Time as Phase: A Dynamic Model of Time Perception

J. Devin McAuley
Computer Science Department
Indiana University
Bloomington, IN 47405
mcauley@cs.indiana.edu

Abstract

In this paper, a dynamic model of human time perception is presented which treats time as phase, relative to the period of an oscillator that adapts its oscillation rate in response to an input rhythm. The adaptive oscillator mechanism is characterized by four fundamental properties: (1) a preferred oscillation rate which captures the notion of a preferred tempo, (2) a fast-acting synchronization procedure which models our ability to perceptually lock onto salient aspects of a rhythm, (3) a decay process to oppose synchronization, and (4) a drift process which causes the preferred rate to gradually drift towards the adapted rate, thereby modeling the context effects of long-term pattern exposure. By assuming that sensitivity to duration is a function of oscillator entrainment to the contextual rhythm, the model provides a qualitative match to data on tempo discrimination, and predicts the types of errors subjects would make on such tasks. These predictions are in agreement with data showing that subjects overestimate short intervals and underestimate long intervals.

Introduction

The perception of rhythm is a dynamic process. It embodies our sensation of movement in time, creates expectations, provides us with a sense of strong and weak "beats", and characterizes an emergent entrainment between the nervous system and the environment. In essence, rhythm plays the role of imposing temporal structure on cognition (Jones, 1976). In spite of the fact that rhythm is so central to cognition, relatively few *computational* models in cognitive science focus on rhythm. This is nowhere more evident than in work on speech recognition. To a lesser degree this is true for work in computational modeling of music cognition which is generally biased towards melody.

Central to an adequate account of rhythm perception is an explanation for how the brain measures time, as it is the temporal pattern of event onsets which primarily determines rhythmic organization (Handel, 1993). However, most computational approaches to rhythm fail even to take into account how time is perceived. Temporal intervals between acoustic-event onsets are typically represented with respect to a reference clock, for example in seconds. However, clock-time is distinct from the way we actually perceive time, as is evident from a review of the relevant time perception literature.

Fairly recently, a number of oscillator models have been proposed as a way to store temporal intervals (Miall, 1989) or more generally as a way to capture the hierarchical structure of rhythms (Torras, 1985; McAuley, 1994; Large & Kolen, 1994). In particular, the *adaptive oscillator* model proposed by McAuley (1994) implicitly measures time as phase, relative to the period of an oscillator that is able to adapt its oscillation rate in response to an input rhythm. Sensitivity to interval duration is a function of how well the oscillator is entrained to the contextual rhythm. This paper further investigates the adaptive oscillator model by showing that it captures important dynamic properties of human time perception. A brief overview of these aspects of time perception is presented in the next section.

Time Perception

How is clock-time related to perceived-time? Many psychoacoustic experiments have probed this question using series of equally spaced "isochronous" tones. For rhythm-like patterns, the interval between tone onsets (IOI) ranges roughly between 0.1 sec and 3.0 sec. In a single-interval experiment, the subject typically compares two stimuli and indicates which is longer. In a multiple-interval experiment, the subject might be asked to detect an interval change that is embedded within an otherwise isochronous sequence. In a tempo discrimination task, the subject has to decide which pattern sounds faster.

For many of the single and multiple interval experiments (Woodrow, 1951; Fraisse, 1963; Allan, 1979; Hirsh, Monahan, Grant, & Singh, 1990), the minimum just-noticeable difference (JND) is 2% – 10%, and usually occurs at a *preferred* IOI value that lies somewhere between 100 and 1000 msec. Worse sensitivity is found for intervals longer and shorter than this preferred IOI value. In tempo discrimination experiments, optimal tempo sensitivity has been found at around 100 msec by Michon (1964) and between 300 and 800 msec by Fraisse (1963) and Drake & Botte (1993). The minimum relative JND in these studies ranged from 2% – 6%. In the Drake and Botte experiments, increasing the number of isochronous intervals was found to improve thresholds, and did so in roughly uniform way across IOI's. Thresholds were raised by adding variability to the tone onsets, thus making the patterns less regular. They conclude it is the regularity of the sequences that improves discrimination thresholds.

Linked with the notion of a *preferred* IOI or tempo is what has been called the *indifference interval.* Intervals shorter than this interval tend to be overestimated and intervals longer than this interval tend to be underestimated. At the preferred tempo, listeners show no such bias, that is, they show indifference (Woodrow, 1951; Allen, 1975; Halpern & Darwin, 1982). In broad terms, listeners attempt to equalize duration estimates with respect to a central value.

The results described in the preceeding paragraphs have been the subject of much debate, as it has been impossible to obtain a consensus among researchers as to the nature of the psychophysical law for time. Preferred tempo is far from invariant. Significant differences within and between individuals exist in preferred tempo and in the shape of threshold curves. Many contextual factors, such as the length of training, order of pattern presentation, and the general mental state of the subject, play an important role in performance. To account for some of this variability, it has often been proposed that subjects show a central tendency in their judgments of time (Hollingworth, 1910; Turchioe, 1948; Woodrow, 1951). This means that a subject's preferred IOI will gravitate towards the mean IOI of the stimulus set. As a test of this hypothesis, Woodrow (1951) evaluated listeners' tendency to over- and underestimate temporal intervals by using a single 1.0 sec IOI standard stimulus in a "which is longer?" task. In support of the central tendency hypothesis, he found that the subjects' indifference interval moved from near 0.8 sec, measured after the first 60 trials to near 0.96 sec, after the second sixty trials. Since then, there has been both support and criticism of the central tendency effect.

Adaptive Oscillator Model

In recent work, McAuley (1994) described a general class of *adaptive oscillator* mechanisms which synchronize their oscillations with input rhythms. Four specific models were investigated, each of which was characterized by the shape of the oscillator's activation function. This paper focuses on a refinement to one of these mechanisms, the adaptive harmonic oscillator, and explores its viability as a model of human time perception. The research presented here emphasizes the processing of *temporal patterns*, rather than event identification. Each event is simply represented as an onset-in-time and an intensity value which can vary between $[0, 1]$ (see Figure 1A).

Unlike descriptive psychological models, the proposed *dynamic* model can be evaluated by simulating specific behavioral experiments from start to finish. This enables the model to predict how many of the contextual aspects of time perception, ranging from length of training to the order of pattern presentation should effect human performance in analogous experiments.

Four fundamental properties characterize the adaptive oscillator model. **(1)** The adaptive oscillator has a preferred oscillation period which models the listener's preferred IOI (or tempo). **(2)** A fast-acting synchronization procedure models the listener's ability to perceptually lock onto salient aspects of a rhythm. Gradient-descent on the synchronization error updates the oscillator's preferred period so that it becomes entrained to periodic components of the input. **(3)** A decay process opposes the synchronization process by forcing the adapted oscillator back to its preferred rate. As a result, it is easy for the model to adapt to tempos that are around the preferred rate, but difficult for it to adapt to tempos that are significantly slower or faster. **(4)** In order to model the context effects of long-term pattern exposure, the preferred period is allowed to drift slowly towards the current adapted period.

Preferred Period For the adaptive harmonic oscillator, the preferred period is embodied within a sinusoidal "activation" function:

$$\phi(t) = (1 + \cos(\frac{2\pi t}{\Omega(n)}))/2.$$

The period of this oscillator $\Omega(n)$ is initialized to its preferred period: $\Omega(0) = p$ (see Figure 1B). On each time-step, the current input $i(t)$ and the current value of the sinusoid $\phi(t)$ sum together to provide a measure of total activity: $a(t) = \phi(t) + i(t)$ The oscillator generates an output "spike" $o(n) = \phi(t)$ each time the total activity $a(t)$ reaches or exceeds a threshold value $\theta = 1.0$, after which the activation function is immediately phase-reset to $t = 0$ (see Figure 1C).

Synchronization and Decay An input pulse that arrives out-of-phase with respect to the zero-phase *spontaneous* firing pattern of the oscillator may, depending on the input intensity, force the oscillator to spike at a phase that is negative (early) or positive (late) in relation to the spontaneous spiking behavior. This phase information can be used to define a spike-driven gradient-descent procedure which synchronizes the oscillator with rhythmic aspects of the input pattern (see Figure 1D). Synchronization error is defined as the squared "temporal distance" between input-forced spikes and spontaneous spikes, which is the squared difference between the threshold θ and the activation $\phi(t)$:

$$E(n) = \frac{1}{2} i(t)(\theta - \phi(t))^2.$$

To minimize the synchronization error at each output spike $o(n)$ (input-forced or spontaneous), the oscillator's period $\Omega(n)$ is adapted by a fraction α that is (negatively) proportional to the partial derivative of the synchronization error $E(n)$ with respect to $\Omega(n)$:

$$\Omega(n+1) = \Omega(n) - \alpha \frac{\delta E(n)}{\delta \Omega(n)}.$$

To include a decay process as part of the update rule, one assumes that the adapted period $\Omega(n)$ is normally distributed about a mean $\overline{\Omega}$ equal to the preferred period. By taking the log of this distribution, one obtains a decay term

$$D(n) = -\frac{1}{2\sigma^2}(1 - i(t))(\Omega(n) - \overline{\Omega})^2$$

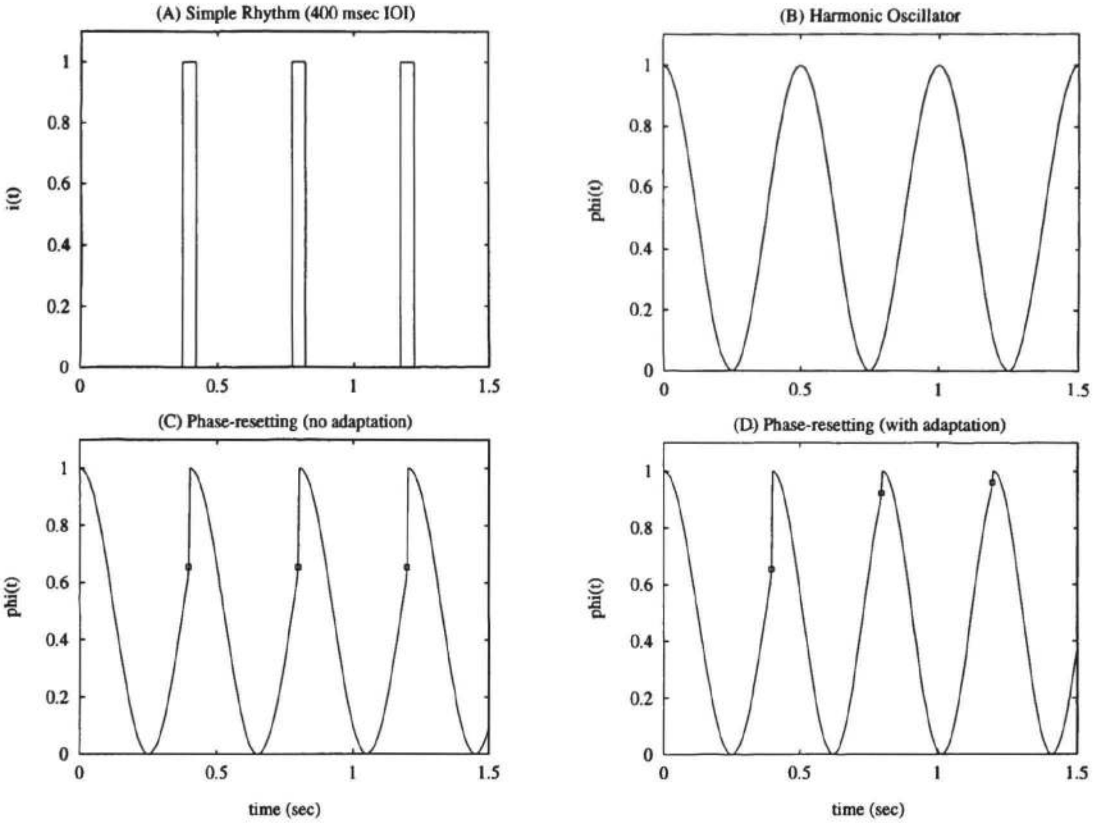

Figure 1: (A) Discrete construction of a simple input rhythm, with each onset occuring every 400. (B) Three periods of a 2.0 Hz harmonic oscillator. (C) Input pulses are added to the harmonic oscillator every 400 msec. Each input pulse causes the oscillator to spike and to be reset to zero phase. Output values, equal to the activation which it spikes, are marked by a square at each such phase-reset. (D) Fast-acting synchronization is applied to the oscillator. Notice that the output values at each phase-reset continue to increase, providing an entrainment measure. The output is approaching a value of 1.0 for which synchronization error is 0.0.

which amounts to a penalty for large differences between the adapted period and the preferred period. Gradient-descent on the decay term pushes the adapted period of the oscillator back towards its preferred period. The standard deviation σ determines how important the decay constraint is. If σ is small, the decay term will be heavily weighted. If it is large, the decay is negligible. The modified update procedure, which incorporates both decay and synchronization is

$$\Omega(n+1) = \Omega(n) - \alpha \frac{\delta E(n)}{\delta \Omega(n)} + \beta \frac{\delta D(n)}{\delta \Omega(n)}.$$

The adaptation rates for synchronization and period decay are α and β respectively.

Drift To model the context-effects of long-term pattern presentation, the preferred period $\overline{\Omega}$ is allowed to "drift" towards the adapted period $\Omega(n)$, although only at a small fraction γ of the synchronization speed:

$$\overline{\Omega}(n+1) = \overline{\Omega}(n) - \gamma(\alpha \frac{\delta E(n)}{\delta \Omega(n)} - \beta \frac{\delta D(n)}{\delta \Omega(n)}).$$

Evaluating the Model

To evaluate the performance of the adaptive oscillator model, it is assumed that human discrimination of rhythm and time is based on entrainment to the stimulus. According to this interpretation, a forced-choice tempo discrimination task requires a listener to synchronize with the standard before having to make a judgment about the comparison.

In the model, each output spike $o(t)$ provides a simple entrainment measure $[0.0, 1.0]$. An output of 0.0 indicates that the current input is 180 degrees out-of-phase with the model's oscillations, much as syncopated notes have the tendency to be felt as out-of-sync with the beat of music. On the other hand, an output of 1.0 corresponds to perfect synchronization between the oscillator's spontaneous firing and the input pulse. We can view the phase response to each input as creating an *attentional phase window* that is symmetric about the zero-phase point, as shown in Figure 2. Successive inputs that fall within this attentional focus will produce larger outputs. For an output of 0.0, we can think of attentional focus as maximally broad, that is, covering the entire range, $[-\pi, \pi]$. As an adaptive oscillator entrains to its input, the attentional focus narrows, so that when it is finally fixed only on the zero-phase point, the model is perfectly synchronized with its input.

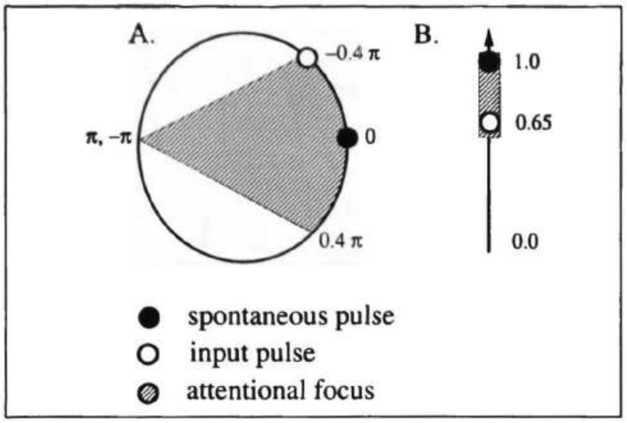

Figure 2: (A) Phase representation of time. Spontaneous pulses of the oscillator always occur at the zero-phase point, as indicated by the solid black circle. Each input pulse perturbs the oscillator at a particular phase, indicated by the open circle, thus providing an estimate of time in relation to the present period of the oscillator. In this example, an isochronous rhythm with a 400 msec IOI has forced a 2Hz oscillator to fire at phase angle of -72 degrees. Panel (B) depicts the corresponding output values which provide a measure of synchronization.

The second assumption is that a series of input pulses resulting in increasing outputs corresponds to a continuation of the same rhythm. On the other hand, a sudden drop in output–a mental "stumble" if you will–indicates that the rhythm has changed. What constitutes a sudden decrease in output? A decision criterion is set which measures the spike-to-spike change in output. A drop in output below this threshold $\lambda(n)$ signals change. In terms of phase, inputs which preserve the same rhythm fall within the attentional phase window. Inputs will fall outside a phase-threshold of this window indicate a change in rhythm. To simulate a two-alternative "which is faster?" task, the model guesses when the output continues to increase through the transition between the standard and comparison patterns. In these cases, the model can't judge which is faster, because both patterns "sound the same". When the model does detect a change in rhythm, the sign of the phase difference indicates whether the comparison pattern is faster (negative relative phase) or slower (positive relative phase).

Tempo Discrimination

Given the above assumptions, one can directly compare the performance of the adaptive oscillator model to the Drake & Botte data using an analogous 2AFC adaptive-tracking training procedure. Simulating the training procedure, as well as the specific psychophysical task is advantageous because it provides both a direct measure of performance in terms of relative JND and permits the investigation of a number of important performance issues that relate to the *training process*, such as the types of errors made, the order of pattern presentation, the range of tempos tested, and the size of of inter-stimulus interval (ISI).

Method In the Drake & Botte experiments, attempts were made to minimize context effects by counterbalancing stimulus presentation order within and between sessions. Consequently, for the purpose of modeling the human subject data, the oscillator's drift parameter γ was set to 0.0. The other parameter settings for this simulation were as follows: $\Omega = 500$msec, $\alpha = 0.1$, and $\beta = 0.1$.

The selected input rhythms were 1-5 interval isochronous patterns having IOI ranging from 300-700 msec in 50 msec intervals, Each input pulse had an intensity of 1.0. On a given trial, the model was presented a standard pattern **A** followed by a variable pattern **B**, and decided, using the measurement procedure outlined above, which was faster. The tempo of the variable pattern differed by a $\pm$ fixed percentage of the standard's. The ISI between the standard and the variable equaled two times the standard's IOI, so as to maintain the regularity of the standard's rhythm.

During a block of trials, the initial ΔIOI was set to 20%. For two consecutive correct responses, the tempo difference was decreased by 1%. For an incorrect response the tempo difference was increased by 1%. This 2AFC adaptive-tracking procedure is shown to converge to a $d' = 1.0$, corresponding to a 70.7% probability of correct detection (Levitt, 1971). Twenty 100 trial blocks were run for each type of pattern (1-5 intervals) and each tempo. Performance on each block was determined by averaging over the last 50 trials, which was analogous to looking at the last 10 reversals. For example, Figure 3A shows one of the 100 trial tracking histories. Notice that several up-down reversals occur well before the end of the first 50 trials. Averaging over the second 50 trials tended to provide a stable measurement of the relative JND. In general, every attempt was made to simulate, as closely as possible, the Drake & Botte study.

Results The relative JND measures from this simulation, averaged over the twenty repetitions of each condition, are shown in Figure 3B. An analysis of variance was run on this simulation data by the number of intervals, repetitions, and tempo, so as to address the main questions investigated in the Drake & Botte study. For comparison with the model data, the subject data is depicted in Figure 4A. In the Drake & Botte experiments, a main effect of interval number is found on threshold. The subject's mean relative JNDs are shown to be smaller for isochronous sequences than for single intervals. Furthermore, these thresholds improve as an increasing function of the number of intervals. No significant further improvement was found beyond four intervals. Similarly, the model data shows a significant effect of interval on threshold ($p < 0.001$). The model's mean relative JNDs (averaged across tempi) for single intervals is 19% and for increasing intervals is 10%, 8.5%, 6.7%, and 6.2%, respectively.

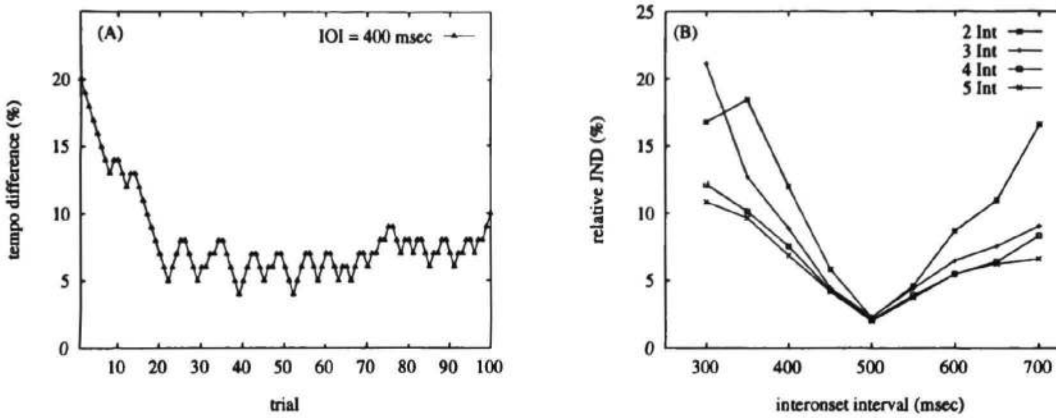

Figure 3: Simulations results of single adaptive oscillator tested on a 2AFC tempo task using adaptive tracking. (A) Tracking history for a model using a 4-tone standard pattern with a 400 msec IOI. (B) Mean relative JNDs for standard tempi ranging from 300-700 msec IOIs for 3-,4-,5-, and 6-tone sequences.

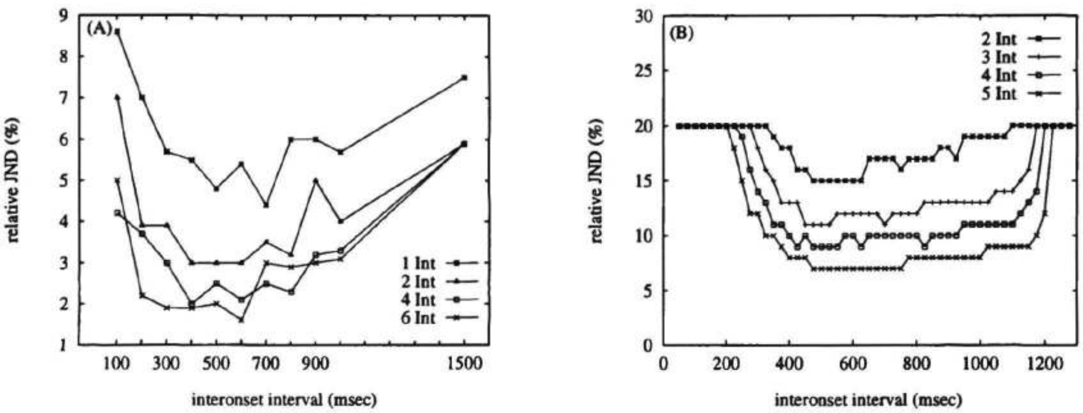

Figure 4: (A) Tempo discrimination data reproduced from Drake & Botte (1993). Mean relative JNDs are shown for sequences containing 2, 3, 5, and 7 tones across tempi ranging from 100-1500 msec IOIs. (B) Simulation results from multiple-adaptive-oscillator model tested on an analogous 2AFC tempo task using adaptive tracking.

Secondly, the Drake & Botte experiments do not confirm Weber's law, which would predict a flat relative JND curve, but instead reveal a U-shaped threshold curve with optimal sensitivity between (300-800 msec). In addition, the listeners show similar U-shaped threshold curves across the single and multiple interval conditions. No significant interaction between tempo and number of intervals is found. Likewise, the model shows a significant U-shaped effect of tempo on mean relative JND ($p < 0.001$). For the single adaptive oscillator, optimal tempo inherently corresponds to the preferred period, with mean relative JND varying from about 2% at the preferred tempo to 22% and 16% at the fastest and slowest tempi, respectively.

There are several important differences between the subject data and the model data. Unlike the Drake & Botte experiments, the model shows a significant interaction between tempo and number of intervals. With a standard pattern that is at the preferred oscillation rate, the number of intervals has no effect on performance because the oscillator is already entrained to it, that is, tempo discrimination for patterns at the preferred rate is always optimal. The second major difference is that a single adaptive oscillator is unable to capture the same range and precision with which humans can discriminate tempo. However, the adaptive oscillator model does exhibit the same *general* U-shaped threshold curves across the single and multiple interval conditions using a simple entrainment principle.

In summary, the model data, generated using a simulated adaptive tracking procedure, provides a *qualitative* match to the Drake & Botte tempo discrimination data. However, capturing the range and precision with which humans can discriminate tempo requires more than one adaptive oscillator. In preliminary simulations using multiple oscillators with preferred periods spanning the space of tempi, a better fit to the data has been obtained (see Figure 4B), although it doesn't adequately capture precision. In this simulation, each oscillator makes an independent tempo judgment which is then weighted by how well that oscillator is entrained to the standard. The weighted sum of these independent judgments provides the probabilistic response of the model.

Predicting Errors in Time Estimation

Because interval duration is measured as a phase-angle that is relative to the adaptive oscillator's period, the model's errors in estimating time are systematic. In agreement with the time perception literature, short intervals, by virtue of causing the oscillator to fire at a negative phase-angle, are overestimated, and long intervals, by virtue of causing the oscillator to fire at a positive phase-angle, are underestimated.

Armed with the knowledge that the adaptive oscillator over- and underestimates time with respect to its preferred period, it should become clear to the reader that the analogous predictions hold for tempo. In particular, in a tempo discrimination task, the mean relative JND should be different for the faster and the slower comparison patterns. In a 2AFC "which is faster?" task, standard pattern tempos that are faster than the preferred oscillator period will be overestimated, thus making it easier for the model to detect a faster comparison pattern than a slower comparison pattern. Analogously, standard tempos that are slower than the preferred tempo will be underestimated, thereby making it easier for the model to detect a slower comparison pattern than a faster comparison pattern. We can compute the skewed JND values directly. Let Ω^s equal the standard IOI and Ω^λ equal the decision threshold interpreted as an IOI. For a standard pattern that is faster than the oscillator ($\Omega^s < \Omega(n)$), the relative JND for the slower comparison pattern is

$$(2|\Omega(n) - \Omega^s| + |\Omega^s - \Omega^\lambda|)/\Omega^s,$$

whereas the relative JND for the faster comparison pattern is

$$|\Omega^s - \Omega^\lambda|/\Omega^s.$$

For a standard that is slower than the oscillator, the equations are reversed.

Conclusions

In this paper, a dynamic model of time perception was presented that provided an explanation for a number of results on single and multiple-interval discrimination, including the Drake and Botte tempo data. Time was measured as phase, relative to the period of an adaptive oscillator. Sensitivity to duration was thereby a function of oscillator entrainment. The results captured by this model included: (1) U-shaped relative JND curves, with optimal sensitivity at the preferred period, (2) improved performance as the number of isochronous intervals increased, and (3) over- and underestimation of short and long intervals relative to the preferred period, which also implied analogous time estimate errors for tempo.

In broad terms, such a dynamic model should help paint a more coherent picture of the nature of human time perception. Results which seem at odds with each other, may be unifiable if time perception is correctly interpreted as a dynamic process, thus providing a foundation for a dynamic theory of rhythm. Ongoing work is directed towards this end.

Acknowledgments

The author thanks Fred Cummins, Mike Gasser, Gary Kidd, Robert Port, and Catherine Rogers for their discussion of this manuscript. This research is supported by NIMH fellowship MH10667.

References

Allan, L. G. (1979). The perception of time. *Perception and Psychophysics, 26*(5), 340–354.

Allen, G. D. (1975). Speech rhythm: its relation to performance universals and articulatory timing. *Journal of Phonetics, 3*, 75–86.

Drake, C. & Botte, M.-C. (1993). Tempo sensitivity in auditory sequences: evidence for a multiple-look model. *Perception and Psychophysics, 54*(3), 277–286.

Fraisse, P. (1963). *The psychology of time.* Lowe and Brydone, London.

Halpern, A. R. & Darwin, C. L. (1982). Duration discrimination in a series of rhythmic events. *Perception and Psychophysics, 31*(1), 86–89.

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., & Singh, P. G. (1990). Studies in auditory timing: 1. simple patterns. *Perception and Psychophysics, 47*(3), 215–226.

Hollingworth, H. L. (1910). The central tendency of judgment. *Journal of Philosophy, 7*, 461–468.

Jones, M. R. (1976). Time, our lost dimension: toward a new theory of perception, attention, and memory. *Psychological Review, 83*, 323–355.

Large, E. W. & Kolen, J. F. (1994). Resonance and the perception of musical meter. *Connection Science.* (in press).

Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *Journal of the Acoustical Society of America, 49*, 467–477.

McAuley, J. D. (1994). Finding metrical structure in time. In Mozer, M. C., Smolensky, P., Touretzky, D. S., Elman, J. L., & Weigend, A. S. (Eds.), *Proceedings of the 1993 Connectionist Models Summer School*, pp. 219–227. Lawrence Erlbaum.

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.

Torras, C. (1985). *Temporal-Pattern Learning in Neural Models.* Springer Verlag, Berlin.

Turchioe, R. M. (1948). The relation of adjacent inhibitory stimuli to the central tendency effect. *The Journal of General Psychology, 39*, 3–14.

Woodrow, H. (1951). Time perception. In Stevens, S. S. (Ed.), *Handbook of Experimental Psychology*, pp. 1224–1236. Wiley, New York.