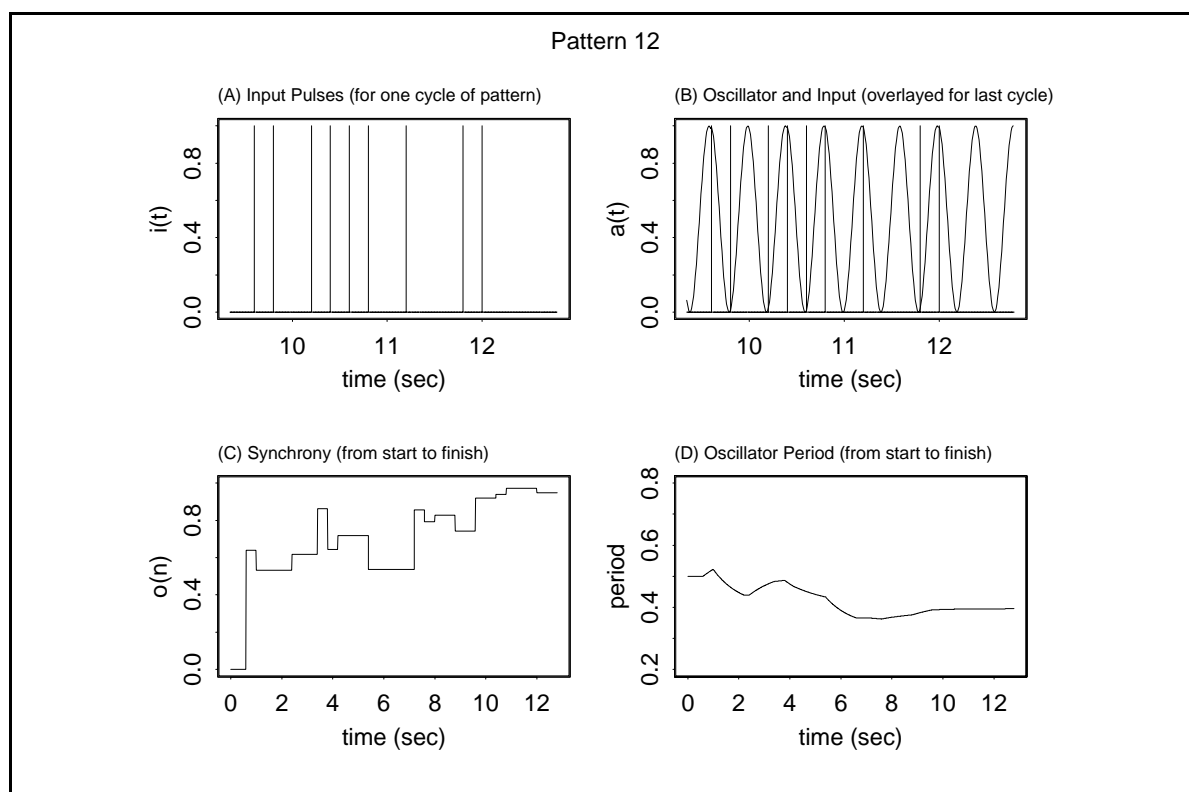


Figure 3.19: Performance of the adaptive oscillator on Pattern 12 in the no-noise condition.

this metric, smaller average outputs correspond to more difficult patterns.

In contrast, as discussed in Chapter 1, Povel and Essens proposed a rule-based clock model which was used to determine the rhythmic complexity of a temporal pattern. Recall that the processing of the Povel and Essens model is in three steps. First, phenomenal accents are determined using three preference rules suggested by Povel and Okkerman (1981) for patterns of physically identical events. Second, all possible clocks, defined as multiples of a fundamental period (in this case 200 ms) are generated up to a clock-period equal to half the cycle duration of the rhythmic pattern (in this case 1600 ms). This generation of all possible clocks includes the generation of all possible alignments of each clock (i.e., all possible downbeats). Third, for each such clock, a rule is applied which scores the amount of dissonance between the clock and the natural patterning of accents determined in step 1. The best clock (i.e., the one with the least dissonance) indicates the preferred beat period and the alignment of the beats for the rhythmic pattern. In the clock model, the dissonance of the best clock is used as a measure of rhythmic complexity. The patterns shown earlier in Figure 3.16 are ordered according to this dissonance measure of rhythmic complexity.

An obvious question is whether the proposed adaptive-oscillator measure of pattern difficulty correlates with listeners' abilities to reproduce rhythmic patterns. To

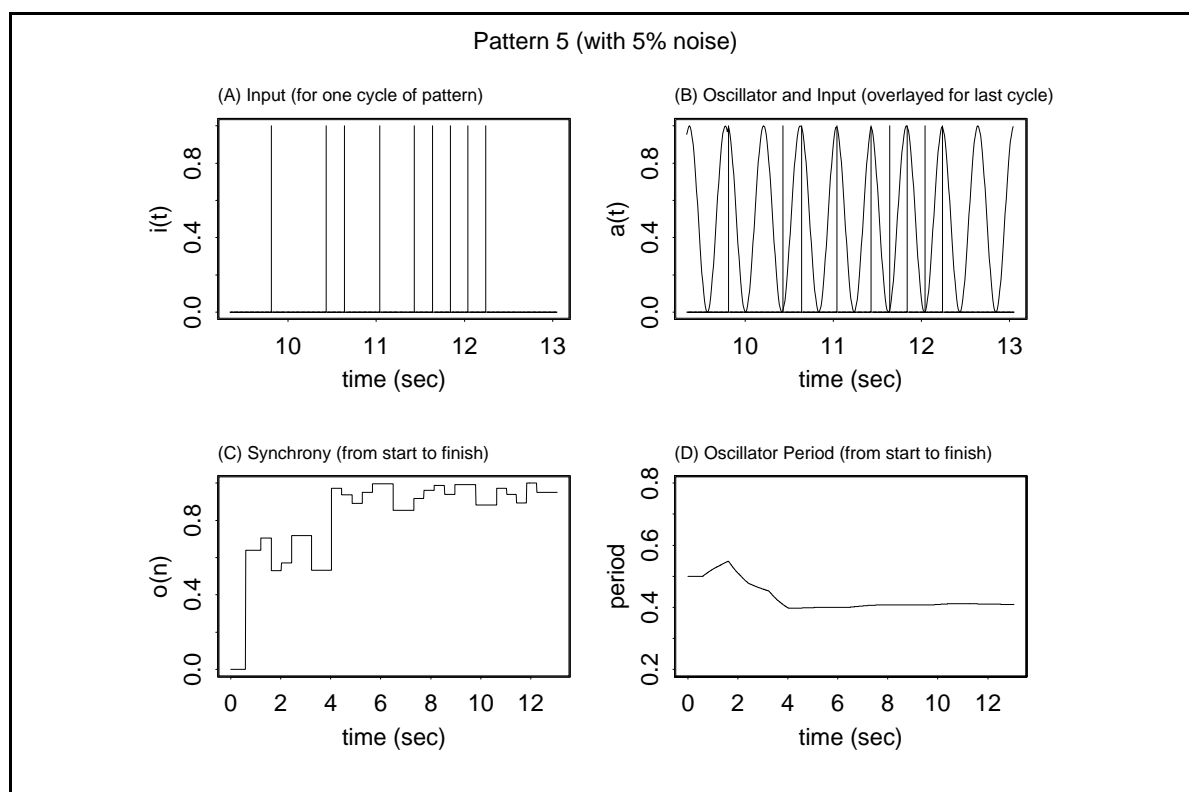


Figure 3.20: Performance of the adaptive oscillator on Pattern 5 in the 5%-noise condition.

test this, pattern difficulty was measured using the proposed synchrony metric for all 35 patterns of the Povel and Essens stimulus set for the 5%-noise condition. Although not perfect, there is a positive correlation between the model's measure of pattern difficulty and the mean timing variability in listeners' reproductions of these patterns, as determined by Povel and Essens (1985). To illustrate one example, Povel and Essens found that Pattern 5 was more accurately reproduced and easier to memorize than Pattern 12. Similar, the adaptive oscillator was entrained by Pattern 5 more quickly than by Pattern 12. The smaller average output for Pattern 12 than for Pattern 5 illustrates that Pattern 12 was more difficult for the model than Pattern 5, in agreement with listeners' performance.

Although this examination of rhythmic complexity using the adaptive oscillator is preliminary, the technique I have introduced offers an alternative way to measure the complexity of rhythmic patterns, one that is potentially better correlated with listeners' abilities to memorize and reproduce those patterns than Povel and Essens rule-based model. In the next chapter, the adaptive-oscillator mechanism proposed here is used as the basis for an Entrainment Model of time perception. This model is evaluated in three tempo-discrimination simulations. The predictions of the model are compared with human listener data for analogous tempo-discrimination tasks,

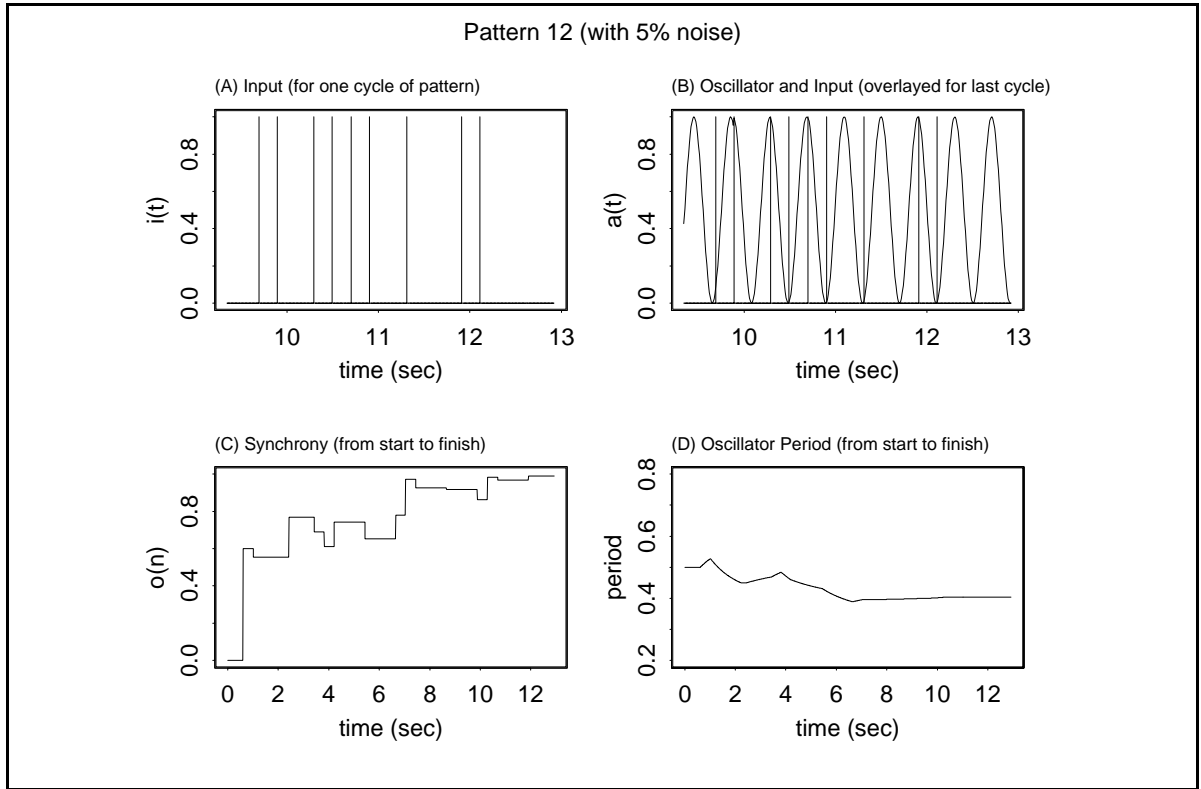


Figure 3.21: Performance of the adaptive oscillator on Pattern 12 in the 5%-noise condition.

providing additional support for the proposed adaptive-oscillator mechanism for the processing of rhythmic patterns.

Chapter 4

The Entrainment Model

Returning now to the entrainment hypothesis, recall that Jones (1976) proposed that the entrainment of attentional periodicities is the basis for the development of expectancies of when in time events in a pattern are likely to occur, guiding the temporal placement of attentional pulses, and thus influencing the overall perception of a pattern, including the time intervals that comprise it. The development of adaptive oscillators in the previous chapter provide an instantiation of this hypothesis, and thus are a useful basis from which to develop an entrainment model of time perception.

As discussed in Chapter 2, the Contrast Model proposed by Jones and Boltz (1989) was also based on this entrainment hypothesis. However, predictions derived from this model were restricted by the data that it attempted to explain, since model predictions were not linked to the dynamics of a functioning oscillator. In particular, Chapter 2 described a set of embedded-interval and tempo-discrimination data that were not intended to be explained by the Contrast Model and which were difficult to see how this model could explain without substantial revision. In addition, these data have proved to be difficult to provide an explanation for with the Multiple-Look Model (Drake and Botte, 1993). It is these problematic data that I will now address in the development of the adaptive-oscillator-based Entrainment Model.

4.1 Model Specification

Three assumptions form the conceptual framework of the Entrainment Model. First, estimates of the duration of a single time interval or the tempo of a pattern of intervals correspond directly to the periods of adaptive oscillators. Second, the detection of a change in duration or tempo is the consequence of abrupt changes in the phase of an adaptive oscillator. Third, the degree of one's ability to detect changes in duration and tempo is related to the degree of entrainment of adaptive oscillators to similar external periodicities. Each of these three assumptions will now be elaborated upon, forming the core of *the time-as-phase hypothesis*.

4.1.1 The Mapping Between External and Internal Periods

Based on the evidence (described in Chapter 2) supporting a linear psychophysical law for time, the proposed Entrainment Model of time perception assumes a linear function of the form

$$\bar{\Omega} = mT + b, \quad (37)$$

relating stimulus time intervals (T) to their subjective estimation ($\bar{\Omega}$) when presented in isolation. Thus, Equation 37 specifies the mapping from external time intervals to internal periodicities, instantiated as adaptive oscillators with resting periods ($\bar{\Omega}$). The parameters m (slope) and b (minimum subjective duration) specify the precise mapping between the external time intervals (T) and the resting period ($\bar{\Omega}$) of the tracking oscillator.

For a linear mapping to capture the overestimation of short intervals and the underestimation of long intervals, as commonly reported in the time perception literature (see Chapter 2), the y-intercept b (or minimum subjective duration) must be non-zero and the slope m must be between 0.0 and 1.0. For all of the model simulations reported in this thesis, I assume a minimum subjective duration (b) of 25 ms and an indifference interval equal to 600 ms. The slope of Equation 37 is then determined by

$$m = 1 - \frac{b}{T}, \quad (38)$$

resulting in $m = 0.9583$ for $b = 25$ ms and $T = 600$ ms.

By linking the concepts of subjective duration and attentional periodicity, we introduce the concepts of an attentional indifference interval and a minimum attentional period (or maximum attentional frequency). It is interesting to note that a minimum subjective duration of 25 ms corresponds, in terms of frequency, to 40 Hertz, a commonly observed frequency of synchronized neural activity in different cortical areas (Baird et al., 1994b; Gray et al., 1989)

4.1.2 Time as Phase

The second assumption of the Entrainment Model concerns the representation of time as phase. Phase is a measure of relative time, in that a change (ΔT) in the duration of a base interval (T) corresponds to a phase change ($\Delta\phi$) that is a fraction of the period (T):

$$\Delta\phi = \frac{\Delta T}{T} \pmod{1}. \quad (39)$$

For the Entrainment Model in which an adaptive oscillator with period Ω is a subjective estimate of the time interval T , the relationship between a time change ΔT and the resultant phase change is given by

$$\Delta\phi = \frac{T + \Delta T}{\Omega} \pmod{1}. \quad (40)$$

Thus, the phase change $\Delta\phi$ registers the effect of ΔT on the adaptive oscillator tracking the isochronous series of time intervals (T). Appropriately, the phase change $\Delta\phi$ (or *phase-difference*) is a phase correlate of the *time-difference* (ΔT); for example, if $\Omega = T$ then a phase-difference ($\Delta\phi$) of zero corresponds to a time-difference (ΔT) that is also zero. With respect to the Entrainment Model, the interesting issue concerns how the relationship between phase differences ($\Delta\phi$) and time differences (ΔT) vary as function of the ratio between the base interval T and the subjective estimate Ω , expressed as $\frac{T}{\Omega}$ (i.e., the amount of over- and underestimation). For $\frac{T}{\Omega} > 1$ time is underestimated, for $\frac{T}{\Omega} < 1$ time is overestimated, and for $\frac{T}{\Omega} = 1.0$ subjective time is identical to clock time (i.e., the base interval T is an indifference interval). In order to investigate how the relationship between time differences and phase differences vary as a function of this ratio ($\frac{T}{\Omega}$), it is useful to rewrite Equation 40 as

$$\Delta\phi = [\frac{T}{\Omega}] + [\frac{\Delta T}{\Omega}] \pmod{1} \quad (41)$$

and to represent phase on $[-0.5, 0.5]$ instead of on $[0, 1]$. In which case, positive phase change *ideally* corresponds to a positive time change $+\Delta T$ and negative phase change ideally corresponds to a negative time change $-\Delta T$. Thus, the phase changes associated with $+\Delta T$ and $-\Delta T$ will be distinguished as $\Delta\phi_+$ and $\Delta\phi_-$, respectively. For the chosen mapping between time interval (T) and the resting period ($\bar{\Omega}$) of the adaptive oscillator, there are three cases to consider.

Case 1: $\frac{T}{\Omega} = 1.0$. For Case 1, $\Omega = T$ with T corresponding to an indifference interval; thus Equation 41 reduces to

$$\Delta\phi = \frac{\Delta T}{T} \pmod{1}. \quad (42)$$

Consequently, lengthening or shortening T by $p\%$ produces the same magnitude of phase change regardless of whether the time change ΔT is positive or negative (i.e., $|\Delta\phi_-| = |\Delta\phi_+|$). To provide a concrete example of Case 1, suppose T is lengthened by 10%, as depicted in Figure 4.1, then $\Delta\phi_+ = 0.1$. On the other hand, if T is shortening by 10%, then $\Delta\phi_- = -0.1$. For both the 10% increase and 10% decrease in duration, the magnitudes of the phase changes are equal ($|\Delta\phi_+| = |\Delta\phi_-| = 0.1$). Moreover, $\Delta T = 0.0$ corresponds to $\Delta\phi = 0.0$.

Case 2: $\frac{T}{\Omega} > 1.0$. For Case 2, $\Omega < T$ and subjective time is an underestimation of clock time. Consequently, the magnitudes of $\Delta\phi_+$ and $\Delta\phi_-$ are not equal for equivalent positive and negative time changes ΔT . Instead, within a limited range of equivalent positive and negative time changes (ΔT), the magnitudes of the resultant phase changes ($\Delta\phi$) are unequal, such that $|\Delta\phi_+| > |\Delta\phi_-|$. In addition, a zero time change ($\Delta T = 0.0$), which in Case 1 corresponded to a zero phase change ($\Delta\phi = 0.0$), corresponds here in Case 2 to a positive phase change ($\Delta\phi > 0.0$). Another effect

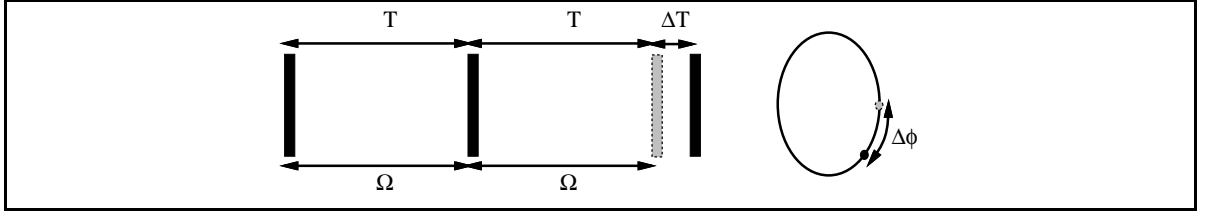


Figure 4.1: Illustration of Case 1: Lengthening of time interval (T) for an oscillator with a period (Ω) equal to T .

of underestimation with respect to the Entrainment Model is that it stretches the mapping between ΔT and $\Delta\phi$, since $\frac{\Delta T}{\Omega}$ in Equation 41 represents a larger fraction of the base interval T than $\frac{\Delta T}{T}$ does, as illustrated in Figure 4.2.

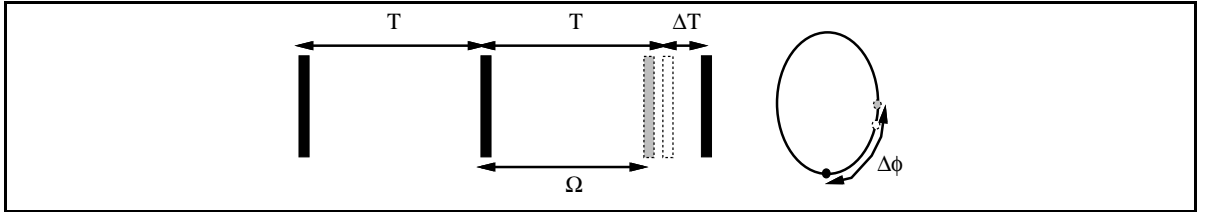


Figure 4.2: Illustration of Case 2: Lengthening of time interval (T) for an oscillator with period (Ω) that underestimates (is shorter than) T .

Case 3: $\frac{T}{\Omega} < 1.0$. For Case 3, which is similar but symmetric to Case 2, $\Omega > T$ and subjective time is an overestimation of clock time. Similar to Case 2, the magnitudes of $\Delta\phi_+$ and $\Delta\phi_-$ are not equal for equivalent positive and negative time changes (ΔT). Instead, symmetric with Case 2, $|\Delta\phi_-| > |\Delta\phi_+|$ within a limited range of ΔTs . Also for Case 3, the phase change ($\Delta\phi$) corresponding to a zero time change is negative, opposite from Case 2. Another effect of overestimation with respect to the Entrainment Model is that it compresses the mapping between ΔT and $\Delta\phi$, since $\frac{\Delta T}{\Omega}$ in Equation 41 represents a smaller fraction of the base interval T than $\frac{\Delta T}{T}$ does, as illustrated in Figure 4.3.

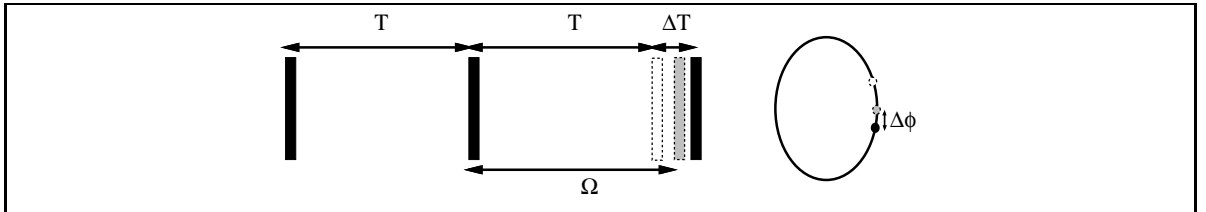


Figure 4.3: Illustration of Case 3: Lengthening of time interval (T) for an oscillator with period (Ω) that overestimates (is longer than) T .

In Chapter 2, we discussed how the over- or underestimation of an isolated time interval can result in negative or positive time errors respectively, in a comparison or reproduction task. For the Entrainment Model, over- and underestimation results in positive and negative phase errors; that is, the phase difference ($\Delta\phi$) corresponding to a zero time change ($\Delta T = 0.0$) is skewed from $\Delta\phi = 0$ to a negative or positive value. This phase is the model's point of subjective (phase) equality (PSE_ϕ), which is analogous to the reported point of subjective time equality for human listeners in comparison and reproduction tasks.

4.1.3 The Just-Noticeable Phase Difference

The third assumption of the Entrainment Model involves the definition of a phase correlate to the just-noticeable time difference (JND), which I will term the just-noticeable phase difference (JND_ϕ). The just-noticeable phase-difference specifies the smallest detectable time change ΔT in a base interval T , within the context of the adaptive oscillator tracking that time interval. JND_ϕ is define as an absolute threshold (independent of the sign of $\Delta\phi$). Thus, if the magnitude of the phase-difference ($\Delta\phi$) corresponding to a time change (ΔT) is greater than JND_ϕ then the time change (ΔT) is detected, otherwise it is not.

In agreement with the entrainment hypothesis, the Entrainment Model assumes that the just-noticeable phase difference decreases as the tracking adaptive oscillator becomes entrained by the input sequence. In Chapter 3, it was shown that the output sequence $o(n)$ provided a dynamic measure of the adaptive oscillator's synchrony with the input sequence that was useful as a "teaching" signal in the learning equations. In the Entrainment Model, the output $o(n)$ is used to modulate the just-noticeable phase difference according to the following rule:

$$JND_\phi = JND_{\phi_{\max}}[1 - o(n)] + o(n)JND_{\phi_{\min}}. \quad (43)$$

The parameters $JND_{\phi_{\min}}$ and $JND_{\phi_{\max}}$ establish the minimum and maximum just-noticeable phase differences. The minimum is set to a phase value near zero (for the reported simulations $JND_{\phi_{\min}} = 0.03$) and the maximum (for the represented range of phases) is set equal to 0.5. Each successive time interval that increases the output, reduces the just-noticeable phase difference. Thus, as the tracking adaptive oscillator is entrained by an input sequence, the time sensitivity of the Entrainment Model improves. In the limit, $o(n) = 1.0$ and $JND_\phi = JND_{\phi_{\min}}$.

In terms of the precise performance of the Entrainment Model on a time discrimination task, the main issue concerns the relationship between JND and JND_ϕ . If we suppose that in the limit, the tracking adaptive oscillator is perfectly entrained by the input sequence, then the period (Ω) of the oscillator is approximately equal to the time interval (T) and $JND_\phi = JND_{\phi_{\min}}$. This is an example of Case 1 since $\frac{T}{\Omega} \approx 1.0$. If Case 1 is true for a range of time intervals (T), then within this range, JND_ϕ is constant. Furthermore, since JND_ϕ is a measure of relative time, constant

JND_ϕ corresponds to constant relative JND (in agreement with Weber's law). Thus, perfect entrainment of the model across all time intervals (T) predicts time sensitivity consistent with Weber's law. A second implication of Case 1 is that thresholds for detecting time increases and time decreases are identical since $|\Delta\phi_+|$ and $|\Delta\phi_-|$ are the same for positive and negative time changes.

However, for the Entrainment Model, Case 1 only applies (initially) at the indifference interval specified by the mapping between external time intervals (T) and the resting period ($\bar{\Omega}$) of the tracking adaptive oscillator. For the suggested linear mapping (Equation 37) the indifference interval was specified as 600 ms. Thus, for intervals (T) greater than 600 ms, $\bar{\Omega} < T$ (Case 2 applies) and for intervals (T) less than 600, $\bar{\Omega} > T$ (Case 3 applies). For these cases, JND_ϕ is not constant and consequently performance does *not* correspond to constant relative JND, as would be predicted by Weber's law. Instead, JND_ϕ is influenced by the amount of under- or overestimation, according to Case 2 or Case 3, respectively. In addition, for time intervals in which Case 2 and Case 3 apply, the thresholds for detecting time increases and time decreases are not identical since $|\Delta\phi_+| \neq |\Delta\phi_-|$ for equivalent positive and negative time changes.

Moreover, the amount of under- or overestimation by the tracking adaptive oscillator, as well as the precise value of JND_ϕ is a dynamic property of the Entrainment Model. As a result, the relationship between JND and JND_ϕ is also dynamic, varying as a function of the ratio $\frac{T}{\bar{\Omega}}$ and the output $o(n)$, which can only be determined by simulating the experimental task.

4.1.4 Parameters

The dynamic relationship between JND_ϕ and JND (performance) will be examined in a series of three tempo-discrimination simulations in Section 4.2. Recall that the output $o(n)$ modulates the just-noticeable phase difference JND_ϕ according to Equation 43. Thus, the performance of the Entrainment Model is controlled by the seven parameters of the adaptive oscillator which influence the dynamics of the output function: α (the entrainment rate), β (the decay rate), γ_{\max} (maximum activation sharpening exponent), o_w (weighting of current phase-reset relative to previous output), $w_i I$ (input coupling strength), γ (the gain on the input impulse-response function), and Θ (the bias on the input impulse-response function). The input coupling strength, the impulse response gain, and the maximum exponent on the activation function are fixed parameters of the model and are specified as $w_i I = 1.0$, $\gamma = 1000.0$, and $\gamma_{\max} = 1.0$, respectively. The impulse response bias Θ , the entrainment rate α , the decay rate β , and the output weight o_w are free parameters of the model. The output weight o_w controls how fast the output changes in response to entrainment or decay of the oscillator, and thus influences how fast the model approaches the minimum just-noticeable phase difference. By specifying the gain γ as a "large number", the input impulse response function effectively reduces to a descending step function

(as is shown in Figure 3.10 of Chapter 3) for which the temporal boundary (IOI_θ) between entrainment and decay is given by

$$IOI_\theta = \left\lceil \frac{\log 0.5}{\Theta} \right\rceil. \quad (44)$$

Consequently, for t less than the threshold IOI_θ , the input entrains the oscillator, but for t greater than the threshold IOI_θ , the oscillator's period decays towards its resting value. In terms of the effect of the output $o(n)$, $t < IOI_\theta$ contributes to increasing the output and $t > IOI_\theta$ contributes to decreasing the output. The overall change in output after each *discrete* input pulse n reflects a combination of the increases and the decreases. The interaction of the entrainment and decay processes, mediated by α and β , as well as by the specific choice of threshold IOI_θ , is critical to the performance of the model.

4.2 Model Predictions

The predictions of the Entrainment Model relating to tempo discrimination are investigated for a series of three tempo-discrimination simulations, in which the tempo sensitivity of the Entrainment Model is compared with human tempo sensitivity. All of the simulations investigate the model's thresholds for detecting changes in the tempo of isochronous sequences. Because the structure of the simulations is mapped so closely to the tempo-discrimination experiments involving human listeners (down to trial-to-trial performance), the model participates in the simulations in much the same way as listeners participate in experiments. The model *performs* each trial of the experiment and the tempo sensitivity of the model is measured using a psychophysical procedure. This method is advantageous because, in addition to providing quantitative measures of performance in terms of mean relative JNDs, it permits the easy investigation of a number of issues related to the *process* of tempo discrimination, that can not be explored with a purely descriptive model. Examples of such issues are: (1) do listeners exhibit systematic errors in tempo discrimination, (2) are there threshold differences between detecting tempo increases and tempo decreases, and (3) how might the temporal structure of each trial affect discrimination thresholds? These are some of the issues that will be explored with the Entrainment Model in the simulations described below.

There are three simulations. Simulation 1 investigates the model's predictions concerning the influence of the number of sequence intervals and the sequence duration on relative JND. Simulation 2 investigates the model's predictions concerning differential sensitivity to tempo increases and decreases. Simulation 3 examines the model's predictions concerning the influence of temporally-directed attending on tempo sensitivity. In addition, Chapter 5 reports data from two listening experiments designed to test the predictions made from simulations 2 and 3, for which no human data was available for comparison.

4.2.1 Tempo Discrimination

All of the simulations evaluate the tempo sensitivity of the Entrainment Model using the 2AFC “which is faster” discrimination paradigm favored by Drake and Botte (1993) and depicted in Figure 4.4. For this task, the model’s tempo judgments are derived in a four step process, intended to model the process by which listeners make similar judgments. In the first step, it is assumed that listeners’ tempo judgments are based on entrainment to the standard sequence, measured by $o(n)$. The computation of $o(n)$ determines the just-noticeable phase-difference JND_ϕ . In the second step, the detection of a tempo change is immediate, following the first *different* interval ($T + \Delta T$) of the comparison sequence with the ending marker of the first different interval establishing the absolute phase-difference $|\Delta\phi|$ (or $|\Delta\phi^r(n+1)|$ in terms of the notation used in Chapter 3). In step 3, if $|\Delta\phi|$ is greater than the threshold JND_ϕ , then a tempo change is detected, otherwise not. In step four, if a tempo change is detected in step 3, then the sign of $\Delta\phi$ establishes the direction of the tempo change (increase or decrease). Negative $\Delta\phi$ indicates that the comparison sequence is faster than the standard sequence; positive $\Delta\phi$ indicates that the comparison sequence is slower (or the standard sequence is faster in terms of the “which is faster” task).

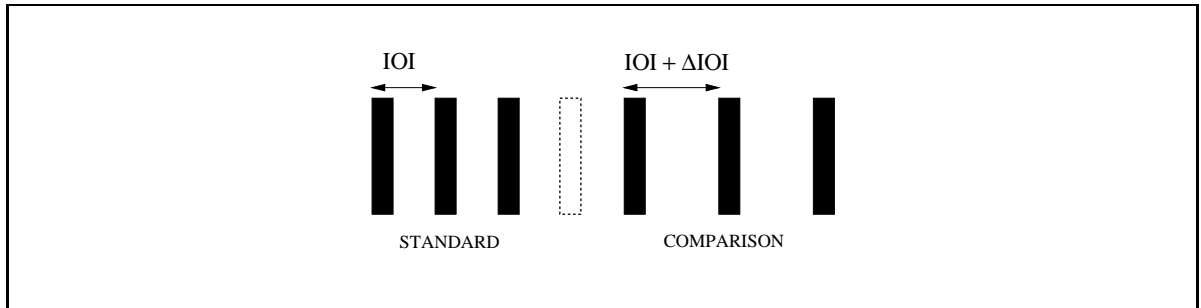


Figure 4.4: Illustration of the “which is faster?” tempo-discrimination task for a 2-interval isochronous standard.

Figure 4.5 shows the response of the adaptive oscillator in such a simulated “which is faster” trial. In this example, the resting period of the oscillator is 600 ms and the input sequence represents a 4-tone standard sequence with a 500-ms IOI followed by a 4-tone comparison sequence that is 20% slower ($\Delta T = +100\text{ms}$). The gap between the standard and comparison sequences is 1000 ms. The output $o(n)$ and the just-noticeable phase difference (JND_ϕ) are computed after each input pulse. At the beginning of the comparison sequence the output $o(5)$ is at 0.851, indicating a fairly high level of synchrony between the oscillator and the input sequence. For $JND_{\phi_{\min}} = 0.03$ and $JND_{\phi_{\max}} = 0.5$, the output value of 0.851 specifies (by Equation 43) a just-noticeable phase difference JND_ϕ of 0.1. The first interval (second input pulse) of the comparison sequence generates a phase-difference of 0.14 (in the Figure, this corresponds to a decrease in the output). Since the phase difference of 0.14 is larger than the output-modulated threshold of 0.1, the model detects the change in tempo

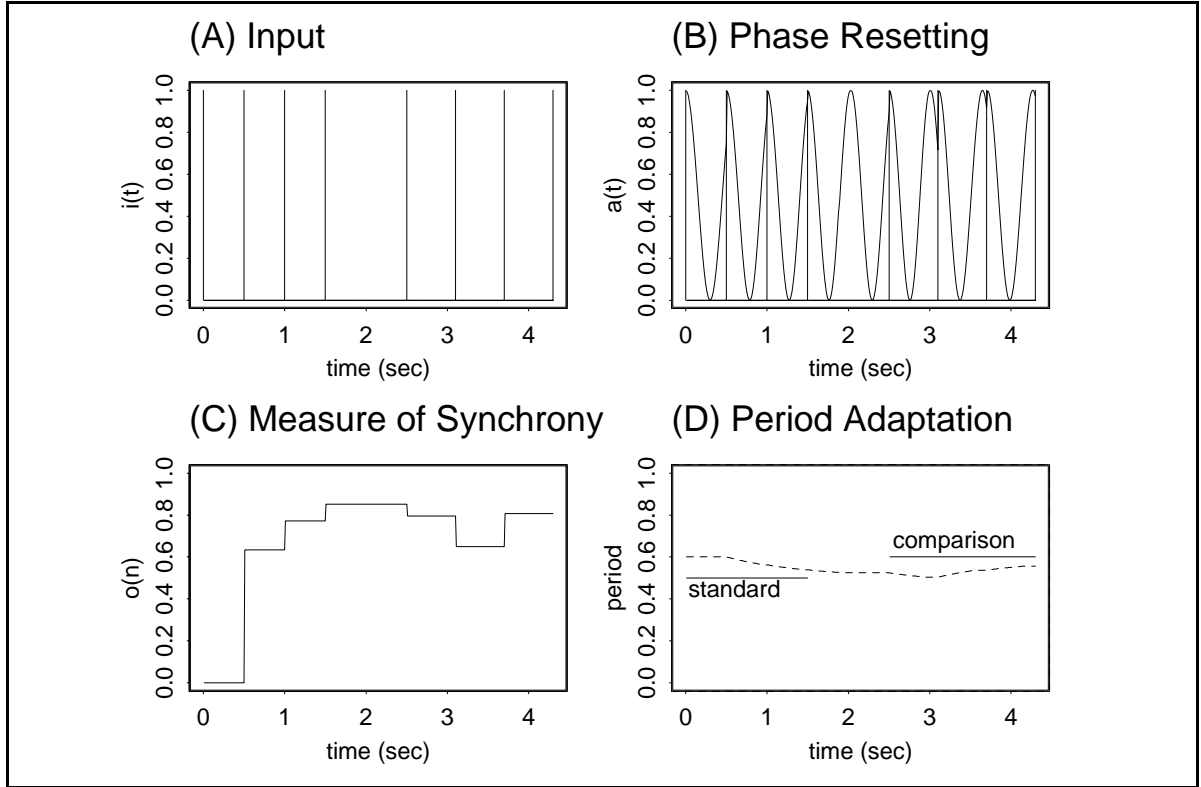


Figure 4.5: Response of adaptive oscillator to a simulated tempo-discrimination trial. The input sequence represents a four-tone standard sequence with a 500-ms IOI followed by a four-tone comparison sequence with a 600-ms IOI (a 20% slower tempo). The gap between the two sequences is 1000 ms, preserving the rhythm established by the standard.

and indicates that the standard sequence is faster, due to the *positive* phase difference detected.

Figure 4.6 illustrates the process of entrainment to a standard sequence in a comparison task using a phase-portrait description. Successive phase portraits show how the adaptive oscillator is entrained by a 500-ms standard sequence (top panel, identical to the simulated trial in Figure 4.5) and a 700-ms standard sequence (bottom panel). The adaptive oscillator's period reflects improved estimates of the standard sequence's IOI as it entrains, resulting in reduced over- and underestimation in the top and bottom panels respectively, . The solid black circle marks the beginning of the oscillator's cycle. Each successive input pulse resets the oscillator at a phase closer to zero phase, as shown by the degree of shading. As the reset phases approach 0.0, the outputs $o(n)$ approach 1.0 and the just-noticeable phase difference approaches its minimum value. In the top panel, the initial oscillator period of 600 ms is an overestimate of the 500-ms IOI. By the fourth input (the end of the standard sequence) the estimate has improved to 525 ms. In the bottom panel, the initial oscillator period of

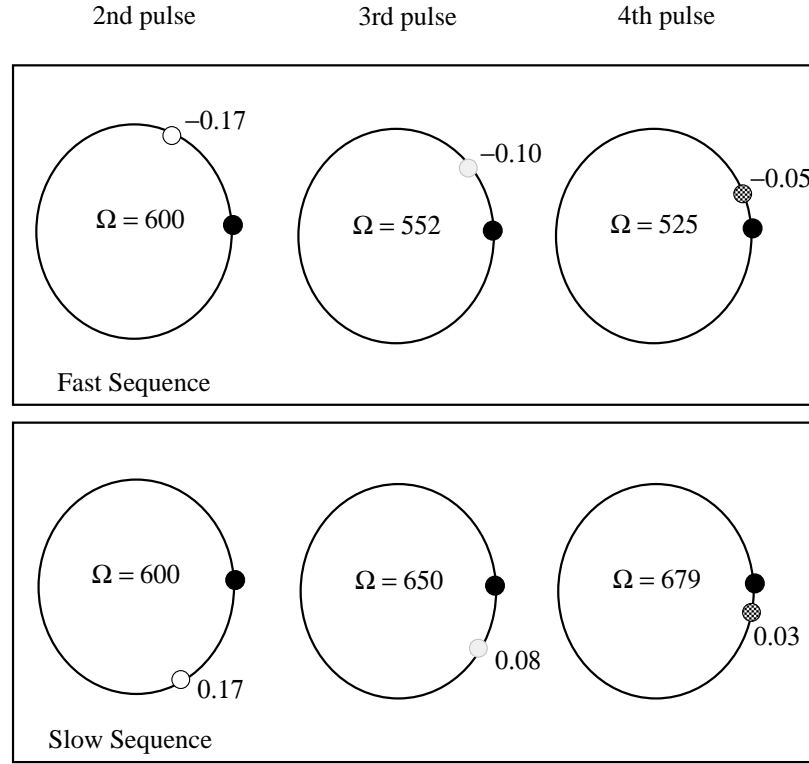


Figure 4.6: Phase portrait description of adaptive-oscillator entrainment by an input sequence representing a 4-tone standard sequence in a comparison task. The top and bottom panels demonstrate improved estimates of the standard’s IOI, for IOIs of 500 and 700 ms, respectively. In both cases, the solid black circle marks the beginning of the oscillator’s cycle. Each successive input pulse resets the phase of the oscillator at a phase closer to its zero-phase, as shown by the degree of shading.

600 ms is an underestimate of the 700-ms IOI. In this case, the subjective estimates improve to 679 ms by the end of the standard sequence.

Figure 4.7 shows the process of detecting a tempo change during the presentation of the comparison sequence. The phase-portrait description illustrates the state of the adaptive oscillator after the first interval (second input pulse) of the comparison sequence with a 600 ms IOI (20% slower than the standard sequence in this case). In Figure 4.6, the oscillator’s estimate is a 5% overestimate of the standard sequence’s IOI of 500 ms. Thus, a “slower” comparison sequence with an inter-onset-interval between 500 and 525 ms will trigger negative phase-differences, falsely indicating “faster” instead of “slower” (if such phase error is large enough and the just-noticeable phase difference is small enough). Analogously, if the model instead underestimated the standard sequence’s IOI by 5%, a comparison sequence with an IOI between 475 and 500 ms could trigger positive phase differences, falsely indicating “slower” instead of “faster”. Thus, in the Entrainment Model, over- and underestimation of time can

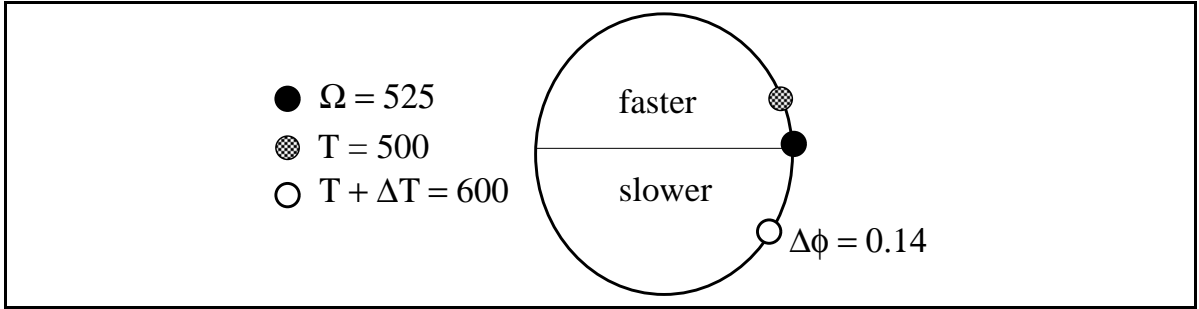


Figure 4.7: Phase-portrait description of the detection of a tempo change during the presentation of the comparison sequence. The comparison interval ($T + \Delta T = 600\text{ms}$) marked by the first two input pulses of the comparison sequence results in a phase difference of $\Delta\phi = 0.14$ (marked by the open circle). Since this phase difference is larger than the just-noticeable phase-difference of 0.1 computed from the standard sequence, the tempo change is detected. Since the phase difference is positive, the tempo of the comparison sequence is perceived as slower than that of the standard.

result in systematic errors in tempo discrimination; i.e., even if a tempo change is detected with respect to $|\phi| > JND_\phi$, the perceived *direction* of the tempo change may be incorrect. This prediction concerning differential sensitivity to increases and decreases in tempo is examined in detail in the second simulation

4.2.2 Simulation 1: Duration & Number of Intervals

Simulation 1 addresses the model's predictions concerning the influence of the number of sequence intervals and the sequence duration on tempo sensitivity for different sequence tempos. The simulation data will be compared directly with the tempo data from Drake and Botte (1993; 1994) and Michon (1964) to address three main issues. First, it was discussed in Chapter 2, that listeners sensitivity to the tempo of an isochronous sequence improves with the number of intervals in that sequence.

Second, tempo sensitivity improves more, with increasing number of intervals, for IOIs shorter than about 300 ms, than for IOIs longer than about 300 ms. This differential improvement in tempo sensitivity results in an extension of the optimal zone of tempo sensitivity to a shorter IOI (Drake and Botte, 1993; Drake and Botte, 1994; Michon, 1964). As discussed in Chapter 2, many researchers have suggested that this potentially heightened sensitivity implies that the processing of intervals shorter than about 300 ms is qualitatively different from the processing of intervals longer than 300 ms, possibly involving distinct mechanisms (Drake and Botte, 1993; Michon, 1964; Hirsh et al., 1990; ten Hoopen et al., 1994).

Third, it was discussed that the maximum number of intervals which lower the relative JND multiplied by the base IOI (which I termed the optimal sequence duration in Chapter 2) is constant at about 1.0 seconds for IOIs shorter than about 300 ms, but

shifts abruptly to about 2.5 seconds for IOIs longer than about 300 ms (Drake and Botte, 1994), providing some additional evidence which suggests that time intervals shorter than 300 are processed differently than those longer than 300 ms.

Method

The stimulus set for Simulation 1 consisted of isochronous sequences with IOIs of 100, 300, 500, 700, 900, 1100, 1300, and 1500 ms. Keeping with the convention established by Michon (1964), I have defined tempo in terms of inter-onset-interval, instead of the number of events per minute. Standard sequences consisted of 1 to 20 intervals. Each input pulse had an amplitude of 1.0, representing a tone onset. On each simulation trial, the model was presented with a standard sequence followed by a comparison sequence that was slightly faster ($\text{IOI} - \Delta\text{IOI}$) or slower ($\text{IOI} + \Delta\text{IOI}$) than the standard. The interval-pattern-interval between the onset on the last input of the standard sequence and the onset of the first input of the comparison sequence was twice the IOI of the standard sequence (i.e., twice the “expected” interval based on an extension of the periodicity of the standard). The model’s task was to judge, using the decision procedure described in Section 4.2.1, which sequence was faster.

On each trial a comparison sequence was randomly selected that was either faster or slower than the standard sequence. Since the specified model is deterministic, the model should always make either a correct or incorrect response for a fixed-percentage tempo change (regardless of direction), unless there is difference between detecting tempo increases and tempo decreases, due to $|\Delta\phi_+| \neq |\Delta\phi_-|$. With such an effective threshold difference between detecting tempo increases and tempo decreases, the model potentially exhibits systematic errors (e.g., always producing a correct response for a 10% tempo increase, while at the same time always producing an incorrect response for a 10% tempo decrease). To test this possibility while maintaining the same experimental paradigm used by Drake and Botte (1993), discrimination thresholds were obtained for each interval condition and for each investigated tempo condition using an adaptive-tracking procedure (Levitt, 1971). Each threshold measurement was based on 100 simulated trials (one block). At the start of each block, the tempo difference between the standard and comparison sequences was set to 20%. If the model made two correct responses, the level was decreased by 1.0%. If the model made a single incorrect response the level was increased by 1.0%. By iterating this procedure, the level converges to a value corresponding to 70.7% correct responses; i.e., the just-noticeable difference in tempo. For each tracking history, relative JND (which reflected an average of the thresholds for tempo increases and tempo decreases) was determined by averaging the tempo differences of the the last 50 trials of each block. Threshold measurements were repeated 20 times for each condition to obtain an average measure of relative JND. Since the reported version of the model is deterministic, the only reason for repeating the measurement procedure 20 times for each condition is to average out potential differences in measured relative JNDs due to the

specific order of tempo-increase and tempo-decrease trials.

Results

The mean relative JNDs obtained in this simulation are shown in Figure 4.8 for all tempo conditions and for 1-, 2-, 4-, and 6-interval sequences. This specific set of interval conditions was chosen for direct comparison with the Drake and Botte tempo data (Drake and Botte, 1993). In order to illustrate the effect on performance of

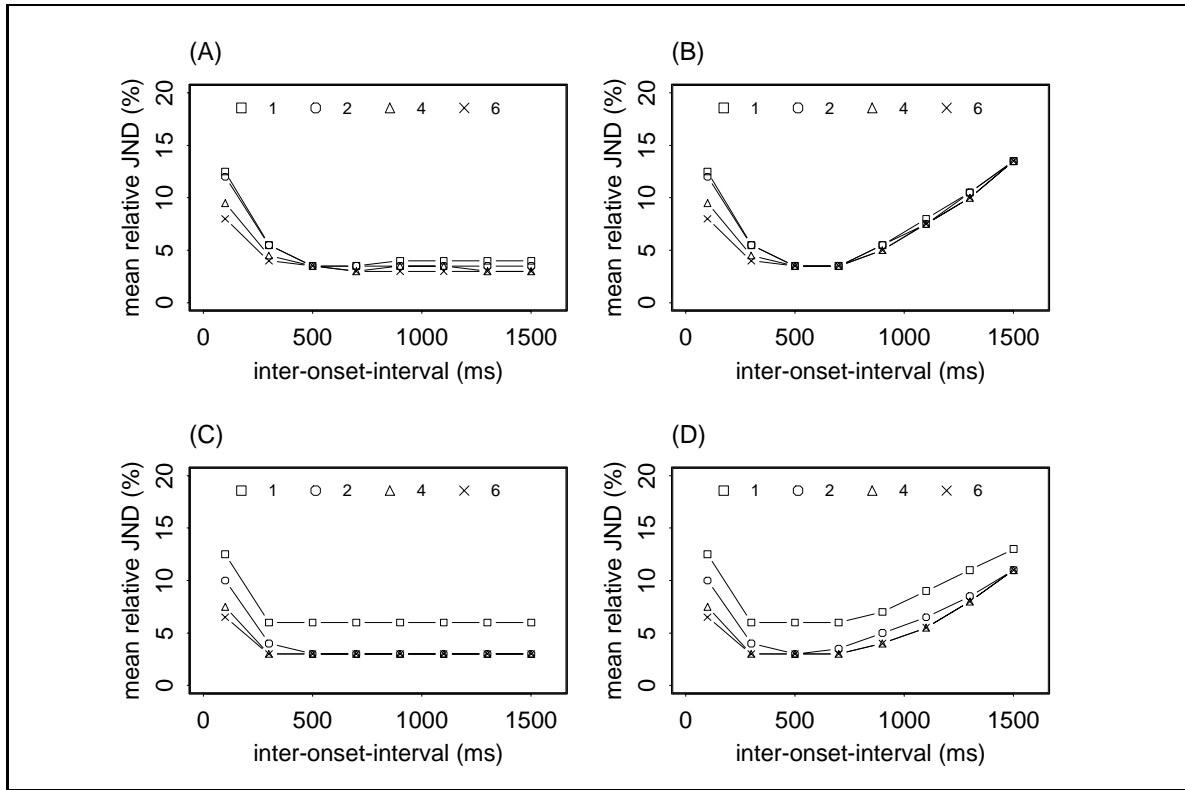


Figure 4.8: The influence of the number of sequence intervals on the model's tempo sensitivity for 8 different tempo conditions. Data for four different sets of parameter values is shown. In (A), $\alpha = 4.0$, $\beta = 0.0$, $w_o = 1.0$, and $\Theta = 0.57$. In (B), $\alpha = 4.0$, $\beta = 0.005$, $w_o = 1.0$, and $\Theta = 0.57$. In (C), $\alpha = 4.0$, $\beta = 0.0$, $w_o = 0.75$, and $\Theta = 0.57$. In (D), $\alpha = 4.0$, $\beta = 0.005$, $w_o = 0.75$, and $\Theta = 0.57$.

varying the free parameters of the model, each of the four graphs depicts the model's performance for a different set of adaptive-oscillator parameter values (as described in the Figure caption). In all of the graphs, $\alpha = 4.0$, $\Theta = 0.57$, and $JND_{\phi_{\min}} = 0.03$. Assuming in the limit that $\Omega = T$, the minimum just-noticeable phase difference of 0.03 corresponds to a relative JND (or $\frac{\Delta T}{T}$) of 3.0%. For $\Theta = 0.57$, the temporal placement of the entrain/decay boundary (IOI_{Θ}) is approximately 1100 ms. The remaining two free parameters, w_o and β , vary systematically from graph to graph.

In Graph A, the decay rate is zero ($\beta = 0.0$) and the output weight is at its maximum value ($w_o = 1.0$). With $w_o = 1.0$, the adaptive oscillator uses only the current phase reset to determine its output. Thus, the just-noticeable phase difference (which is modulated by the the output) is specified by each successive phase reset. As can be seen in Graph A, in this minimal version of the model (without decay or phase memory), the relative JND is fairly constant for IOIs between 300 and 1500 ms (in agreement with Weber's law) but increases abruptly for those shorter than 300 ms. Increasing the number of intervals does improve the relative JND, but much more dramatically for IOIs shorter than 300 ms than those longer than 300 ms. These model data are similar to the experimental data discussed in Chapter 2, which researchers argued support different processing of time intervals shorter than about 300 ms from the processing of those longer than about 300 ms. (Drake and Botte, 1993; Hirsh et al., 1990; Michon, 1964; Schulze, 1989; ten Hoopen et al., 1994)

However, the model data show that this assumption of distinct processing of short and long IOIs is not necessary to produce such a result. The Entrainment Model uses a single mechanism that is unchanged for fast and slow sequences. The observed differences between the model's performance on short and long IOIs, shown in Graph A, are due entirely to the dynamics of entrainment. That is, in this case, the longer IOIs provide a longer period of time to entrain (in between input pulses) relative to the shorter IOIs. For the longer IOIs maximum sensitivity is attainable by the Entrainment Model in one or two cycles, but for the shorter IOIs, maximum sensitivity is only attained by the Entrainment Model after many cycles. This issue is further explored in Graph B.

In Graph B, all parameters are the same as in Graph A, except decay which is now non-zero ($\beta = 0.005$). The pattern of results in this case is the same as in Graph A, except that the relative JND curve is now U-shaped, with approximately three zones of tempo sensitivity, as reported by Drake and Botte (1993): (1) a zone of maximum sensitivity between about 300 and 900 ms; (2) a zone of reduced, but potentially heightened sensitivity for IOIs shorter than 300 ms for which increasing the number of sequence intervals reduces the relative JND; and (3) a zone of lesser sensitivity for IOIs longer than about 900 ms. This increase in relative JNDs for IOIs longer than 900 ms is due to the addition of period decay. The 900-ms "boundary" between zone (1) and zone (3) is due to the temporal placement of the entrain/decay "boundary" (IOI_θ) at about 1100 ms. For IOIs longer than 1100 ms, period decay begins to have a significant effect. For IOIs shorter than 1100 ms, each onset arrives before the entrain/decay boundary and thus the effect of period decay is much less.

In Graph C, all parameters are the same as in Graph A (including zero decay), except that the output weight is decreased ($w_o = 0.75$). Decreasing w_o , causes the output $o(n)$ to reach the asymptote at 1.0 more slowly. That is, the model gradually accumulates evidence of its synchronization with the standard sequence based on a memory of the previous outputs as well as the current reset phase. The output $o(n)$ can be viewed as a measure of the model's confidence. Thus, by decreasing the output

weight, the model's confidence is built up more slowly. In addition, changes in the just-noticeable phase difference accrue more slowly, causing the just-noticeable phase difference to reach its minimum value $JND_{\phi_{\min}}$ later. Since $JND_{\phi_{\min}}$ is reached later, the improvement due to increasing the number of intervals is much more pronounced than in Graph B for all of the IOI conditions. For all of the IOIs except the shortest (IOIs < 300 ms), most of the improvement in the relative JND occurs between one and two intervals. However, for the shortest IOIs, relative JNDs continue to decrease even for the 6-interval sequences.

Graph D combines the parameter change in Graph B (non-zero decay, resulting in an entrain/decay boundary) and the parameter change in Graph C (output weight less than 1.0, resulting in a more gradual improvement in tempo sensitivity). Including both period decay and a gradual output function, turns out to be critical for providing a comprehensive explanation of the Drake and Botte tempo data (Drake and Botte, 1993; Drake and Botte, 1994). Consistent with the Drake and Botte (1993) data (see Figure 2.7 for comparison), the Entrainment Model shows a U-shaped relative JND curve with the same three zones of tempo sensitivity specified in the discussion of Graph C. In addition, identical to the Drake and Botte (1993) tempo data, the Entrainment Model's improvement in tempo sensitivity is most dramatic for IOIs shorter than 300 ms, whereas *no* improvement in tempo sensitivity is observed for the 1500-ms IOI, after increasing the standard sequence to two intervals.

In order to account for their data with the Multiple-Look Model, Drake and Botte (1993) assumed that listeners only integrate timing information within a limited temporal window. Thus, with a fixed temporal window (d), the number of useful intervals (n) for any IOI condition is determined by $\frac{d}{IOI}$. Thus, for short IOIs, the number of useful "looks" is larger than for long IOIs. To determine listeners temporal window size (d), Drake and Botte (1994), as discussed in Chapter 2, specifically investigated the extent to which increasing the number of intervals improves discrimination thresholds. For a range of IOIs, they determined the interval number (n) for which an additional interval failed to lower the relative JND (i.e., the addition of the interval exceeded the listener's temporal window). In Chapter 2, I identified this duration ($d = nIOI$) as the *optimal sequence duration*. For IOIs < 300 ms, Drake and Botte (1994) reported an optimal sequence duration of approximately 1 sec, whereas for IOIs > 300 ms, they reported an optimal sequence duration of approximately 2.5 sec. Drake and Botte suggested that the discontinuity in these data provide even stronger evidence that short intervals are processed by a different mechanism than long intervals.

In order to test this claim, optimal sequence durations were determined for the Entrainment Model for the parameter settings in Graph D. These data were then compared with the "temporal window" data reported by Drake and Botte (1994) (as shown in Figure 4.9). In agreement with the Drake and Botte data, the optimal sequence duration of the model varied as a function of tempo with an abrupt transition from about 1.0 second to 2.5-3.0 seconds at the 500-ms IOI. Thus, this modeling

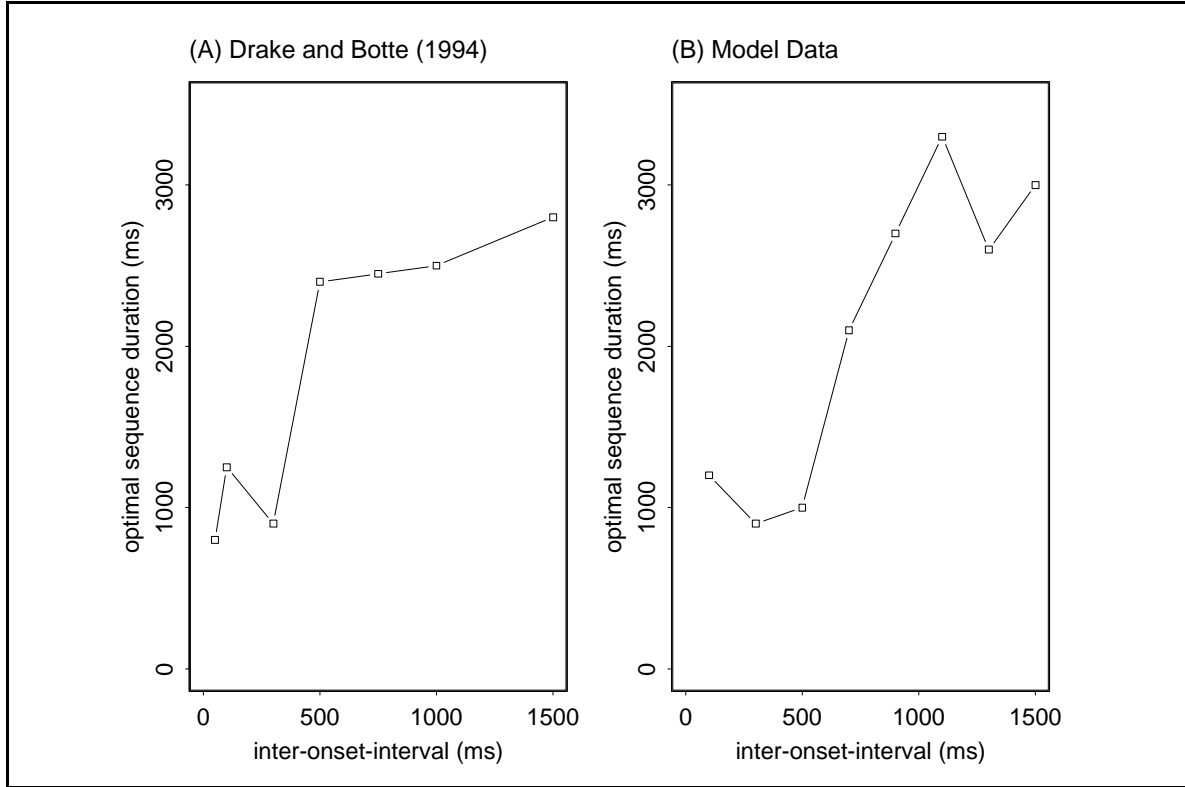


Figure 4.9: Interpretation of model performance in terms of optimal sequence duration (or temporal window size according to Drake and Botte (1993; 1994)).

result shows that it is not necessary to posit distinct temporal window sizes in order to explain listeners' abilities to discriminate tempo. Instead, these tempo data can be explained in terms of the adaptive-oscillator-based Entrainment Model, in which performance is based on the interaction of two dynamic processes: period entrainment and decay. The corresponding optimal JNDs obtained for the Entrainment Model are shown in Figure 4.10 in comparison with the optimal relative JNDs from Drake and Botte (1994) and the earlier tempo data from Michon (1964).

4.2.3 Simulation 2: Direction of Tempo Change

Previous models of time perception, including, Drake and Botte's Multiple-Look Model (Drake and Botte, 1993; Drake and Botte, 1994) and Jones and Boltz's Contrast Model, have assumed no difference between the relative JNDs for increases in tempo ($-\Delta IOI$) and those for decreases in tempo ($+\Delta IOI$). In the specification of the Entrainment Model in Section 4.1, it was shown that the relationship between the JND and JND_ϕ is influenced by the amount of under- or overestimation of the base interval (T). Consequently, even if a tempo change is detectable ($|\Delta\phi| > JND_\phi$), the perceived direction of the tempo change (increase or decrease) may be incorrect (i.e.,

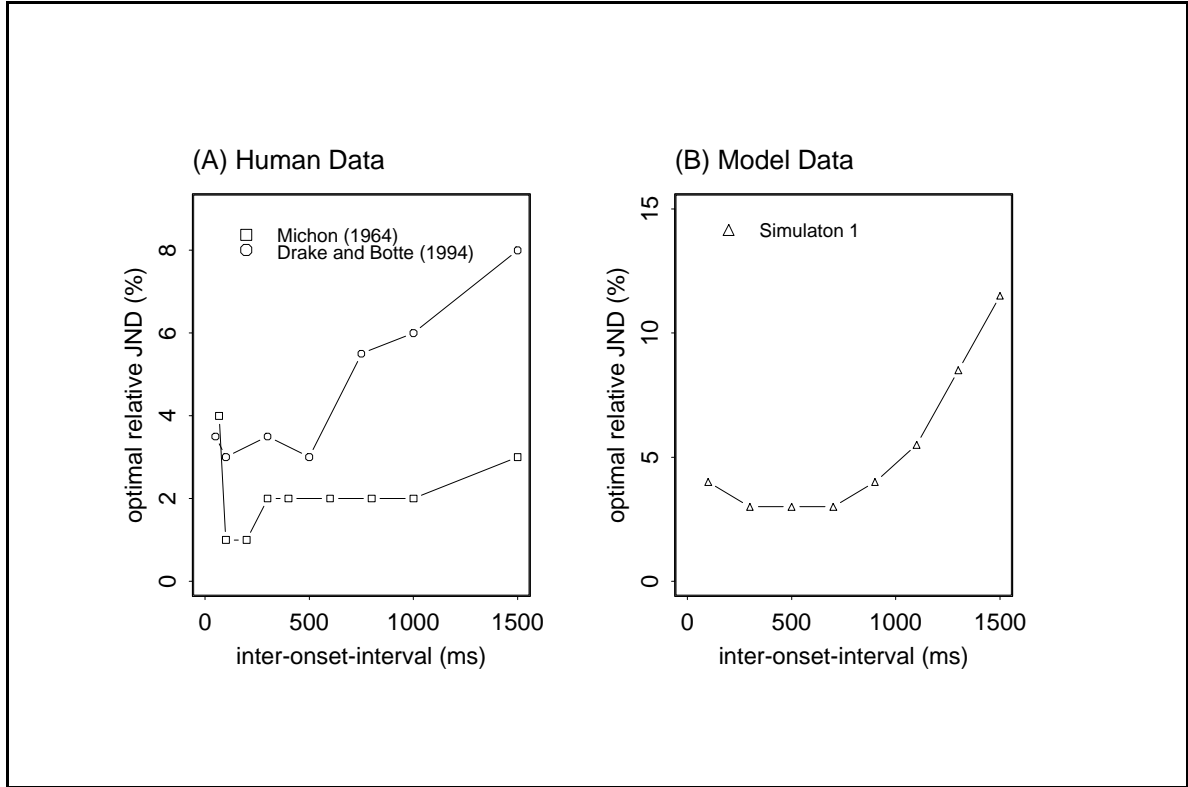


Figure 4.10: Maximum human and model tempo sensitivity for increases in the number of sequence intervals.

the sign of the phase change $\Delta\phi$ may not correctly specify whether the comparison sequence is faster or slower than the standard sequence). Thus, distortions in the relationship between JND and JND_ϕ due to $\frac{T}{\Omega} > 1.0$ (Case 2) or $\frac{T}{\Omega} < 1.0$ (Case 3) predict systematic asymmetries in tempo discrimination, as suggested by Simulation 1. Simulation 2 specifically addresses the predicted differences in relative JNDs for increases and decreases in tempo.

Method

The stimulus set used in Simulation 2 was the same as in Simulation 1. Relative JNDs were established separately for increases and decreases in tempo, for each interval number and IOI condition, by gradually reducing the tempo difference ($+\Delta IOI$ or $-\Delta IOI$), in 1% steps, until the model made an incorrect response. The relative JND for a decrease in tempo was the smallest $+\Delta IOI$ that the model was able to detect, and the relative JND for an increase in tempo was the smallest $-\Delta IOI$ that the model was able to detect. Thus, the combined measure of relative JND obtained in Simulation 1 via an adaptive-tracking procedure which assumed no difference between thresholds for detecting an increase or decrease in tempo, should be an average of of

the separate measures of relative JND obtained here.

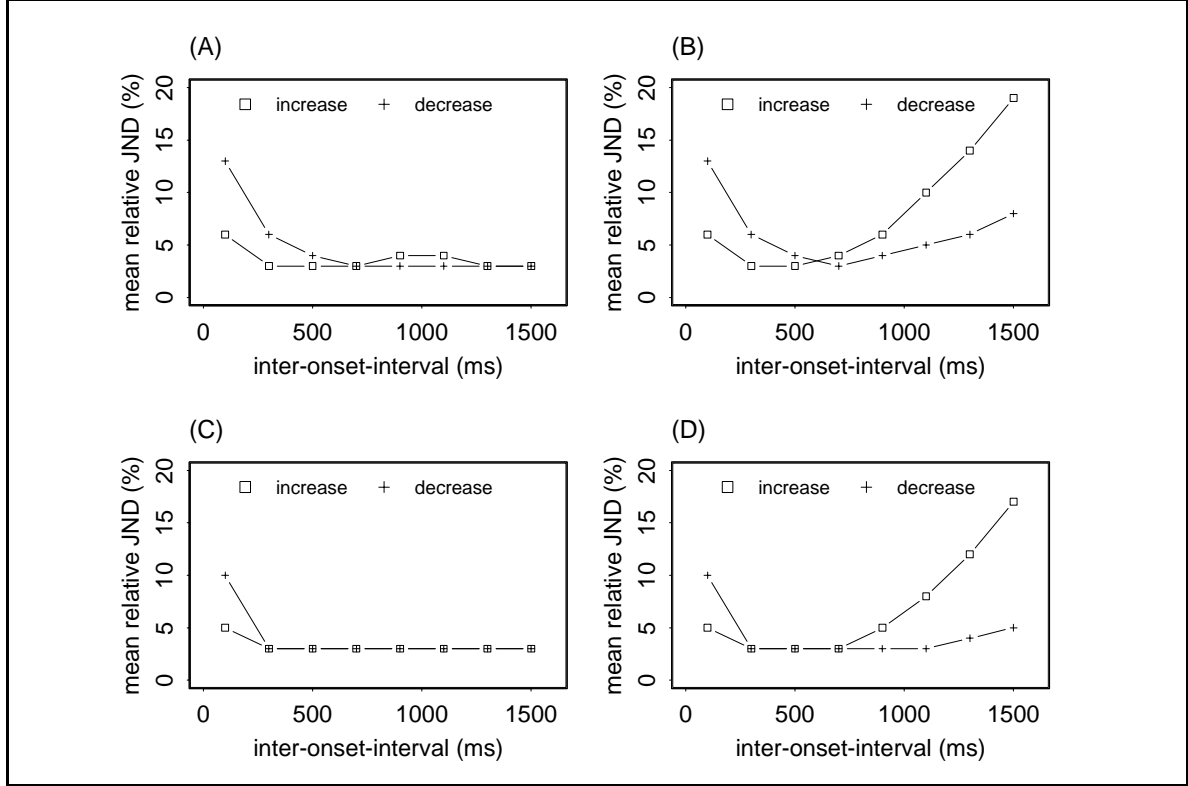


Figure 4.11: Differences between thresholds for detecting increases and decreases in tempo for the same four sets of parameter values as in Simulation 1.

Results

Mean relative JNDs for tempo increases and tempo decreases are shown in Figure 4.11, for the four-interval sequences only, for the same four sets of parameter values used in Simulation 1. Relative JNDs for the tempo increases and the tempo decreases are indicated by the open squares and the crosses, respectively.

The main feature of all four graphs is that for shorter IOIs, the relative just-noticeable tempo increase is lower than the relative just-noticeable tempo decrease. Similarly, in Graphs B and D for longer IOIs, the relative just-noticeable tempo decrease is lower than the relative just-noticeable tempo increase. In all cases, there is no difference in tempo sensitivity for an intermediate range of IOIs. For Graphs B and D, differential sensitivity to increases and decreases in tempo results in a crossing pattern of relative JNDs.

For each of the Graphs, the observed differences in relative JNDs for increases and decreases in tempo are due to the amount of over- or underestimation of the standard sequence's IOI. Since for all cases, short IOIs are initially overestimated and long

IOIs are initially underestimated, the relative JND curves for increases and decreases in tempo are initially in a crossing pattern. However, as the Entrainment Model successfully reduces the amount of over- or underestimation, the difference between the thresholds for detecting increases and decreases in tempo is also reduced. With no decay (Graphs A and C), the model becomes perfectly entrained by the standard sequence for IOIs longer than about 300 ms ($\Omega = \text{IOI}$), but not for IOIs shorter than about 300 ms ($\Omega > \text{IOI}$). Thus, for these cases, relative JNDs for increases and decreases in tempo are the same for IOIs longer than about 300 ms, but not for IOIs shorter than about 300 ms. With decay (Graphs B and D), the model is unable to be perfectly entrained by the longest IOIs (at least for four-interval standard sequences). For these cases, relative JNDs for tempo increases and tempo decreases are different at both the shortest and longest IOIs, with tempo increases easier to detect at the short IOIs and tempo decreases easier to detect at the long IOIs, resulting in a crossing relative JND pattern.

Thus, the Entrainment Model with decay (which was found to be necessary for modeling the data in Simulation 1) predicts that, for fast sequences, listeners should be more sensitive to a tempo change for a comparison sequence that is faster than the standard sequence; whereas, for slow sequences, listeners should be more sensitive to a tempo change for a comparison sequence that is slower than the standard sequence. However, as the listener becomes successfully entrained by the standard sequence (i.e., with more intervals in standard sequence), differences between thresholds for detecting increases and decreases in tempo should be diminished. If supported by experimental data, this prediction places strong constraints on possible models. In particular, if experimental data shows threshold differences for increases and decreases in tempo with the same crossover pattern as predicted by the Entrainment Model, then (1) the class of multiple-look models (Drake and Botte, 1993; Drake and Botte, 1994; Schulze, 1989) would no longer seem to be a viable approach to tempo perception, and (2) models of time perception based on dynamic-attending theory such as the Jones and Boltz (1989) Contrast Model would require an adaptive oscillator similar to the one proposed in this thesis in order to display appropriate behavior. At present, the Contrast Model assumes no difference between relative JNDs for tempo increases and tempo decreases.

4.2.4 Simulation 3: Temporally-Directed Attending

In studies of tempo discrimination a fixed gap is typically maintained between the onset of the last event of the standard sequence and the onset of the first event of the comparison sequence. The duration of this silent interval is usually a multiple of the standard sequence's IOI, thus extending the periodicity of the standard sequence. In terms of the temporal structure of the 2AFC tempo-discrimination task, one implication of the entrainment hypothesis is that the tempo of the comparison sequence should be less well resolved when the comparison sequence arrives 'out-of-phase' with

respect to the periodicity established by an isochronous standard. To this end, Simulation 3 examines the Entrainment Model's predictions concerning discrimination thresholds for systematic 'out-of-phase' variations in the onset of the comparison sequence.

Method

In Simulations 1 and 2, the gap between the last input pulse of the standard sequence and first input pulse of the comparison sequence (i.e., the inter-pattern-interval or IPI) was twice the IOI of the standard sequence. To examine the effects of temporally-directed attending on the tempo-discrimination thresholds of the Entrainment Model, relative JNDs were determined for a range of inter-pattern-interval conditions for three-interval isochronous standard sequences for IOIs of 100, 300, 500, 700, 900, 1100, 1300, and 1500 ms. The inter-pattern-interval (IPI) varied between 100% of the IOI of the standard sequence to 300% of the IOI of the standard sequence in 10% steps. In terms of phase, each 10% step is equivalent to a phase-step of 0.1 (one tenth of standard's IOI). Thus, for IPIs of 150%, and 250% of the standard sequence's IOI, the comparison sequence's arrived 180 degrees out-of-phase (a phase of 0.5) with respect to its expected temporal location (zero phase) established by the IOI of the standard sequence. Discrimination thresholds were determined for each combination of IOI and IPI using the adaptive-tracking procedure described in Simulation 1. Like Simulation 2, relative JNDs for increases and decreases in tempo were measured independently. Thus, the reported relative JNDs are an average of the relative JNDs obtained for increases and decreases in tempo for the respective condition.

Results

The relative JNDs obtained in this simulation are shown in Figure 4.12 for all IPIs (phase conditions) and for all IOIs. Only the parameter values from Simulation 1 that gave the "best fit" to the Drake and Botte data were tested (i.e., those used to obtain data in Panel (D) in Figure 4.8). These data are displayed slightly differently than those shown for the previous two simulations. The onset phase of the comparison sequence is represented on the x -axis as a percentage of the standard sequence's IOI (i.e., the units on the x -axis are independent of the standard sequence's IOI). As in the previous simulations, relative JND is represented on the y -axis. Each line corresponds to a different IOI for the standard sequence, as labeled in the legend.

There are three primary properties of the Entrainment Model to identify in this graph. First, for all of the IOIs, relative JND varies as a sinusoidal function of the onset phase of the comparison sequence. Relative JNDs are lowest when the onset of the comparison sequence is expected (i.e., occurs at a multiple of the standard's IOI: 100%, 200%, or 300%) and highest when the onset of the comparison sequence is most unexpected (i.e., occurs 180 degrees out-of-phase with the expected periodicity of the

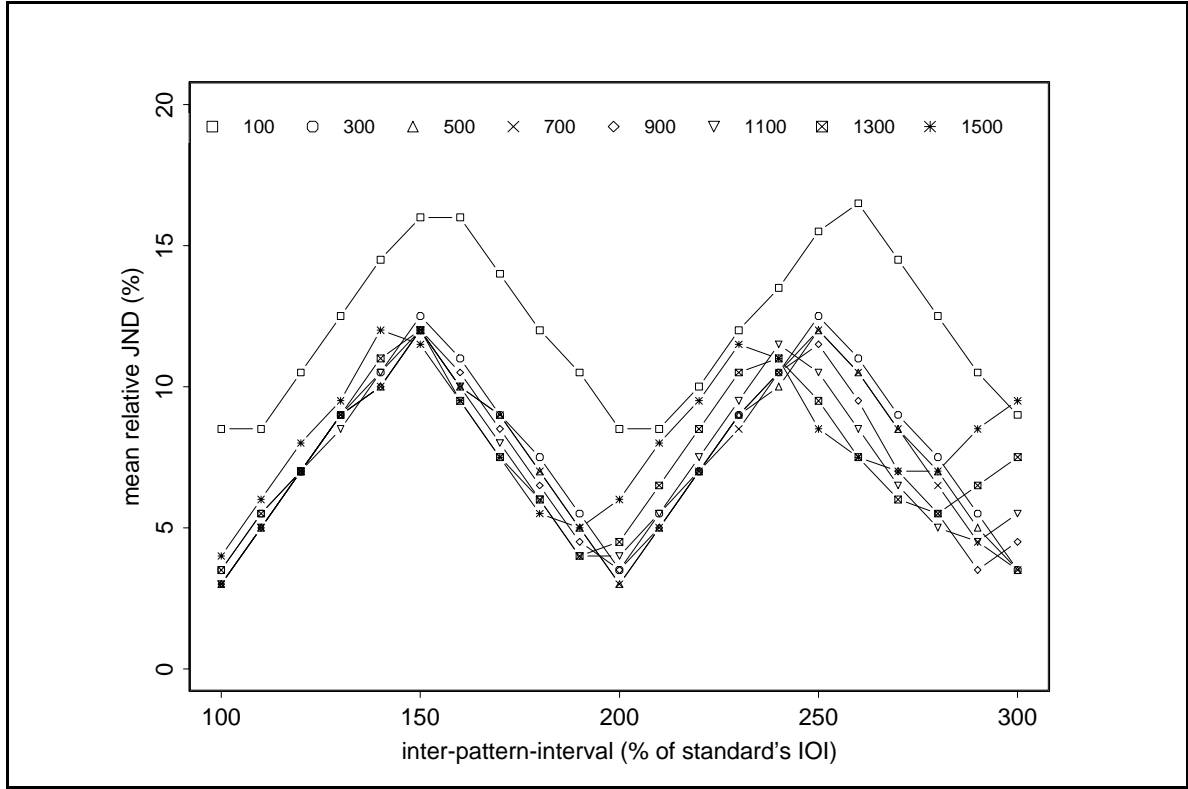


Figure 4.12: The effect of variations in the inter-pattern-interval on the model's tempo sensitivity: Each line indicates a different IOI condition, as specified.

standard: 150% or 250% of the standard sequence's IOI). Second, for longer IOIs, the sinusoidal-shaped-relative-JND curve is slightly phase-shifted towards shorter IPI conditions. That is, the lowest relative JNDs occur for IPIs that are slightly less than 100%, 200%, or 300% of the standard sequence's IOI. This phase-shift in the “expected” onset of the comparison sequence is due to the underestimation that occurs for the longer IOIs. For example, if the model underestimates the standard's IOI by 10%, the “expected” onset of the comparison sequence, based on multiples of the subjective IOI, will be 90%, 180%, or 270% of the actual IOI of the standard. Third, as the *absolute* magnitude of the IPI increases, the discrimination performance of the model worsens for corresponding phase offsets. For example, performance is worse overall for the 200% condition than for the 100% condition. The degraded performance observed as a function of the magnitude of the IPI is due to the gradual “decay” of the oscillator's period back to its resting value. In addition, performance degrades more for longer IOIs than for shorter IOIs because of the temporal placement of the entrain/decay boundary. Thus, in addition to predicting that tempo differences between the standard and comparison sequences should be better resolved when the onset of the comparison sequence is at an “expected” temporal location than when it

is at an “unexpected” temporal location, the Entrainment Model predicts that short-term memory for the tempo of a sequence should be better for tempos close to the resting rate of the tracking adaptive oscillator than for tempos that are not.

4.3 Model Summary

In this chapter, I specified an Entrainment Model of time perception based on the proposed adaptive oscillator. The main purpose of developing this model was to evaluate its predictions concerning factors which influence listeners’ abilities to detect changes in the tempo of isochronous sequences. The modeling efforts (in Simulation 1) focused on a set of human data from several tempo-discrimination experiments (Drake and Botte, 1993; Drake and Botte, 1994; Michon, 1964) which were difficult to explain with the Multiple-Look Model and not intended to be explained by the Contrast Model. The Entrainment Model provided a parsimonious explanation for these data in terms of the interaction of the entrainment and period-decay processes. This modeling work was motivated by the belief that in order to explain the processing mechanisms underlying complex rhythmic behaviors, we must first be able to explain the mechanism responsible for the perception of simple rhythms, such as isochronous sequences. It should be apparent to the reader, that isochronous sequences provide a rich, yet simple, test bed to probe the nature of human tempo perception. This chapter closes with a summary of the Entrainment Model, enumerating its assumptions and predictions.

Assumptions

1. The Entrainment Model assumed a linear psychophysical law for time for which short intervals are overestimated and long intervals are underestimated with respect to an intermediate indifference interval. The resting period of the adaptive oscillator, *initially* determined by the linear mapping, provided the model with subjective estimates of isolated time intervals.
2. The Entrainment Model assumed that listeners’ internal representation of a time interval is as a phase angle. The relationship between the time difference (ΔT) and the phase difference ($\Delta\phi$) triggered by ΔT was determined by the amount, the adaptive oscillator over- or underestimates the time interval T . Only when $\Omega = T$ did a zero time-difference ($\Delta T = 0.0$) correspond to a zero phase-difference ($\Delta\phi = 0.0$).
3. The Entrainment Model assumed that there is a phase analog of the just-noticeable time difference (JND), which I termed the just-noticeable phase difference (JND_ϕ). The just-noticeable phase difference specified the smallest detectable time difference ΔT in a base interval T , within the context of the adaptive oscillator tracking that time interval.

4. The Entrainment Model assumed that the just-noticeable phase difference JND_ϕ decreases as a listener becomes entrained by the stimulus pattern. This assumption was incorporated into the following rule

$$JND_\phi = JND_{\phi_{\max}}[1 - o(n)] + o(n)JND_{\phi_{\min}}$$

in which the output $o(n)$ of the adaptive oscillator modulates the just-noticeable phase difference between a maximum ($JND_{\phi_{\max}}$) and a minimum ($JND_{\phi_{\min}}$) value.

5. The Entrainment Model's tempo judgments were derived in a four step process intended to model the process by which listeners make similar judgments. First, listeners' tempo judgments were based on entrainment to the standard sequence (measured by $o(n)$). Second, the detection of a tempo difference was *immediate* following the first *different* interval ($T + \Delta T$) of the comparison sequence, with the first different interval establishing the phase difference $\Delta\phi$. Third, if the phase difference ($\Delta\phi$) was greater than the just-noticeable phase difference (JND_ϕ), then the tempo difference was detected, otherwise it was not. Fourth, the sign of $\Delta\phi$ established the direction of the tempo difference (i.e., whether the comparison sequence was faster or slower than the standard).

Predictions Based on the above assumptions, the Entrainment Model made the following predictions concerning human tempo sensitivity.

1. The Entrainment Model predicted that relative JND is U-shaped as a function of the IOI, with a zone of optimal sensitivity (smallest relative JNDs) centered on the indifference interval for the initial linear mapping. The width of the zone of optimal sensitivity depended on the values for the entrainment- and decay-rate parameters, the temporal placement of the entrain/decay boundary, and the initial degree of over- and underestimation specified by the linear mapping.
2. The Entrainment Model predicted that increasing the number of intervals in the standard sequence lowers relative JNDs for all IOI conditions, but especially for the shorter IOIs conditions. As a result, the zone of optimal sensitivity extended to shorter IOIs when the number of intervals in the standard sequence increased. For the reported simulation data, improvement in thresholds with increasing number of intervals was most dramatic for IOIs shorter than 300 ms.
3. The Entrainment Model predicted that the *optimal sequence duration* (which was defined for each IOI as the number of intervals which lower the relative JND multiplied by the IOI) is approximately sigmoidal. For the reported simulation data, the optimal sequence duration was about 1.0 second for IOIs shorter than 500 ms and was about 2.5-3.0 seconds for IOIs longer than 900 ms, with monotonically increasing values in between 500 and 900 ms.

4. The Entrainment Model predicted that the relative JNDs for tempo increases should be lower than those for tempo decreases for fast sequences (short IOIs); and conversely the relative JNDs for tempo decreases should be lower than those for tempo increases for slow sequences (long IOIs). However, as the listener is entrained by the standard sequence (through having more isochronous intervals), differences in tempo sensitivity for tempo increases and tempo decreases should be diminished.
5. The Entrainment Model predicted that tempo differences between the standard and comparison sequence should be better resolved when the onset of the comparison sequences occurs at an “expected” temporal location (based on the periodicity of the standard sequence) than when it occurs at a “unexpected” temporal location. Relative JNDs should be lowest when the onset of the comparison sequence occurs at a multiple of the standard sequence’s IOI (i.e., an inter-pattern-interval of 100%, 200%, 300%, etc, of the standard sequence IOI) and should be highest when the onset of the comparison sequence occurs 180 degrees out-of-phase with respect to the expected periodicity of the standard sequence (i.e., an inter-pattern-interval of 150%, 250%, etc, of the standard sequence IOI). Therefore, relative JNDs should vary as a sinusoidal function of the onset phase of the comparison sequence.
6. With regard to Prediction 5, the Entrainment Model predicted phase delays and phase advances in the sinusoidal function depending on the degree of over- or underestimation of the standard sequence’s IOI, respectively.
7. Also with regard to Prediction 5, the Entrainment Model predicted that short-term memory for the tempo of a sequence should be better for tempos close to the listener’s preferred rate than for tempos substantially faster or slower than the listener’s preferred rate.

Predictions 1, 2, and 3 of the Entrainment Model accounted for the tempo-discrimination data reported by (Drake and Botte, 1993; Drake and Botte, 1994; Michon, 1964). In order to even partially account for these data with the Multiple-Look Model, Drake and Botte assumed that listeners are only able to use multiple looks within a limited temporal window and that that this window has two sizes: 1.0 seconds for IOIs shorter than about 300 ms and 2.5 seconds for IOIs longer than about 300 ms. Even with these assumptions, Drake and Botte were unable to explain why increasing the number of intervals in the standard sequence improves thresholds by a much greater amount for the short IOIs than for the long IOIs. Concordantly, it has been repeatedly suggested that short and long intervals (or fast and slow sequences) are processed differently, perhaps by distinct mechanisms (Drake and Botte, 1993; Drake and Botte, 1994; Hirsh et al., 1990; Michon, 1964; Schulze, 1989; ten Hoopen et al., 1994). The Entrainment Model, however, provided a single-mechanism explanation for these data.

Predictions 4, 5, 6, and 7 of the Entrainment Model concerned relatively unexplored experimental territory, for which no human data was available for comparison. The following chapter, reports results from two listening experiments designed to test aspects of these predictions.

Chapter 5

Two Tempo Discrimination Experiments

5.1 Overview of the Listening Experiments

The two listening experiments presented in this chapter were designed to test predictions of the Entrainment Model regarding listeners' abilities to detect differences in the tempo of isochronous sequences, as discussed in the previous chapter. The first experiment compared listeners' abilities to detect increases and decrease in tempo, in order to test the model prediction that for short IOI conditions listeners' are better able to detect faster comparison sequences (tempo increases) than slower comparison sequences (tempo decreases), and for long IOI conditions, listeners' are better able to detect slower comparison sequences than faster comparison sequences. Moreover, as a listener is entrained by the standard sequence (e.g., by that sequence having more isochronous intervals), this differential sensitivity should be diminished. Testing this prediction required a slight modification to the “which is faster” paradigm used by Drake and Botte (1993), in order to obtain an unbiased measurement of relative JND. In the present experiment, listeners heard a standard sequence followed by two comparison sequences (instead of one) and judged which of the comparison sequences was different in tempo from the standard, as illustrated in Figure 5.1.

If the “which is faster” task with a single comparison sequence was used with adaptive tracking instead of the “which is different” task with two comparisons, then listeners could adopt the fixed strategy of either responding that the comparison sequence was faster or that the standard sequence was faster. As a result, such a listener would either always correctly detect tempo increases or always correctly detect tempo decreases, introducing a response bias, since separate tracks are maintained for tempo-increase and tempo-decrease trials. Using two comparisons in the “which is different” task eliminates this possibility.

Experiment 2 tested the model prediction that tempo differences between the standard and comparison sequences should be better resolved when the onset of the

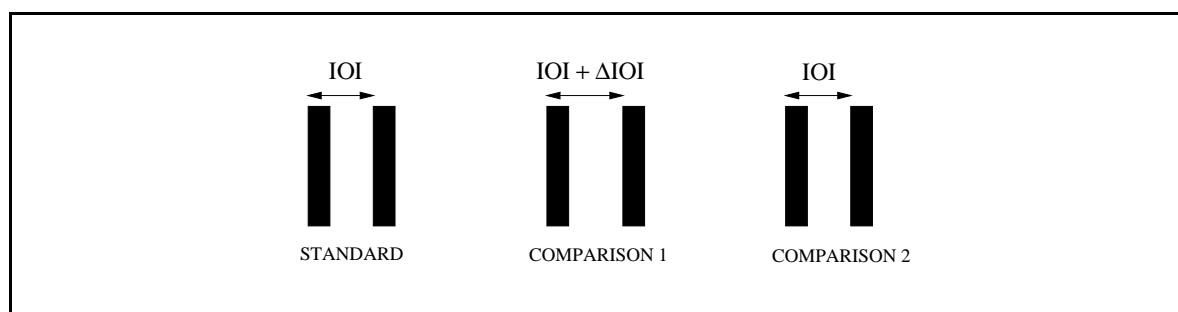


Figure 5.1: The “which is different” task used in Experiment 1. It is illustrated here for 1-interval standard and comparison sequences.

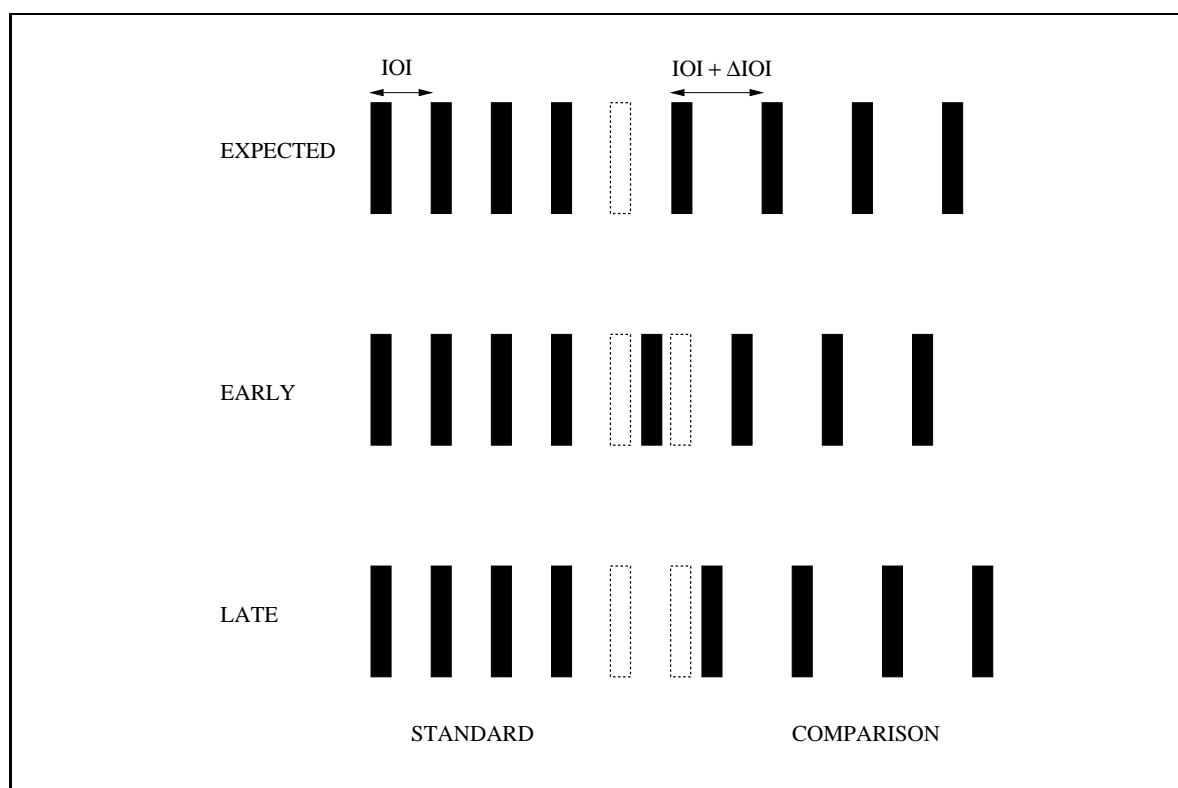


Figure 5.2: The “which is faster” task for conditions in which the onset of the comparison sequence is at “early”, “late”, and “expected” temporal locations defined by $2 \times \text{IOI}$ of the standard sequence.

comparison sequence occurs at an “expected” temporal location (based on an extension of the periodicity of the standard sequence) than when it occurs at a “unexpected” temporal location. Thus, Experiment 2 evaluated listeners’ abilities to detect tempo differences for onset conditions in which the onset of the comparison sequence was “early”, “late”, or at the “expected” temporal location defined by twice the IOI

of the standard sequence, as illustrated in Figure 5.2. Since this experiment did not separate “faster” and “slower” trials, differentiating thresholds for unexpected and expected conditions did not introduce a response bias to the which-is-faster paradigm, as was the case in Experiment 1. Thus, the simpler which- -is-faster paradigm used by Drake and Botte (1993) was used in Experiment 2.

5.2 Experiment 1: Direction of Tempo Change

5.2.1 Rationale

As stated above, the primary purpose of Experiment 1 was to compare listeners’ abilities to detect increase and decreases in tempo for the same IOI conditions. Predictions of the Entrainment Model regarding differential sensitivity to increases and decreases in tempo can be subdivided into four parts. First, for short IOIs, the relative JND for an increase in tempo should be lower than the relative JND for a decrease in tempo. Second, for an intermediate range of IOIs, the relative JNDs for increases and decreases in tempo should be approximately the same. Third, for even longer IOIs, the relative JND for a decrease in tempo should be lower than the relative JND for an increase in tempo. And finally, as the number of intervals in the standard sequence is increased, the observed differential sensitivity between detecting an increase in tempo and a decrease in tempo should diminish, especially for the short IOI conditions. All four parts of this predictions are either beyond the intended scope or not consistent with the predictions of previous theories of time perception (Creelman, 1962; Divenyi and Danner, 1977; Drake and Botte, 1993; Kristofferson, 1980; Michon, 1964; Jones and Boltz, 1989; Schulze, 1989). Experimental data supporting *any* parts of this prediction pose potential explanatory problems for these models.

A secondary purpose of Experiment 1 was to confirm predictions of the Entrainment Model, already supported by data from previous studies of tempo discrimination (Michon, 1964; Drake and Botte, 1993; Drake and Botte, 1994). In particular, to provide data which confirms the existence of three “zones” of time sensitivity, confirms that increasing the number of intervals in a sequences lowers the relative JND, and suggests that, in the limit, increasing the number of intervals in a sequence, improves the relative JND more for the short IOI conditions than for the longer IOI conditions, as discussed in the previous chapter.

A third purpose of Experiment 1 was to investigate the effect of musical training on tempo sensitivity, testing the controversial claim that musical training improves time sensitivity.

5.2.2 Method

Subjects. Nine subjects, five male and four female participated in Experiment 1. All subjects were students at Indiana University, reported normal hearing, and had

a wide range of musical training.

Stimuli. The stimuli consisted of 1- and 3-interval isochronous sequences, each presented at four different standard tempos. There was one fast tempo (the 100-ms IOI condition), two intermediate tempos (the 400- and 700-ms IOI condition) and one slow tempo (the 1000-ms IOI condition). Again, keeping with the convention established by Michon (1964), I have defined tempo in terms of the IOI, instead of the number of sequence events per unit time. Sequences were composed of 440-Hz 50-ms tones, with the number of tones and IOI specified by the interval and tempo condition. Four tempo (IOI) conditions combined with two interval conditions produced eight possible stimulus conditions. Sequences within each standard-2AFC trial were separated by an inter-pattern-interval (IPI) that was equal to twice the IOI of the standard, so that the onset of both comparison sequences occurred at an expected temporal location.

Apparatus. All possible experimental trials were generated prior to the experiment using the C-sound software package developed at the MIT media lab (Vercoe, 1986). The pre-generated trials were saved in named files on disk, for later playback to subjects at a comfortable listening level. By pre-generating the listening trials, the timing within each trial was not subject to variability inherent in real-time synthesis. Subjects made responses at a Silicon Graphics workstation and listened via headphones (Koss TD/75) in a quiet listening environment.

Procedure. On each trial, the subject first heard the standard sequence at the tested tempo followed by two comparison sequences, one of which was presented at a slightly different tempo from the standard. The subject's task was to indicate which of the two comparison sequences was different in tempo from the standard. Responses to each trial were entered by the subject on the computer keyboard. The next trial did not begin until a response was entered and the return key was pressed. For each stimulus condition, the adaptive-tracking procedure developed by Levitt (1971) was used to measure separate discrimination thresholds for tempo increases and tempo decreases. If the subject correctly detected two successive tempo increases, then the next tempo increase was diminished by 1%, whereas, an incorrect response led to an increase of 1% in the next tempo-increase trial. This algorithm converges to a tempo difference that the listener is able to detect with 70.7% reliability. The same adaptive procedure was applied to tempo-decrease trials, resulting in simultaneous interleaved tracks.

The initial tempo difference was 12%. In a block of 80 random trials, the IOI condition of the standard sequence remained fixed, and there were exactly 40 tempo increases and 40 tempos decreases to one of the two comparison sequences within each trial. Every 10 trials contained exactly 5 increases and 5 decreases. Relative JNDs were computed by averaging the last six reversals of each 40-trial track. Relative

JNDs were also obtained by averaging the last 20 trials of each track, in order to test the reliability of the reversal measure, but no significant difference between the two measurement procedures was found.

Each listener participated in 4 experimental sessions, with each session consisting of JND measurements for each of the four IOI conditions for one of the interval conditions (1 or 3). Repeat relative JND measurements were obtained for both interval conditions on different days. Each JND measurement took between 10 and 20 minutes, with a short rest break at the half-way mark of each block and a somewhat longer rest break between blocks. The sequence of interval conditions across experimental sessions and the order of IOI conditions within each session were counterbalanced between listeners.

5.2.3 Results

A five-factor analysis of variance (ANOVA) was run on the JNDs obtained in the experiment, with factors of musical training (three levels), number of sequence intervals, experimental session (first or second measurement of relative JNDs for either the 1- or 3-interval condition), IOI condition, and direction of tempo change (increase or decrease). Relative JNDs for all listeners and all sessions were in the analysis. Any effect due to practice should show up in the ANOVA as a significant effect of session (first versus second). The dependent variable was relative JND, which can be interpreted as a Weber fraction.

These data are first discussed with respect to our secondary purpose. Figure 5.3 shows the mean relative JNDs (averaged across all subjects) obtained for the 4 IOI conditions for 1- and 3-interval sequences. In order to compare these tempo data with the Drake and Botte (1993) tempo data (see Figure 2.7), the relative JND for a tempo increase and a tempo decrease were averaged for each IOI condition. Consistent with the distinctions made between three zones of time sensitivity, the ANOVA demonstrated a main effect of tempo [$F(3, 18) = 24.76, p < 0.001$]. (1) For all listeners, relative JNDs were lowest for the two intermediate IOI conditions (400 and 700 ms). The mean relative JNDs for the 400- and 700-ms IOI conditions, combining data from the 1- and 3-interval sequences, were 4.9% and 5.3% respectively, in fairly close agreement with Weber's law. (2) For the short IOI condition (100 ms) relative JND increased sharply to 10.3%. (3) For the long IOI condition (1000 ms), relative JND increased to 6.5%, suggesting a much more gradually decrease in the relative JND for longer IOIs than for shorter IOIs. Thus, the overall shape of the relative JND curve was U-shaped as a function of IOI, with an optimal zone of tempo sensitivity for IOI conditions between 400 and 700 ms. This is in agreement with predictions of the Entrainment Model, and consistent with the 300- to 900-ms optimal zone of tempo sensitivity reported by Drake and Botte (1993).

The ANOVA demonstrated a main effect of the number of intervals [$F(1, 6) = 117.8, p < 0.001$], in agreement with the predictions of the Entrainment Model

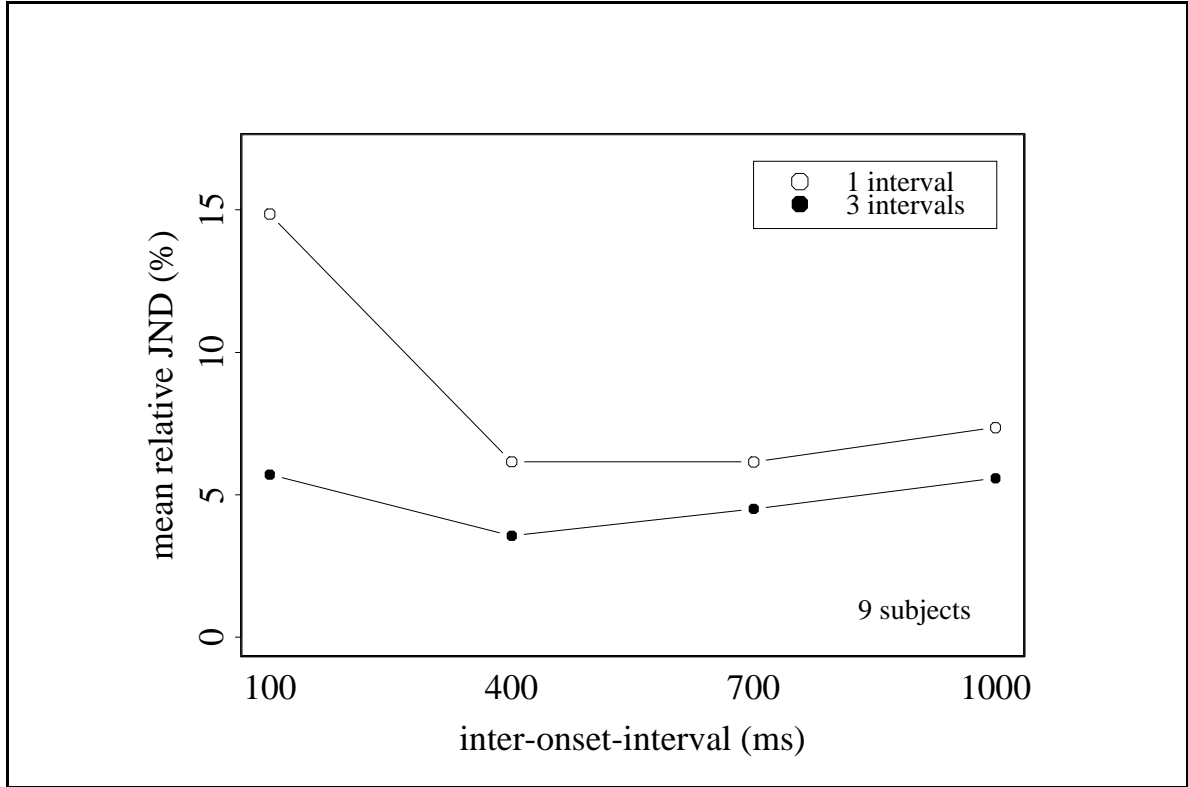


Figure 5.3: Experiment 1: Mean relative JNDs for the 100-, 400-, 700-, and 1000-ms IOI conditions, for 1- and 3-interval sequences. In this figure, the relative JNDs, determined separately for increases and decrease in tempo, are combined as an average to compare with Drake and Botte (1993; 1994) who did not distinguish between increases and decreases in tempo.

and with previous experimental results (e.g., Drake and Botte, 1993; Hirsh et al., 1990; Schulze, 1989). For the 1-interval sequences, the mean relative JND was 8.6%, whereas for the 3-interval sequences, the mean relative JND was 4.8%. For each IOI condition examined separately, the mean relative JND was lower for the 3-interval sequences than for the 1-interval sequences. In general, these data indicate poorer tempo sensitivity overall, than found by Drake and Botte (1993). One possible reason for this difference in overall performance levels is that Drake and Botte used listeners that were highly experienced with this type of tempo task. In addition, Drake and Botte (1993) did not include data from the first (practice) session in the analysis, whereas data from all sessions were included in the present analysis.

As illustrated in Figure 5.3, the ANOVA also demonstrated a significant interaction between the number of intervals in the sequence and the IOI condition [$F(3, 18) = 23.6, p < 0.001$], in agreement with the predictions of the Entrainment Model and with previous experimental results (e.g., Drake and Botte, 1993; Hirsh et al., 1990; Schulze, 1989). When the number of intervals in the sequences was

increased from 1 to 3, relative JND decreased most for the 100-ms IOI condition (9.2%) and least for the 700-ms and 1000-ms IOI conditions (only about 1.7%). The decrease in the relative JND for the 400-ms IOI condition was 2.7%, intermediate to the improvement observed for the shorter and longer IOI conditions.

With respect to the third purpose of Experiment 1, a wide range of listener tempo sensitivity was observed. Figure 5.4 shows the effect of musical training on overall tempo sensitivity in this experiment. Listeners were grouped into three categories based on their musical backgrounds. Three of the listeners were classified as non-musicians because they had little or no musical training. Three of the listeners were classified as amateur musicians. The amateur musicians all played a musical instrument and had less than 10 years of formal musical training. Three of the listeners were classified as professional musicians. Two of the three had extensive formal musical training, both having obtained degrees from Indiana University in music performance. The third listener classified as professional had played the drums for more than 10 years and performed professionally as part of a rock band.

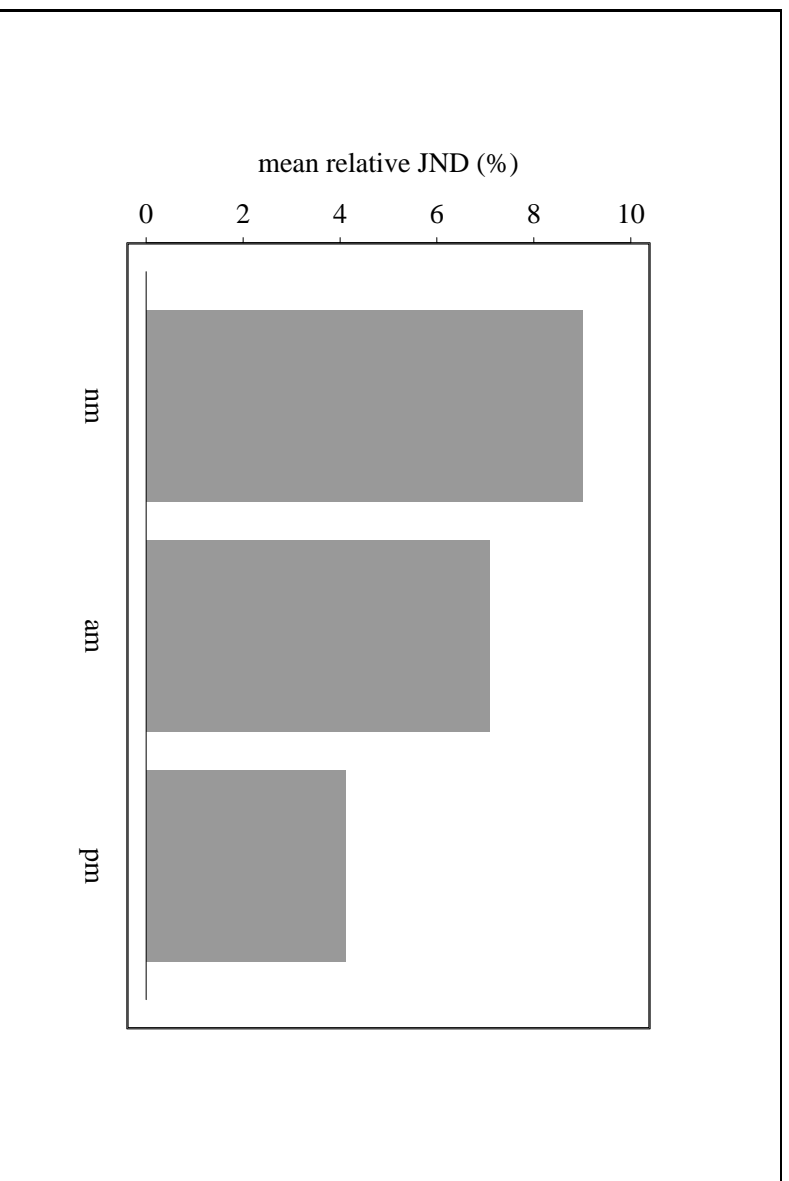


Figure 5.4: Experiment 1: The ANOVA demonstrated a main effect of musical training. Relative JNDs were lower for musicians (pm) than for amateur musicians (am) which were lower than those of non-musicians (nm).

With this classification, the ANOVA demonstrated a significant main effect of musical training. The mean relative JNDs for the professional musicians, amateur musicians, and non-musicians, averaged across experimental session, tempo, and number of intervals, was 4.1%, 7.1%, and 9.0%, respectively. For the zone of optimal sensitivity (the 400-ms and 700-ms IOI conditions) the standard deviation of the thresholds for all listeners was approximately 3.3% with the best listener (a professional musician), reliably detecting a 2.0% change for both intermediate IOI conditions (i.e., detecting a 8 ms change for the 400-ms IOI condition and a 14 ms change for the 700-ms IOI condition). Thus, musical training appears to significantly improve tempo sensitivity. However, this conclusion is premature, given that only 9 listeners participated in the experiment. It is difficult to factor out the effects that attention and effort might have on performance. Musically trained listeners may feel pressure to do well on a tempo-discrimination task and thus may try harder to do well in the experiment than do musically untrained listeners. In addition, given several more sessions of practice, non-musicians may reach threshold levels close to those found for musicians.

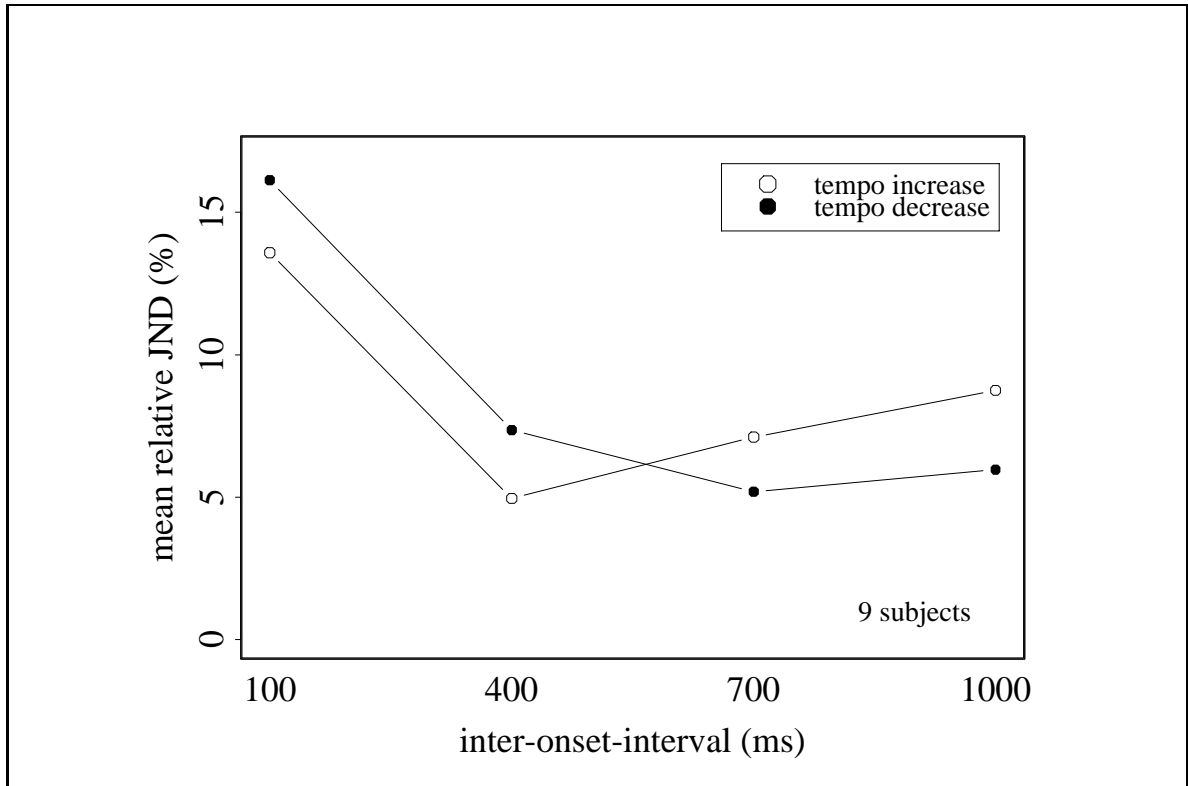


Figure 5.5: Experiment 1: Mean relative JNDs for tempo increases and tempo decreases for 1-interval sequences for all IOI conditions.

Returning to the primary purpose of this experiment, the data were evaluated with respect to predictions of the Entrainment Model concerning differential sensitivity to increases and decreases in tempo. Figure 5.5 shows the relative JNDs obtained for

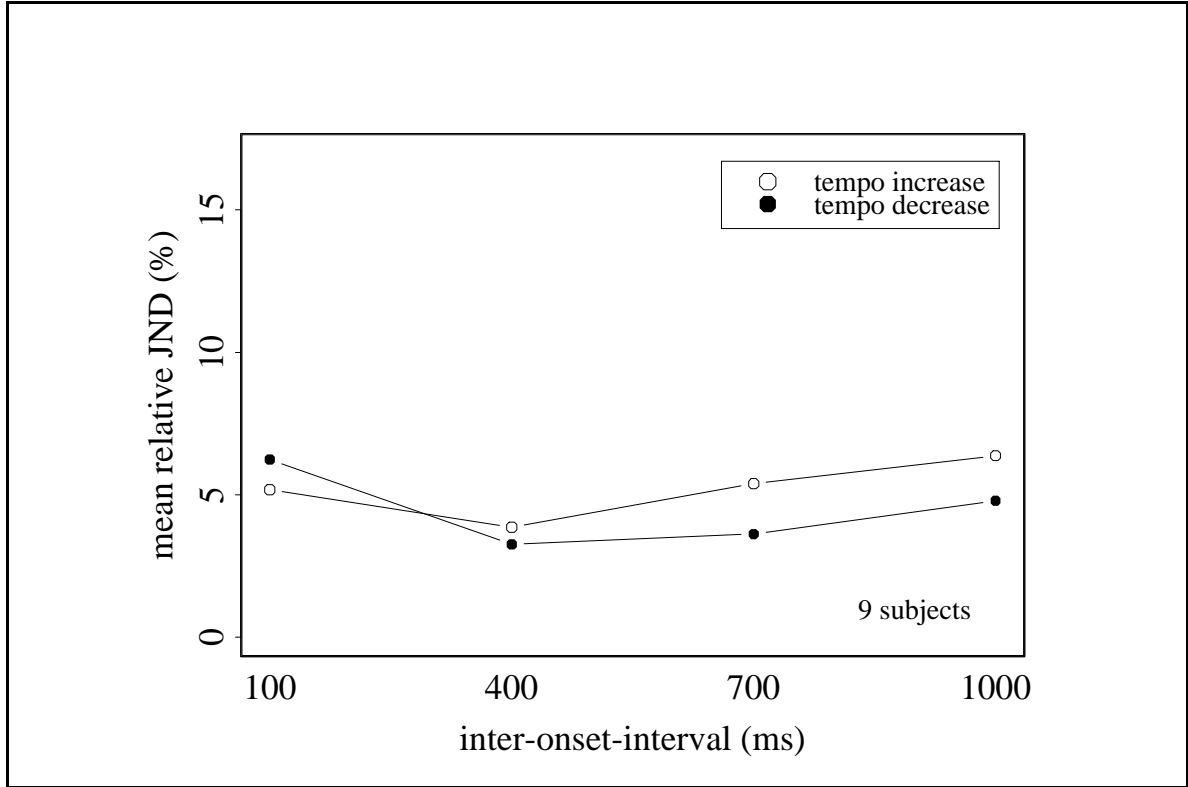


Figure 5.6: Experiment 1: Mean relative JNDs for tempo increases and tempo decreases for 3-interval sequences for all IOI conditions.

increases and decreases in tempo for the 1-interval sequences. The relative JNDs obtained for increases and decreases in tempo for the 3-interval sequences are shown in Figure 5.6. The five-way ANOVA demonstrated a significant interaction between tempo and the direction of the tempo change (increase versus decrease) [$F(3, 18), p < 0.01$], as predicted by the Entrainment Model. For the 1-interval sequences, mean relative JNDs were lower for the tempo increases than for the tempo decreases, for the 100-ms and 400-ms IOI conditions. For the longer IOI conditions (700 ms and 1000 ms), mean relative JNDs were instead lower for the tempo decreases than for the tempo increases. As a function of the IOI condition, the relative JND curves for tempo increases and tempo decreases crossed for an IOI condition somewhere between 400 and 700 ms.

The ANOVA also demonstrated a 3-way interaction between the IOI condition, the direction of the tempo change, and the number of intervals in the sequence. For the 3-interval sequences, the differences between relative JNDs for tempo increases and tempo decreases were diminished for the short IOI conditions. For the 100- and 400-ms IOI conditions, the *difference* between the relative JNDs for increases and decreases in tempo was less than 1.0%, reduced from approximately 2.5% for the 1-interval sequences. On the other hand, for the 700- and 1000-ms IOI conditions,

the approximate 2.0% difference between the relative JNDs for tempo increases and tempo decrease was maintained.

Thus, the data from Experiment 1 support *all* four parts of model's predictions regarding differential sensitivity to increases and decreases in tempo (see Panel (D) of Figure 4.11 for comparison with the simulation data): (1) for short IOIs, the relative JND for a tempo increase was lower than the relative JND for a tempo decrease; (2) for an intermediate range of IOIs, the data suggested that relative JNDs for tempo increases and tempo decreases are approximately the same; (3) for the longest IOIs, the relative JND for a tempo decrease was lower than the relative JND for a tempo increase; and (4) when the number of intervals in the sequences were increased from one to three, the observed differences in the relative JNDs for tempo increases and tempo decreases, were diminished, especially for the shorter IOI conditions. These data provide strong support for the Entrainment Model, and at the same time pose potential explanatory problems for many other models (Creelman, 1962; Divenyi and Danner, 1977; Drake and Botte, 1993; Kristofferson, 1980; Michon, 1964; Jones and Boltz, 1989; Schulze, 1989).

5.3 Experiment 2: Dynamic Attending

5.3.1 Rationale

With regard to dynamic attending, the Entrainment Model predicted that tempo differences between the standard and comparison sequence should be better resolved when the onset of the comparison sequence occurs at an “expected” temporal location (based on an extension of the periodicity of the standard sequence) than when it occurs at an “unexpected” temporal location. Thus, the primary purpose of Experiment 2 was to compare measurements of relative JNDs for onset conditions in which the comparison sequence was at “expected” and “unexpected” temporal locations.

Previous tempo discrimination studies have usually maintained a fixed inter-pattern-interval (IPI) or have assumed that varying the IPI of the comparison sequence has no effect on relative JND. No study, as far as I've been able to determine, has specifically examined the influence of the onset phase of the comparison sequence on relative JND for tempo discrimination. A secondary purpose of this experiment was to examine individual differences in sensitivity to the onset-phase of the comparison sequence, including differences due to musical training.

5.3.2 Method

Subjects. Nine new subjects, three male and six female, participated in Experiment 2. As in Experiment 1, all subjects were students at Indiana University, reported normal hearing, and had a wide range of musical training. For purposes of consistency with Experiment 1, subject selection was constrained so that the set of listeners

consisted of three non-musicians, three amateur musicians, and three professional musicians (determined according the criteria outlined in Experiment 1).

Stimuli. Only the three-interval condition and the 400-ms IOI condition were tested in this experiment. All tones in the sequence had a frequency of 440 Hertz and lasted 50 ms. This stimulus sequence was selected because the relative JND for the 400-ms IOI condition in Experiment 1 was found to be lowest for the majority of the listeners. Also, most listeners were equally sensitive to increases and decreases in the tempo of standard sequence at this rate. It was assumed that if variations in the onset-phase of the comparison sequence influence relative JNDs, then this effect was most likely to occur for the best tempo of the listeners (which was estimated to be near the 400-ms IOI condition for the listeners in Experiment 1) and that at the best rate, temporally-directed attending should not be influenced by whether the listeners were detecting an increase or a decrease in tempo.

Four out-of-phase “unexpected” conditions and one in-phase “expected” condition were selected for comparison. In the in-phase control condition, the IPI was 800 ms, equal to twice the 400-ms IOI of the standard sequence. For two “early” conditions, the IPIs were 680 and 560 ms, 15 and 30 percent shorter than the 800-ms “expected” IPI. For two “late” conditions, the IPIs were 920 and 1040 ms, 15 and 30 percent longer than the 800-ms “expected” IPI.

Apparatus. The apparatus was the same as that used in Experiment 1.

Procedure. The procedure used in Experiment 2 was identical to Drake and Botte (1993). Each subject heard the standard sequence at the tested tempo followed by a comparison sequence that was presented at a slightly faster or slower tempo. The subject’s task was to indicate which of the two sequences was faster. Which-is-faster responses to each trial were entered by the subject on the computer keyboard. The next trial did not begin until the response was entered and the return key was pressed.

Relative JNDs were determined for each of the five onset-phase conditions using the adaptive-tracking procedure of Levitt (1971) to interleave tracks for each condition. Thus, the subject had to make two correct judgments for a specific onset-phase condition before the tempo difference (between the standard and comparison sequences) for that onset-phase condition was decreased by 1%. Similarly, an incorrect response for that onset-phase condition resulting in a 1% increase in the tempo difference between the two sequences. The initial tempo difference between the two sequences for each onset-phase condition was 12%.

Each track lasted 64 trials for a total of 320 trials in the experimental session. The session lasted about an hour and the listener received short rest breaks every 40 trials. On each trial the onset-phase condition was randomly selected, but had the constraint that all 5 IPIs occurred twice every 10 trials. Consequently, each onset-phase (IPI) was equally represented by 1/5 of the trials. Listeners participated in two

identical experimental sessions.

Relative JNDs were measured by averaging the last six reversals of each 64-trial track. Relative JNDs were also obtained by averaging the last 32 trials of each track, to test the reliability of the reversal measure, but no significant difference between the two measurement procedures was found.

5.3.3 Results

A three-factor ANOVA was run on the JNDs obtained in this experiment, with factors of musical training (three levels), session (1st versus 2nd), and onset phase of the comparison sequence. The data from all listeners and both experimental sessions were included in the analysis. Any practice effects should show up in the ANOVA as a significant effect of session. As in Experiment 1, relative JND was the dependent variable.

The ANOVA demonstrated a main effect of the onset-phase of the comparison sequence [$F(4, 24) = 3.14, p < 0.05$]. For the expected condition, the relative JND was lower than the mean relative JND for the two early conditions (2.29% for the expected condition compared with 2.9% for the two early conditions), as predicted by the Entrainment Model. In contrast, mean relative JNDs for the the two late conditions were even slightly lower than those for the expected condition (1.98% for the two late conditions compared with 2.29% for the expected condition).

Possible reasons for this discrepancy between the model's predictions and the observed data will be considered in the discussion section (Section 5.4). A relevant comment made by several of the subjects is worth mentioning here. Listeners were *not* told that the onset of the comparison sequence would either occur “early”, “late”, or “in the rhythm” of the standard sequence, but instead they were essentially told to concentrate on determining which of the standard and comparison sequences was faster, independent of the gap separating them. Even so, several of the listeners commented that when the comparison sequence arrived “early” it was difficult to compare the tempo of the comparison sequence with that of the standard sequence. Thus, even though the expected condition occurred on only 1/5 of the trials, listeners internalized this temporal expectancy, deciding to use the term “early” to describe some of the comparison sequence onsets! The term “late” was also used by the listeners to describe some of the comparison sequences, but not by as many of the listeners. It may be the case that for the late onsets, listeners were able to delay their temporal expectation of the onset of the comparison sequence.

The ANOVA also demonstrated a significant interaction between the onset-phase of the comparison sequence and the experimental session (1st versus 2nd) [$F(4, 24) = 3.34, p < 0.05$], indicating that there was a significant practice effect specific to particular onset-phase conditions, as illustrated in Figure 5.7. Figure 5.7 compares the mean relative JNDs (averaged across all listeners) obtained from the first session (Panel A) with those obtained from the second session (Panel B). The mean relative

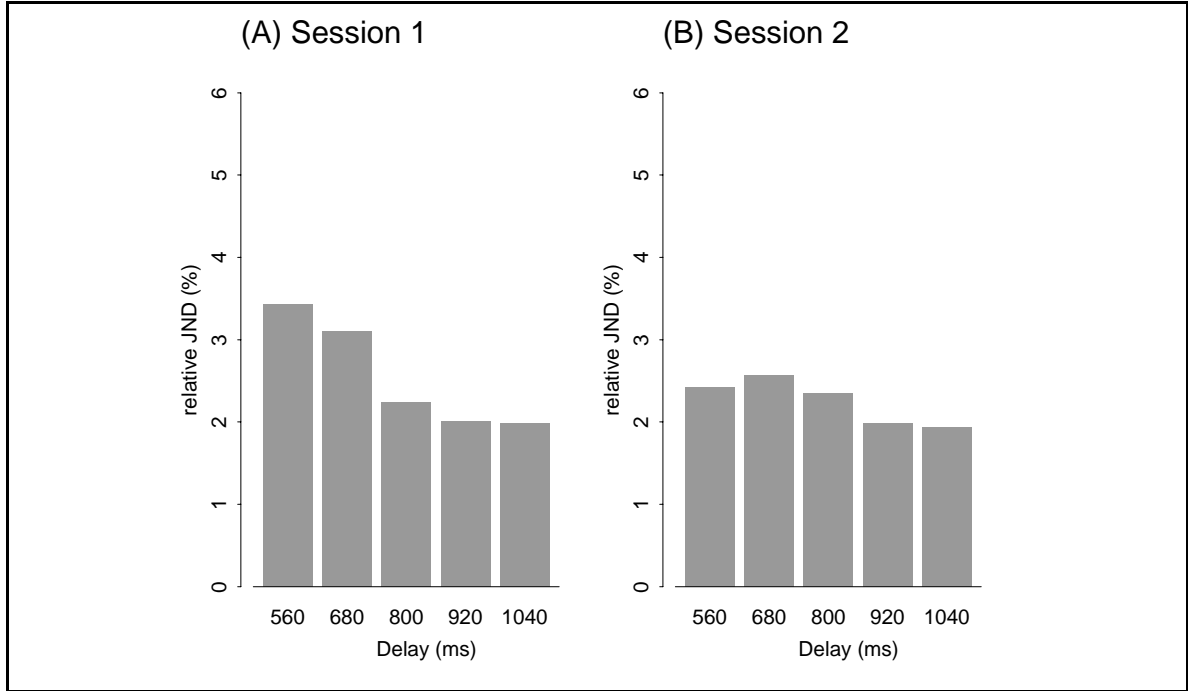


Figure 5.7: Experiment 2: Comparison of mean relative JNDs obtained for all onset-phase conditions in the first session (Panel A) with those obtained in the second session (Panel B).

JNDs for the two early conditions significantly decreased from Session 1 to Session 2. For the 560-ms IPI condition, the relative JND decreased from 3.43% in first session to 2.41% in the second session. For the 680-ms IPI condition, the relative JND decreased from 3.09% in the first session to 2.57% in the second session. Thus, the data from just the second session (Panel B) is weaker support for the Entrainment Model and, more generally weaker support for temporally-directed attending, than the data from just the first session.

Figure 5.8 shows the relative JNDs for each listener averaged across session. Listeners were grouped according to musical background: non-musicians (top-row), amateur musicians (middle row), professional musicians (bottom row). A wide range of individual differences were observed, but no significant main effect of musical training was found [$F(2, 6) = 1.4, p > 0.3$] (i.e., differences in tempo sensitivity in Experiment 2 did not depend on the amount of musical training the listener had). However, the two listeners with the highest overall relative JNDs (worst performance) were Subject's 1 and 2—both non-musicians.

The same sequence of relative JNDs across phase conditions was not observed for all subjects. However, subjects 1, 2, 6, 7, and 9 were fairly prototypical: relative JNDs for the early conditions were higher than those for the control condition (in agreement with the Entrainment Model), but relative JNDs for the late conditions

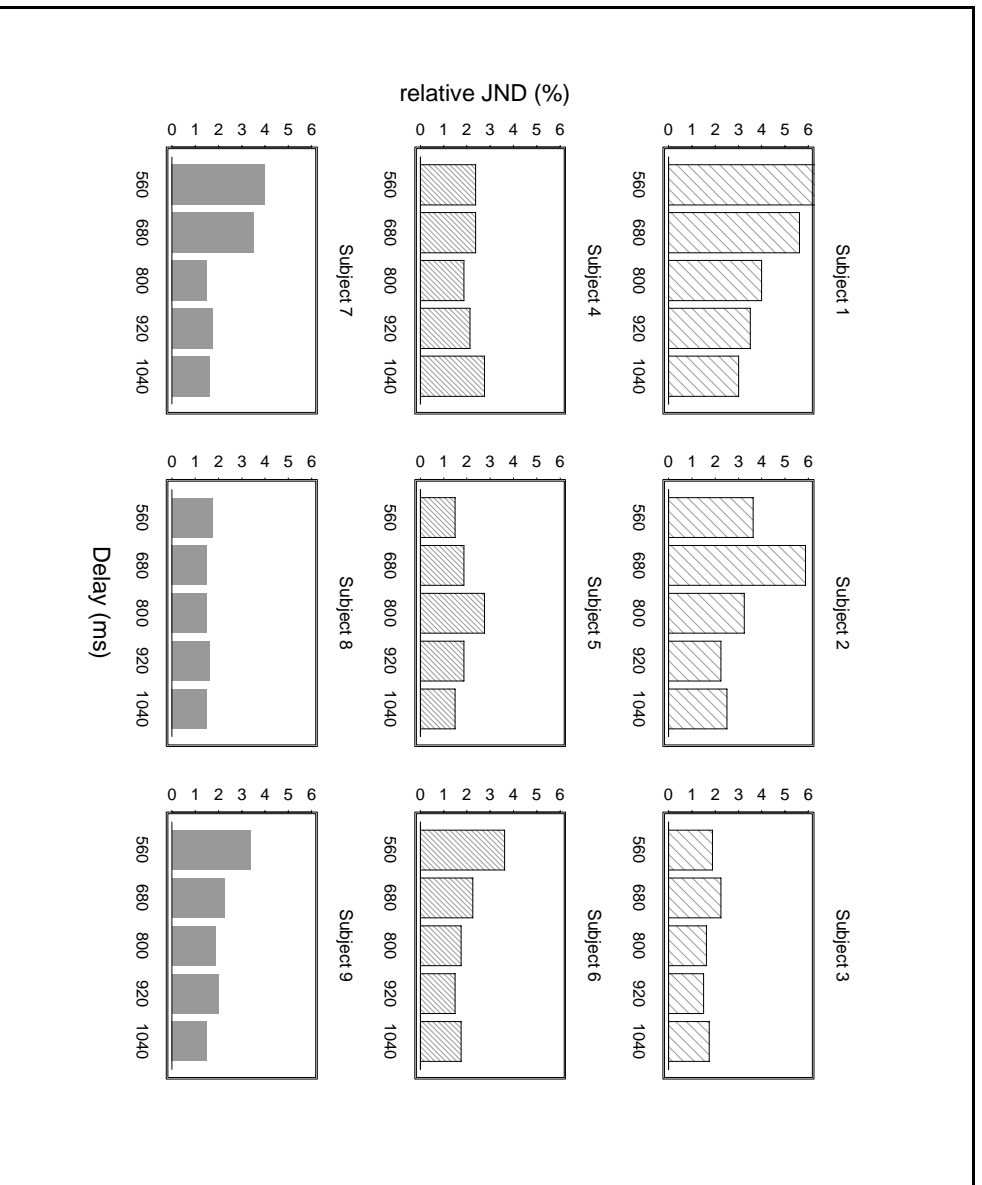


Figure 5.8: Experiment 2: Mean relative JNDs for each listener averaged for sessions 1 and 2.

were approximately equal to those for the control condition (not in agreement with the predictions of the Entrainment Model). Subject 3, 4, and 8 performed similarly in all phase conditions, suggesting that variations in the inter-pattern-interval had no effect on the tempo sensitivity of these listeners. The data from Subject 5 was completely opposite from the predictions of the Entrainment Model (relative JND for the expected condition was higher than the relative JNDs for all of the unexpected conditions), although the differences between the relative JNDs for all onset-phase conditions was still only slight. The implications of the observed individual differences will be evaluated in the discussion section that follows which includes a discussion and summary of the main results from both experiments.

5.4 Discussion

Returning to Experiment 1, recall that the primary purpose of Experiment 1 was to investigate differential sensitivity to increases and decreases in tempo. In accordance with this purpose relative JNDs were measured separately for increase and decreases in tempo, for IOI conditions of 100, 400, 700, and 1000 ms for sequences with one and three intervals for a set of nine listeners. There were four main results from these data with respect to the primary purpose of the experiment. First, for the shorter IOI conditions (100-ms and 400-ms) the relative JND for a tempo increase was lower than the relative JND for a tempo decrease. Second, for an intermediate range of IOI conditions (somewhere between the tested 400-ms and 700-ms conditions), the data suggested that relative JNDs for tempo increases and tempo decreases are approximately the same. Third, for longer IOI conditions (700-ms and 1000-ms) the relative JND for a tempo decrease was lower than the relative JND for a tempo increase. Finally, increasing the number of intervals in the sequences from one to four diminished the difference between the relative JND for a tempo increase and the relative JND for a tempo decrease, but mainly for the shorter IOI conditions (100-ms and 400-ms). For the four-interval sequences, differential sensitivity to increases and decreases in tempo was negligible for the 100- and 400-ms IOI conditions. These four main results provide direct experimental support *all* four parts of the Entrainment Model's predictions regarding differential sensitivity, as discussed in the beginning of this chapter. Furthermore, these data pose explanatory problems for previous models of time perception which have assumed no inherent difference between detecting a time increase and detecting a time decrease (Creelman, 1962; Drake and Botte, 1993; Divenyi and Danner, 1977; Kristofferson, 1980; Jones and Boltz, 1989; Schulze, 1989).

Additional tempo-discrimination studies which replicate these data are needed to further substantiate this claim. In particular, additional experiments should include more IOI conditions and increase the number of interval conditions, as well as include a larger number of subjects. By increasing the number of interval conditions, it would be possible to determine optimal sequence durations separately for increases and decreases in tempo. Since the Multiple-Look Model of Drake and Botte (1993; 1994) assumes no difference between detecting increases and decreases in tempo, it would also predict that window duration (optimal sequence duration) should not depend on whether the listener is detecting an increase or decrease in tempo. The Entrainment Model also predicts that there should be no difference between the optimal sequence duration for detecting a tempo increase and that for detecting a tempo decrease, since reducing the just-noticeable phase difference depends *only* on entrainment to the standard sequence, and not on whether the comparison sequence is faster or slower than the standard. In contrast, data from Experiment 1 suggests that increasing the number of intervals in the sequences reduces the relative JND for a tempo decrease *more* than the relative JND for a tempo increase. For detecting tempo increases the relative JND reduced from 8.6% for 1-interval sequences to 5.2% for 3-interval

sequences, whereas for tempo decreases the relative JND reduced from 8.7% to 4.5%. Although this difference in the magnitude of the change was only marginally significant [$F(1, 6) = 5.687, p < 0.1$], it implies that the optimal relative JND for a decrease in tempo will be reached with fewer intervals in the sequence than the number of intervals in the sequence required to attain the optimal relative JND for an increase in tempo. If true, this result would necessitate revisions to the Entrainment Model.

In order to compare these data with prior tempo studies (Michon, 1964; Drake and Botte, 1993; Drake and Botte, 1994), which assumed no differential sensitivity, the relative JNDs obtained in Experiment 1 for increases and decrease in tempo were averaged. These data were found to be consistent with the three main conclusions from these prior tempo studies: (1) there exist three zones of tempo sensitivity (shorter than 300 ms, between 300 ms and 900 ms, and longer than 900 ms); (2) relative JND decreases as a function of the number of intervals in the sequence; and (3) increasing the number of intervals, lowers the relative JND more for the short IOI conditions (IOIs less than 300 ms) than for the longer IOI conditions (IOIs greater than 300 ms). This agreement between the data from Experiment 1 and the earlier tempo data suggests that the relative JNDs obtained in the earlier studies actually reflects an average of the relative JNDs for tempo increases and tempo decreases. Thus, time-discrimination data obtained by testing only positive *or* negative time changes necessitates careful scrutiny, especially if conclusions are made from comparisons with data obtained by testing both positive and negative time changes.

Returning to Experiment 2, relative JNDs were determined for two unexpected-early conditions and for two unexpected-late conditions relative to the expected 800-ms IPI, in order to test predictions of the Entrainment Model regarding dynamic attending. In partial agreement with the model's predictions, relative JNDs for the two early conditions were higher than those for the expected condition. In contrast, the data from the late conditions were inconsistent with the model's predictions (i.e., relative JNDs for the two late conditions were not higher than those for the expected condition). In addition, the poorer tempo sensitivity observed for the early conditions was substantially eliminated in the second experimental session (i.e., practice on the early conditions helped).

In terms of the Entrainment Model, this practice effect suggests that the subjects were able to develop a listening strategy which enabled them to “not pay attention to” the temporal location of the onset of the comparison sequence when comparing the tempos of the two sequences. Recall that one assumption of the Entrainment Model was that the detection of a tempo change is immediate following the first interval of the comparison sequence. Thus, having multiple-intervals in the comparison sequences permitted the listeners to ignore the first interval of the comparison (the onset of which was sometimes out-of-phase with respect to expected onset), and base their tempo judgments only on the remaining intervals in that sequence. That is, having multiple intervals in the comparison sequence enabled the listeners to phase-reset and entrain to the remainder of the comparison sequence before having to make a tempo

judgment.

One possible approach to testing this hypothesis would be to repeat Experiments 1 and 2, but include only one interval in the comparison sequences. Consequently, if the detection of a tempo change is immediate following the first interval of the comparison sequence then data from this revised version of Experiment 1 (using only one-interval comparison sequences) should agree with data reported in this thesis (which was obtained using the same number of intervals in both the standard and comparison sequences). Using only one-interval comparison sequences in Experiment 2 would make it impossible for subjects to develop the listening strategy of ignoring the first interval of the comparison sequence, and thus would clarify the issue of whether the tempo of “expected” comparison sequences is inherently easier to discriminate than the tempo of “unexpected” comparisons.

One alternative theory for the present Experiment-2 data is that relative JNDs are a decreasing function of the *absolute* duration of the inter-pattern-interval, instead of a sinusoidal function of the onset phase of the comparison sequence. Perhaps, the reason that the relative JNDs for the “early”, “expected”, and “late” conditions were approximately ordered from high to low has nothing to do with early, expected, or late conditions, but is instead due to the amount of time required to *process* the tempo of the standard sequence. Thus, for the early conditions, the listeners didn’t have enough time to process the standard sequence before the onset of the comparison sequence. This processing-time explanation of the data seems unlikely since even in the worst case the inter-pattern-interval was 560 ms which is substantially longer than the 200 ms duration reported as the time required to identify components of an auditory sequence (Warren, 1993). Furthermore, listeners verbal reports of the difficulty of the task lend credence to the entrainment hypothesis and not to the processing-time hypothesis, since several of the listeners, without knowledge of the different phase conditions for the onset of the comparison sequence, decided to use the terms “early” and “late” to describe experimental trials corresponding to “early” and “late” conditions.

In order to distinguish between these two hypotheses (entrainment versus processing-time), Experiment 2 needs to be replicated using one interval comparison sequences (as suggested above) and a larger set of onset-phase conditions which spans more than one cycle and includes multiple “expected” conditions such as 2xIOI, 3xIOI, etc. Data from this replication of Experiment 2 should either provide stronger support for the sinusoidal pattern of relative JNDs predicted by the Entrainment Model or provide evidence that the processing-time argument is a better explanation of the data.

An additional unresolved issue from Experiment 2 is the substantial difference in the pattern of relative JNDs observed between listeners across phase conditions. A replication of Experiment 2 should also increase the number of investigated IOI conditions. Additional predictions of the Entrainment Model suggest that varying the IOI of the standard sequence should produce phase-advances or phase-delays in the observed sinusoidal pattern of relative JNDs, depending on the amount of over-

or and underestimation of the standard sequence's IOI. Possible then, the substantial individual differences with respect to the phase conditions reflect different “best” tempos of the listeners. The model also suggested that the tempo-discrimination performance of the listeners should degrade gracefully for tempos near their “best” tempo, while degrade less gracefully for tempos faster or slower than this best rate; i.e., memory of the standard's tempo depends on the tempo. An experiment combining a larger number of IOI conditions with a set of phase conditions spanning fairly wide range of IPI durations would provide data to test these additional predictions.

Chapter 6

Conclusions

The approach taken in this thesis has been that the development of a successful computational model of human rhythm perception must first address the perception of time intervals which comprise rhythmic patterns. Thus, modeling time perception was considered a necessary step in the development of a comprehensive model of rhythmic pattern processing. Towards this goal, an Entrainment Model of time perception was developed in Chapters 4 and 5 using the adaptive oscillator proposed in Chapter 3 as the entrainment mechanism. The model was evaluated by comparing its performance on simulated tempo-discrimination experiments to the performance of human listeners in analogous experiments, and by conducting two original listening experiments to test predictions of the model for which no human performance data was available for comparison.

The contributions of this research are twofold. First, the Entrainment Model contributes to an improved understanding of human time perception. Second, the adaptive-oscillator mechanism contributes to the development of a computational model of rhythm perception that addresses the temporal constraint on rhythmic pattern processing, the problem of timing variability, and the perception of time.

6.1 Contributions of the Entrainment Model

The Entrainment Model makes three main contributions that improve our understanding of human time perception. First, the Entrainment Model provides a single-mechanism explanation for a myriad of time-perception data that researchers had previously interpreted as strong evidence in favor of the hypothesis that short intervals (those shorter than 300 ms) are processed differently than long intervals (those longer than 300 ms) (Hirsh et al., 1990; Michon, 1964; Schulze, 1989; ten Hoopen et al., 1994). Drake and Botte (1993) further supported this claim by showing that increasing the number of intervals in an isochronous sequence improves tempo-discrimination thresholds much more for short-IOI sequences than for long-IOI sequences. Using this line of argumentation, the tempo-discrimination data reported in Chapter 5 could

also be used to support the differential processing claim. In addition, Drake and Botte (1994), in evaluating the extent to which increasing the number of intervals in a sequence improves tempo sensitivity, found evidence suggestive of two temporal windows, one of about 1.0 seconds for the processing of short intervals and one of about 2.5 seconds for the processing of long intervals. However, as I demonstrated in Chapter 4, it is not necessary to posit distinct processing of short and long intervals to account for these data. Instead, these data can be explained by the Entrainment Model through the dynamic interaction of the entrainment and period-decay processes.

The second contribution of the Entrainment Model concerns its predictions regarding differential sensitivity to increases and decreases in tempo. In Chapter 4, it was predicted that for short IOIs (fast sequences), listeners should be more sensitive to tempo increases than to tempo decreases, whereas for long IOIs (slow sequences) tempo sensitivity should be better for tempo decreases than for tempo increases. In addition, this differential sensitivity to increases and decreases in tempo should be reduced as the number of isochronous intervals increases, especially for the faster sequences. This predicted interaction was due to the asymmetric interaction of entrainment with period-decay for fast and slow sequences, resulting in superior entrainment for fast sequences. Thus, the reduction in the amount of overestimation for fast sequences is more than the reduction in the amount of underestimation for slow sequences. The tempo data reported in Chapter 5 from Experiment 1 were consistent with these predictions regarding differential sensitivity to increases and decreases in tempo, and thereby pose explanatory problems for those models which do not distinguish between positive and negative temporal deviations, such as multiple-look models (e.g. Drake and Botte, 1993).

One consequence of the model's predictions regarding differential sensitivity to increases and decreases in tempo is that in a 2AFC "which is faster" task, listeners should sometimes make systematic reversal-type errors in estimating a difference in tempo, depending on the degree of over- or underestimation of the pattern tempo and the magnitude of the tempo difference that they are being asked to detect. Thus, for a fast standard sequence listeners may indicate that the comparison sequence is faster than the standard when it is actually slower; and conversely, for a slow standard sequence listeners may indicate that the comparison sequence is slower than the standard when it is actually faster. The data obtained in a pilot experiment for Experiment 1, that used the "which is faster" task, demonstrated such reversal errors in tempo discrimination. Further data is needed to substantiate this claim.

The third contribution of the Entrainment Model concerns its predictions regarding the application of Jones's (1976) *entrainment hypothesis* to listeners' tempo sensitivity. In Chapter 4, it was predicted that for an isochronous standard sequence, listeners should be more sensitive to a tempo difference in the comparison sequence when its onset is at an expected temporal location (based on extending the periodicity of the standard sequence) than when its onset is at an unexpected location.

To explore this prediction, Experiment 2 investigated the effect of systematic deviations in the temporal onset of a comparison sequence on listeners' ability to detect differences in tempo between standard and comparison sequences. Consistent with the model's predictions, it was found that the tempo of the comparison sequence was less well resolved when it occurred at an "early" temporal location than when it occurred at the expected location. However, inconsistent with the model's predictions, tempos in the "late" unexpected conditions were resolved equally as well as tempos in the expected conditions. For all onset conditions, a wide range of individual differences was observed, and thus only weak generalizations from these data were possible.

In summary, the performance of the adaptive-oscillator-based Entrainment model was evaluated in three simulated tempo-discrimination experiments and compared with human performance in analogous listening experiments. In these simulations, the interaction of entrainment and decay processes, as mediated by the input impulse-response function, was found to be critical to modeling the human tempo data. Additional listening experiments, such as those proposed in Section 5.4 of Chapter 5 are intended to clarify the results from Experiment 2 to help to direct further development of the Entrainment Model of time perception, and to constrain the form of the adaptive oscillator in the development of a comprehensive model of rhythm perception.

6.2 Contributions of the Adaptive Oscillator

In addressing the problem of modeling rhythm perception, the adaptive-oscillator mechanism makes several important contributions. Foremost, adaptive-oscillator models of rhythmic pattern processing address the perception of time intervals that comprise rhythmic patterns, and thus do not assume musical notation as input. Second, adaptive oscillators process rhythmic patterns via entrainment and thus the perception of beats emerges as the pattern evolves over time. Finally, since the adaptive oscillator's period is modified by its input, the adaptive oscillator also addresses the problem of timing variability, both intrinsic random variability and intended "expressive" variability (such as gradually speeding up or slowing down).

Toward the development of an adaptive-oscillator model of human rhythm perception, the ability of a single adaptive oscillator to be entrained by rhythmic patterns of varying complexity was tested in Chapter 4 on the Povel and Essens (1985) set of rhythmic patterns, which vary on a complexity scale correlated with listeners' ability to memorize and reproduce those patterns. Two questions were posed with regard to the entrainment of the adaptive oscillator. First, can the adaptive oscillator lock onto a beat period consistent with listeners' perception of beats, in spite of timing variability in the tested patterns. Second, can the adaptive oscillator align its beats appropriately with those patterns. These questions were addressed for four timing-variability (or noise) conditions: 0% noise added to each interval, 5% added noise,

7% added noise, and 10% added noise.

All of the tested patterns do evoke a sense of periodic beats in the perceiver. If asked to “beat along” with these patterns, most listeners tap out beats approximately every 400 or 800 ms, consistent with a 2/4 musical meter. For the adaptive oscillator, it was found that in the no-noise condition, entrainment was successful for 80% of the patterns, with the adaptive oscillator locking onto an appropriate beat period of 400 ms. The alignment of oscillator beats corresponded with natural accents for the majority of these cases. For the 20% of the cases in which the adaptive oscillator was not entrained by the rhythmic pattern, its period oscillated between 400 and 600 ms. However, for the 5%-noise condition, the entrainment of the adaptive oscillator improved from 80% to 100% correct. Thus, for the 20% of the cases in which the adaptive oscillator’s period vacillated between 400 and 600 ms, temporal variability in the intervals of the rhythmic pattern helped the adaptive oscillator achieve a stable period. That is, noise *improved* the ability of the adaptive oscillator to be entrained by rhythmic patterns. Additional timing variability ($> 5\%$) gradually reduced the ability of the adaptive oscillator to be entrained by the tested patterns, with performance on the 10%-noise condition slightly below that on the no-noise condition.

Thus, contrary to the claims of Large and Kolen (1995) regarding phase-resetting models, the proposed adaptive oscillator was shown to display appropriate behavior without requiring strong assumptions concerning phenomenal accentuation. However, a number of questions remain concerning the relative merits of phase-resetting models and those which adjust phase incrementally as advocated by Large and Kolen (1995). In addition to the question of whether to phase-reset or not to phase-reset, different choices for activation function, output function, and period-coupling can significantly influence the entrainment dynamics of the resultant model. Obviously, the parameter space of *possible* models is quite large.

Continued comparison of model performance with that of human listeners will help resolve the issue of appropriate parameterization of adaptive oscillators. In Chapter 3, the average output of the adaptive oscillator was proposed as a measure of pattern difficulty that could be compared with listener-based measures of rhythmic complexity (e.g., for the temporal patterns in Povel and Essens (1985)). Although comparison with listeners’ ability to memorize and reproduce the patterns in Experiment 1 of Povel and Essens (1985) was only preliminary, it suggested that the average output of the proposed adaptive oscillator may provide a more accurate measure of pattern difficulty, with regard to human listeners’ judgments, than the rule-based metric proposed by Povel and Essens (1985). Additional comparisons with listener-based measurements of rhythmic complexity may suggest important revisions to the form of the proposed output function.

6.3 Evolving Adaptive Oscillators

One potential method for searching the parameter space of possible adaptive oscillators would be to use a genetic algorithm (Goldberg, 1989; Holland, 1975) to evolve adaptive oscillators that are tailored for specific environments of rhythmic patterns, much in the same way that genetic algorithms have been used to evolve connectionist networks to solve specific tasks, such as in learning the weights of a network that controls the movement of a small robot towards a fixed light source (Meeden, 1994).

The concept of a genetic algorithm is based on the principles of evolution, operating on a population of individuals, in which each individual represents a suggested solution to a given problem. To solve that problem based only on the randomly generated suggested solutions, the genetic algorithm evolves new generations of individuals (solutions) through a process of natural selection and reproduction. Individuals are selected for reproduction according to their “fitness,” a measure of solution “goodness” for the given problem. In the process of selection, those individuals with greater fitness are more likely to be chosen to reproduce than those with lesser fitness. Selected parents reproduce by recombining their information. Through this process, the genetic algorithm attempts to maximize the fitness of the population, and thus to obtain an improved set of candidate solutions.

In order to apply genetic algorithms to the parameterization problem for adaptive oscillators, each adaptive oscillator could be coded as a real-valued vector, analogous to the vector coding of network weights used by Meeden (1994). Fitness would be measured by running the coded adaptive oscillator on a test set of rhythmic patterns, selected from the particular environment, and then by averaging the output measure of synchrony for those patterns. As advocated by Meeden (1994), the next generation of individuals would be obtained through a technique called tournament selection, which uses a mutation of the fitter parent to generate new individuals (Meeden, 1994). By this process, the goal is to evolve adaptive oscillators which entrain optimally according to specific criteria (e.g., human performance data) to whatever environment of rhythmic patterns they are situated in. Thus, there are two time scales of learning: (1) in the short term, each oscillator adapts its period in response to rhythmic patterns of stimulation; (2) however, at the same time, but at a much slower rate (over generations of oscillators), the *mechanism* of the adaptive oscillator is evolving to meet the requirements of a specific environment. The usefulness of applying such evolutionary search techniques to the parameterization problem of adaptive oscillators is an open question, one which I plan to explore.

6.4 The Role of Timing Variability

The demonstrated beneficial effect of noise for the entrainment of the adaptive oscillator supports an intriguing hypothesis: brain structures involved in the perception of rhythm may be able to take advantage of temporal variability during entrainment. Of

course, the evolution of any such perceptual mechanism would be constrained by the temporal sensitivity requirements of the organism. That is, it is only beneficial for an organism to evolve perceptual mechanisms whose attunement to rhythmic patterns of stimulation is heightened by temporal variability that is less than the *minimum* temporal variability that must be detectable by that organism for its successful interaction with the environment. The proposed adaptive oscillator models the dynamics of one such perceptual mechanism.

The effect of noise on the entrainment of the adaptive oscillator is thus far preliminary, being demonstrated for only the Povel and Essens stimulus set. However, in terms of the dynamics of adaptive oscillation it is clear why noise can help entrainment. Without timing variability, it was shown that there are instances when a repeated temporal pattern will induce a cyclic pattern of changes to the adaptive oscillator's period, forcing the adaptive oscillator to bounce within a parameter region of the Arnold map bordering two stable periodic attractors. In the reported examples, the adaptive oscillator's period bounced between 400 and 600 ms, which, with respect to the 200-ms micro-pulse of the Povel and Essen's stimulus set, corresponded to potentially stable 1:2 and 1:3 entrainment ratios, respectively. However, with temporal variability added to each interval, the cyclic pattern of period changes was disrupted, enabling the effective winding number of the adaptive oscillator to change, such that it entered only a single stable region of entrainment. For the tested patterns, this corresponded to the more stable 1:2 entrainment region.

With regard to timing variability, an outstanding issue concerns quantifying the effect of noise on adaptive-oscillator entrainment. Future investigations of this effect will address how the adaptive oscillator's parameterization and the rhythmic structure of the input influence the effect of noise on entrainment. For example, in what conditions it would be pushed to 1:3 entrainment instead of 1:2. This additional research would complement a previous analysis of the beneficial effect of input noise on a recurrent neural-network model of short-term active memory (McAuley and Stampfi, 1994).

6.5 The Perception of Meter

This thesis has focused on the entrainment of a single adaptive-oscillator processing unit by rhythmic input patterns, and on the perception of time intervals within an isochronous context. However, as discussed in Chapter 1, the perception of rhythm in music and language is hierarchical, with beats perceived at different time scales (or metrical levels), resulting in a sense of strong and weak beats (or meter). Modeling the perception of musical meter with adaptive oscillators necessarily involves the interaction of many oscillators with a range of intrinsic periods, reminiscent of the BeatNet model of Scarborough, Miller, and Jones (1990). However, unlike those in the BeatNet model, adaptive oscillators have variable beat periods.

Two adaptive-oscillator models for the perception of musical meter have been

proposed (Large, 1994; Large and Kolen, 1994; McAuley, 1994a). Other than the differences, discussed in Chapter 3, between the phase-resetting oscillator proposed here and the Large and Kolen oscillator, the two models for meter perception are conceptually the same. Both models consist of a bank of adaptive oscillators with a range of intrinsic periods. Each oscillator receives input from a single input channel, but not from any of the other oscillators; that is, the oscillators do not interact with each other. By having a range of intrinsic periods, the oscillators can attain different entrainment ratios in response to rhythmic input patterns. For example, one oscillator may be 1:1 entrained with the beat of a rhythmic pattern, while the other may be 1:2 entrained by that pattern, indicative of a 2/4 musical meter. Thus, multiple oscillators can extract multiple metrical levels. These models have been successfully applied to both simple rhythms and to polyrhythms that exhibit variability in their timing (Large, 1994; Large and Kolen, 1994; McAuley, 1994a).

A main problem with these models is that unwanted harmonics are sometimes elicited by the entraining rhythmic pattern. For example, although 1:1 and 1:2 entrainment are observed for a rhythmic pattern with a stable 2/4 meter, 1:3, 2:3, or some other entrainment ratio, inconsistent with a 2/4 meter, might also be observed due to the initial spacing of the oscillator's intrinsic periods. Since the oscillators do not interact with each other, there is no way to "shut off" these spurious metrical levels (with respect to human performance). Moreover, it is not yet clear in what ways the oscillators *should* interact. Eventual interactions between oscillators should be guided by human performance data for rhythms of increasing complexity, beginning with the perception of isochronous patterns and moving towards the perception of polyrhythmic patterns.

6.6 Closing Thoughts

By assuming that an understanding of the perception of time intervals which comprise rhythmic patterns is a necessary step in the development of a computational model of rhythm perception, I have taken a modeling approach that is distinct from many of those in AI. That is, with regard to many AI approaches to rhythm perception (or more generally to many AI approaches to music perception or speech recognition), I have addressed what might commonly be considered the "non-AI" part of the problem (Brooks, 1991), since such approaches assume that durations are obtained by an unspecified pre-processor (Port et al., 1995). However, from this thesis, the intricate nature of human time perception and its consequent importance to understanding human rhythm perception should be clear. Thus, in developing a computational model of human rhythm perception it is not sufficient to abstract time to the level of musical notation.

Rather, there is a growing body of psychological data based on the entrainment hypothesis that suggests that the ramifications of human time-perception abilities extend to high-level aspects of cognition such as attention, language comprehension,

and memory. Thus, the time-perception experiments described in this thesis involving isochronous auditory sequences suggest an important behavioral probe of the entrainment underlying fundamental aspects of cognition. Furthermore, the development of a computational model of rhythm perception based on adaptive oscillators—one that is molded by the data from these time-perception experiments—is a valuable step towards a method of modeling in cognitive science that is firmly grounded in time. Furthermore, by developing a computational model that is inspired by both single-neuron models (Torras, 1985) and psychology theory (Jones, 1976), continued research is informed by both neuroscience and psychology, and thus will improve our understanding of how the macro-level phenomenon of the perception of rhythm might emerge through complex micro-level interactions in the nervous system.

Acknowledgments

Coming to Indiana University as a freshman in the fall of 1984, I was very unsure of what my “major” would be. I took classes in music, astronomy, mathematics, and French before discovering computer science. I had no idea that ten years later, I would be completing a joint Ph.D. in computer science and cognitive science. During the time I’ve spent in Bloomington, I have had many wonderful experiences, as well my share of those that I would not like to repeat. Many people have shared in both the good and the bad, and I would like to take this opportunity to thank some of these people who have helped, encouraged, and inspired me, during my studies at IU.

I’ve known Mitzi Lorentzen since my first year in Bloomington. I would like to especially thank her for encouraging me during a difficult first couple of years at IU. Her self-determination and hard work have always been an inspiration.

I am very grateful to the School of Music to have had the opportunity to play violin in several of the many orchestras over the years, the highlight of which was a performance of Bernstein’s Mass at Tanglewood, celebrating Leonard Bernstein’s 70th birthday; this was an unforgettable experience. I would especially like to thank Monique Mead for her friendship during this exciting period of time.

As a undergraduate, I also had the opportunity to participate in an IU overseas study program in Dijon, France, during which I learned at least as much French out of the classroom as in the classroom. I would like to thank Patrick Mayette and Lucinda Branaman for the many shared adventures in Europe, and more importantly for their continued friendship in the years since.

Jennifer Lill and I somehow took the same undergraduate computer science classes. Her endless energy, her sense of humor, and her obsession with finding all programming “bugs” made many all-nighters in the unrenovated Lindley Hall much more enjoyable than they otherwise would have been.

I took C311 (programming languages) with Dan Friedman my senior year. It was Dan Friedman’s inspiring style of teaching that convinced me to go to graduate school. My first graduate class (C611), also with Dan Friedman, remains the most difficult (and probably the most enjoyable) class I have taken. It is in that class that I learned exactly how many hours there are in a weekend. I am very grateful to Dan for teaching me the joy of programming in Scheme, and for his sound advice during my first year in graduate school.

For two years, I worked as an Associate Instructor for George Springer for C201

(introductory programming). I am thankful for all that I learned from him about teaching and for his continued kindness and support.

Following my teaching assignments, my soon-to-be thesis director, Bob Port, hired me as a research assistant on a ONR funded grant to study neural network models of audition. I am very grateful to him for kindling my interest in human audition, for all of the many discussions we've had on temporal processing, for his enthusiasm and support, and for his friendship. I will never forget the time Bob rushed to the emergency room after I had broken my leg, carrying a research paper that we were working on.

I took my first course in Artificial Intelligence from Mike Gasser, the co-director of my thesis. With frequent meetings, he has been a constant influence throughout my graduate studies, allowing me to work independently, but also providing important direction. It was from his initial suggestions, that I began research on modeling rhythm perception.

I would also like to thank the other members of my committee, Joe Stampfli, Jonathan Mills, and Gary Kidd. Both Joe and Jonathan have encouraged me and provided helpful comments during my thesis research. I am grateful to Gary for the many hours we spent discussing the Entrainment Model and the two experiments described in Chapter 5. In addition, I am thankful for all that I have learned from Charles Watson about psychoacoustics and science in general.

The Bloomington Symphony Orchestra has provided a much needed escape from the rigors of graduate school. I am indebted to Sarah Nelson and Juli Enzinger, whom I've known since joining the orchestra five years ago, and whom I will miss greatly now that I am leaving.

I've also appreciated my interaction with past and present members of the Artificial Intelligence Lab, including Sven Anderson, Doug Blank, Fred Cummins, Doug Eck, Susan Fox, Paul Kienzle, Lisa Meeden, John Nienart, Cathy Rogers, Raja Sooriamurthi, and Keiichi Tajima, who have provided insightful comments on papers and talks, answers to numerous technical questions, as well as moral support. From Sven and Lisa, I learned that it is possible to finish. Thanks also to Fred and Keiichi for hosting monthly poker games, enabling us to test our intuitive knowledge of probabilities.

Cathy Rogers deserves special mention, since in addition to being a lab-mate and friend, she is my wife. Consequently, she had to put up with me both at work and at home during the writing of this thesis. Cathy patiently listened to me talk about my thesis for months on end, commented on countless drafts, took on more than her share of household duties, and provided unwavering emotional support. It is difficult to imagine finishing this thesis without her.

Geoff Cashman, Keri Hensley, and Tim McCune also helped keep me sane during the thesis-writing stage, whether it was to get a quick bite to eat, catch a late movie, or go for a hike on the weekend.

Finally, I would like to thank the administrative and systems staff in the Computer

Science Department, who have been very helpful and knowledgeable. In particular, I would like to single out Pam Larson and Nancy Garrett whom have answered innumerable questions, and helped me to navigate the bureaucratic obstacles of graduate school.

This research was supported by a National Research Service Award from the National Institute of Mental Health: 46-267-01/02.

Bibliography

- Abel, S. M. (1972). Discrimination of temporal gaps. *Journal of the Acoustical Society of America*, 52(2):519–524.
- Abraham, R. H. and Shaw, C. D. (1992). *Dynamics: The Geometry of Behavior. Part 1: Periodic Behavior*. Addison-Wesley, Reading, Massachusetts.
- Allan, L. G. (1979). The perception of time. *Perception and Psychophysics*, 26(5):340–354.
- Allan, L. G. and Gibbon, J. (1994). A new temporal illusion or the TOE once again? *Perception and Psychophysics*, 55(2):227–229.
- Allan, L. G. and Kristofferson, A. B. (1974). Psychophysical theories of duration discrimination. *Perception and Psychophysics*, 16:26–34.
- Arnold, V. I. (1983). *Geometrical Methods in the Theory of Ordinary Differential Equations*. Springer-Verlag, New York.
- Baird, B., Troyer, T., and Eeckman, F. H. (1994a). Attention as selective synchronization of oscillating cortical sensory and motor associative memories. In Eeckman, F., editor, *Neural Systems Analysis and Modeling*, Norwell, Ma. Kluwer. (in press).
- Baird, B., Troyer, T., and Eeckman, F. H. (1994b). Gramatical inference by attentional control of synchronization in an oscillating Elman network. In Hanson, S. J., Cowan, J., and Giles, C. L., editors, *Advances in Neural Information Processing Systems 6*, pages 67–75. Morgan Kaufman.
- Bharucha, J. J. and Pryor, J. H. (1986). Disrupting the isochrony underlying rhythm: An asymmetry in discrimination. *Perception and Psychophysics*, 40(3):137–141.
- Bindra, D. and Waksburg, H. (1956). Methods and terminology in time estimation. *Psychological Bulletin*, 53:155–159.
- Bloch, B. (1942). Studies in colloquial Japanese IV: Phonemics. *Language*, 26:86–125.

- Bobko, D. J., Shiffman, J. G., and Shiffman, H. R. (1977). The perception of brief temporal intervals: Power functions for auditory and visual stimulus intervals. *Perception*, 6:703–709.
- Brooks, R. A. (1991). Intelligence without representaton. *Artificial Intelligence*, 47:139–159.
- Chistovitch, L. A. (1959). Discrimination of the time interval between two short acoustic pulses. *Soviet Physics: Acoustics*, 5:493–496.
- Church, R. M. and Broadbent, H. A. (1990). Alternative representations of time, number, and rate. *Cognition*, 37:55–81.
- Clarke, E. C. (1989). The perception of expressive timing in music performance. *Psychological Research*, 51:2–9.
- Creelman, C. D. (1962). Human discrimination of auditory duration. *Journal of the Acoustical Society of America*, 34(5):582–593.
- Deutsch, D. (1983). The generation of two isochronous sequences in parallel. *Perception and Psychophysics*, 34(4):331–337.
- Divenyi, P. L. and Danner, W. F. (1977). Discrimination of time intervals marked by brief acoustic pulses of various intensities and spectra. *Perception and Psychophysics*, 21(2):125–142.
- Drake, C. and Botte, M.-C. (1993). Tempo sensitivity in auditory sequences: Evidence for a multiple-look model. *Perception and Psychophysics*, 54(3):277–286.
- Drake, C. and Botte, M.-C. (1994). The measure and modelization of tempo sensitivity in auditory sequences. *Journal of the Acoustical Society of America*, 95(5):2966.
- Drake, C. and Palmer, C. (1993). Accent structure in musical performance. *Music Perception*, 10(3):343–378.
- Efron, R. (1973). An invariant characteristic of perceptual systems in the time domain. In Kornblum, S., editor, *Attention and Performance IV*. Academic Press, New York.
- Eisler, H. (1975). Subjective duration and psychophysics. *Psychological Review*, 82:429–450.
- Espinoza-Varas, B. and Watson, C. (1986). Temporal discrimination for single components of nonspeech auditory patterns. *Journal of the Acoustical Society of America*, 80(6):1685–1694.

- Fechner, G. T. (1966). *Elements of Psychophysics*. Holt, Rinehart, and Winston, New York. (Trans. H. E. Adler).
- Fraisse, P. (1948). Les erreurs constantes dans la reproduction de courts intervalles temporels. *Archives de Psychologie*, 32:161–176.
- Fraisse, P. (1956). *Les Structures Rhythmiques*. Publication Universitaires de Louvain, Louvain.
- Fraisse, P. (1963). *The Psychology of Time*. Lowe and Brydone, London.
- Fraisse, P. (1978). Time and rhythm perception. In Carterette, E. C. and Friedman, M. P., editors, *Handbook of Perception VIII: Perceptual Coding*, pages 203–254. Academic Press, New York.
- Fraisse, P. (1982). Rhythm and tempo. In Deutsch, D., editor, *The Psychology of Music*, pages 149–180. Academic Press, New York.
- Gabrielsson, A. (1973). Similarity ratings and dimension analyses of auditory rhythm patterns. I. *Scandinavian Journal of Psychology*, 14:138–160.
- Getty, D. J. (1975). Discrimination of short temporal intervals: A comparison of two models. *Perception and Psychophysics*, 18:1–8.
- Glass, L. and Mackey, M. C. (1988). *From Clocks to Chaos: The Rhythms of Life*. Princeton University Press, New Jersey.
- Goldberg, D. E. (1989). *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley.
- Goldstone, S., Boardman, W. K., and Lhamon, W. T. (1959). Intersensory comparisons of temporal judgments. *Journal of Experimental Psychology*, 57(4):243–248.
- Goldstone, S., Lhamon, W. T., and Boardman, W. K. (1957). The time sense: Anchor effects and apparant duration. *The Journal of Psychology*, 44:145–153.
- Gray, C. M., Konig, P., Engel, A. K., and 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. and Darwin, C. L. (1982). Duration discrimination in a series of rhythmic events. *Perception and Psychophysics*, 31(1):86–89.
- Handel, S. (1989). *Listening: An Introduction to the Perception of Auditory Events*. Bradford Books/MIT Press, Cambridge, Mass.
- Handel, S. (1993). The effect of tempo and tone duration on rhythm discrimination. *Perception and Psychophysics*, 54(3):370–382.

- Hirsh, I. J., Monahan, C. B., Grant, K. W., and Singh, P. G. (1990). Studies in auditory timing: 1. Simple patterns. *Perception and Psychophysics*, 47(3):215–226.
- Holland, J. (1975). *Adaptation in Natural and Artificial Systems*. University of Michigan Press, Ann Arbor, Michigan.
- Hollingworth, H. L. (1910). The central tendency of judgment. *Journal of Philosophy*, 7:461–468.
- John, E. R. (1967). *Mechanisms of Memory*. Academic Press, New York, N.Y.
- John, E. R. (1972). Switchboard versus statistical theories of learning and memory. *Science*, 177:850–864.
- Jones, M. R. (1976). Time, our lost dimension: Toward a new theory of perception, attention, and memory. *Psychological Review*, 83:323–355.
- Jones, M. R. (1981). A tutorial on some issues and methods in serial pattern research. *Perception and Psychophysics*, 30(5):492–504.
- Jones, M. R. (1987). Dynamic pattern structure in music: Recent theory and research. *Perception and Psychophysics*, 41(6):621–634.
- Jones, M. R. and Boltz, M. (1989). Dynamic attending and response to time. *Psychological Review*, 96:459–491.
- Jones, M. R., Boltz, M., and Kidd, G. (1982). Controlled attending as a function of melodic and temporal context. *Perception and Psychophysics*, 32(3):211–218.
- Jones, M. R., Kidd, G., and Wetzell, R. (1981). Evidence for rhythmic attention. *Journal of Experimental Psychology: Human Perception and Performance*, 7:1059–1073.
- Jones, M. R. and Yee, W. (1993). Attending to auditory events: the role of temporal organization. In McAdams, S. and Bigand, E., editors, *Thinking in Sound: The Cognitive Psychology of Human Audition*, pages 69–112. Oxford University Press.
- Kidd, G. (1989). Articulatory-rate context effects in phoneme identification. *Journal of Experimental Psychology: Human Perception and Performance*, 15(4):736–748.
- Kidd, G. (1993). Temporally directed attending in the detection and discrimination of auditory pattern components. *Journal of the Acoustical Society of America*, 93(4):2315.

- Kidd, G. (1994). The influence of temporal deviations on the perception of auditory pattern components. *Journal of the Acoustical Society of America*, 95(5):2966.
- Kidd, G., Boltz, M., and Jones, M. R. (1984). Some effects of rhythmic context on melody recognition. *American Journal of Psychology*, 97(2):153–173.
- Killeen, P. R. and Weiss, N. A. (1987). Optimal timing and the Weber function. *Psychological Review*, 94(4):455–468.
- Kinchla, J. (1972). Duration discrimination of acoustically defined intervals in the 1- to 8-sec range. *Perception and Psychophysics*, 12:318–320.
- Kristofferson, A. B. (1977). A real-time criterion theory of duration discrimination. *Perception and Psychophysics*, 21(2):105–117.
- Kristofferson, A. B. (1980). A quantal step function in duration discrimination. *Perception and Psychophysics*, 27(4):300–306.
- Large, E. W. (1994). Dynamic representation of musical structure. Ph.D. Thesis, The Ohio State University.
- Large, E. W. and Kolen, J. F. (1994). Resonance and the perception of musical meter. *Connection Science*, 6:177–208.
- Large, E. W., Palmer, C., and Pollack, J. B. (1995). Reduced memory representations for music. *Cognitive Science*, 19:53–96.
- Lehiste, I. (1977). Isochrony reconsidered. *Journal of Phonetics*, 5:253–263.
- Lerdahl, F. and Jackendoff, R. (1983). *A Generative Theory of Tonal Music*. MIT Press, Cambridge, MA.
- Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *Journal of the Acoustical Society of America*, 49:467–477.
- Longuet-Higgins, C. and Lee, C. (1982). The perception of musical rhythms. *Perception*, 11:115–128.
- Martin, J. G. (1972). Rhythmic (hierarchical) versus serial structure in speech and other behavior. *Psychological Review*, 79(6):487–509.
- McAuley, J. D. (1993). Learning to perceive and produce rhythmic patterns in an artificial neural network. Technical Report 371, Computer Science Department, Indiana University.

- McAuley, J. D. (1994a). Finding metrical structure in time. In Mozer, M. C., Smolensky, P., Touretzky, D. S., Elman, J. L., and Weigend, A. S., editors, *Proceedings of the 1993 Connectionist Models Summer School*, pages 219–227, Hillsdale, New Jersey. Lawrence Erlbaum Associates.
- McAuley, J. D. (1994b). Time as phase: A dynamic model of time perception. In *Proceedings of the Sixteenth Annual Meeting of the Cognitive Science Society*, pages 607–612, Hillsdale, New Jersey. Lawrence Erlbaum Associates.
- McAuley, J. D. and Stampfli, J. (1994). Analysis of the effects of noise on a model for the neural mechanism of short-term active memory. *Neural Computation*, 6(4):668–678.
- Meeden, L. (1994). Toward planning: Incremental investigations into adaptive robot control. Ph.D. Thesis, Indiana University.
- Miall, C. (1989). The storage of time intervals using oscillating neurons. *Neural Computation*, 1(3):359–371.
- Michon, J. A. (1964). Studies on subjective duration: I. Differential sensitivity in the perception of repeated temporal intervals. *Acta Psychologica*, 22:441–450.
- Miller, B. O., Scarborough, D. L., and Jones, J. A. (1988). A model of meter perception in music. In *Proceedings of the Tenth Annual Meeting of the Cognitive Science Society*, pages 717–723, Hillsdale, New Jersey. Lawrence Erlbaum Associates.
- Monahan, C. B. and Hirsh, I. J. (1990). Studies in auditory timing: 2. Rhythmic patterns. *Perception and Psychophysics*, 47(3):227–242.
- Moore, B. C. J. (1989). *An Introduction to Psychology of Hearing*. Academic Press, San Diego, third edition.
- Mowbray, G. H. (1956). Sensitivity to changes in the interruption rate of white noise. *Journal of the Acoustical Society of America*, 28:106.
- Nakajima, Y., ten Hoopen, G., and Takayuki Sasaki, G. H. (1992). Time-shrinking: A discontinuity in the perception of auditory temporal patterns. *Perception and Psychophysics*, 51(5):504–507.
- Pike, K. L. (1945). *The Intonation of American English*. University of Michigan Press, Ann Arbor, Michigan.
- Pollack, I. (1952). Auditory flutter. *American Journal of Psychology*, 65:544.

- Port, R., Cummins, F., and McAuley, J. D. (1995). Naive time, temporal patterns and human audition. In Port, R. and van Gelder, T., editors, *Mind as Motion: Explorations in the Dynamics of Cognition*, pages 339–372. MIT Press, Cambridge, Mass.
- Port, R. F., Dalby, J., and O'Dell, M. (1987). Evidence for mora timing in Japanese. *Journal of the Acoustical Society of America*, 81(5):1574–1585.
- Povel, D. J. (1981). Internal representation of simple temporal patterns. *Journal of Experimental Psychology: Human Perception and Performance*, 7(1):3–18.
- Povel, D. J. and Essens, P. (1985). Perception of temporal patterns. *Music Perception*, 2(4):411–440.
- Povel, D. J. and Okkerman, H. (1981). Accents in equitone sequences. *Perception and Psychophysics*, 30(6):565–572.
- Scarborough, D. L., Miller, B. O., and Jones, J. A. (1990). PDP models for meter perception. In *Proceedings of the Twelfth Annual Meeting of the Cognitive Science Society*, pages 892–899, Hillsdale, New Jersey. Lawrence Erlbaum Associates.
- Schulze, H. H. (1989). The perception of temporal deviations in isochronic patterns. *Perception and Psychophysics*, 45(4):291–296.
- Shroeder, M. (1991). *Fractals, Chaos, Power Laws*. Freeman, New York.
- Slobada, J. A. (1985). *The Musical Mind: The Cognitive Psychology of Music*. Clarendon Press, Oxford.
- Small, A. M. and Cambell, R. A. (1962). Temporal differential sensitivity for auditory stimuli. *American Journal of Psychology*, 75:401–410.
- Sorkin, R., Boggs, G. J., and Brady, S. L. (1982). Discrimination of temporal jitter in patterned sequences of tones. *Journal of Experimental Psychology: Human Perception and Performance*, 8(1):46–57.
- Sternberg, S., Knoll, R. L., and Zukofsky, P. (1982). Timing by skilled musicians. In Deutsch, D., editor, *The Psychology of Music*, pages 181–239. Academic Press, Orlando, Florida.
- Stevens, C. and Wiles, J. (1994). Representation of tonal music: A case study in the development of temporal relationships. In Mozer, M. C., Smolensky, P., Touretzky, D. S., Elman, J. L., and Weigend, A. S., editors, *Proceedings of the 1993 Connectionist Models Summer School*, pages 228–235, Hillsdale, New Jersey. Lawrence Erlbaum Associates.

- Stevens, S. S. (1951). Mathematics, measurement and psychophysics. In Stevens, S. S., editor, *Handbook of Experimental Psychology*, pages 1–49. Wiley, New York.
- Stevens, S. S. (1975). *Psychophysics: Introduction to its Perceptual, Neural, and Social Prospects*. Wiley and Sons, New York.
- Strogatz, S. H. and Stewart, I. (1993). Coupled oscillators and biological synchronization. *Scientific American*, pages 102–109.
- ten Hoopen, G., Boelaarts, L., Gruisen, A., Apon, I., Donders, K., Mul, N., and Akerboon, S. (1994). The detection of anisochrony in monaural and interaural sound sequences. *Perception and Psychophysics*, 56(1):110–120.
- ten Hoopen, G., Hilkhuisen, G., Vis, G., Nakajima, Y., Yamauchi, F., and Sasaki, T. (1993). A new illusion of time perception - II. *Music Perception*, 11(1):15–38.
- Thatcher, R. and John, E. R., editors (1977). *Functional Neuroscience*. Lawrence Erlbaum Associates, New York.
- Torras, C. (1985). *Temporal-Pattern Learning in Neural Models*. Springer Verlag, Berlin.
- Torras, C. (1986). Neural network model with rhythm-assimilation capacity. *IEEE Transactions on Systems, Man, and Cybernetics*, 16:680–693.
- Treffner, P. J. and Turvey, M. T. (1993). Resonance constraints on rhythmic movement. *Journal of Experimental Psychology: Human Perception and Performance*, 19(6):1221–1237.
- Treisman, M. (1963). Temporal discrimination and the indifference interval: Implications for a model of the internal clock. *Psychological Monographs: General and Applied*, 77(13).
- Turchioe, R. M. (1948). The relation of adjacent inhibitory stimuli to the central tendency effect. *The Journal of General Psychology*, 39:3–14.
- Vercoe, B. L. (1986). C-sound. Technical report, Experimental Music Studio, Media Laboratory, Massachusetts Institute of Technology.
- von der Malsberg, C. and Schneider, W. (1986). A neural cocktail-party processor. *Biological Cybernetics*, 54:29–40.
- Wang, D. L. (1995). Emergent synchrony in locally coupled neural oscillators. *IEEE Transactions on Neural Networks*. (in press).

- Warren, R. M. (1993). Perception of acoustic sequences. In McAdams, S. and Bigand, E., editors, *Thinking in Sound: The cognitive psychology of human audition*, pages 37–68. Oxford Science Publications.
- Watson, C. S. (1973). Psychophysics. In Wolman, B. B., editor, *Handbook of General Psychology*, pages 275–306. Prentice-Hall.
- Winfree, A. T. (1980). *The Geometry of Biological Time*. Springer-Verlag, New York.
- Woodrow, H. (1951). Time Perception. In Stevens, S. S., editor, *Handbook of Experimental Psychology*, pages 1224–1236. Wiley, New York.
- Yetson, M. (1976). *The stratification of musical rhythm*. Yale University Press, Yale.