This is a brief summary of the author's PhD dissertation,
Timing in Temporal Tracking.
Leiden University, 1967
Published by Van Gorcum, Assen (The Netherlands)

Timing in Temporal Tracking[1]

John A. Michon

Abstract

The temporal structure of synchronization behaviour is studied in a quantitative way, and an information processing model is described which provides a common frame of reference for research on timing, as well as a conceptualization of the human 'time sense'. Several experimental results are discussed, among others the way in which subjects follow a sudden change in the rate of presentation of stimuli. The interrelations between temporal and non-temporal information is touched upon.

Introduction

In many circumstances people are required to make their actions coincide with a sequence of external events. Working at a conveyor belt the worker has to carry out his perceptual-motor routines in ways that leave little room for playful variations in timing *A fortiori* this applies to musical ensemble playing: the art of performing

---

[1] This text is based in part also on Michon, J. A., & Van der Valk, N. J. L. (1967). A dynamic model of timing behavior. *Acta Psychologica, 27*, 204-212; and Michon, J. A. (1968). A model of some temporal relations in human behavior. *Psychologische Forschung, 31*, 287-298. I am greatly indebted to Ir. Nico van der Valk, who made a truly indispensible contribution the development of the underlying model and to the mathematical analysis of the data.

together requires a special sensitivity for the temporal relations between successive notes and phrases (Rasch 1981). A major aspect of mastering a motor skill is, in fact, the acquisition of a correct timing pattern for the successive elements of the movement plan. Mastering a skill consists to a large extent of storing a temporal plan that can be carried out independently of the particular pattern of muscular innervation. It is, for instance, known from physiology that simple repetitive movements are carried out with different muscle innervation each time, presumably in order to postpone fatigue (Seashore, 1951).

Only if the temporal plan is stored independently of the execution rules can the marvels of timing displayed by musicians, dancers and sportsmen arise. The temporal plan may in part be intertwined with the spatial plan, however, and so give rise to kappa or tau effects.[2]

Here I propose a way of dealing with the temporal structure of discrete sensori-motor performance quantitatively, a way which also implies an explicit model of the 'time sense'—a rather metaphorical expression for the mechanisms which enable man to store and retrieve sequences of durations.

The synchronization task studied—key tapping—is a very simple form of the sensori-motor skill required in such complicated performance as playing Bach or Beethoven. In key tapping the temporal structure is the paramount characteristic, whereas spatial melodic and other elements have been reduced to a bare minimum. Key tapping in synchrony with a sequence of clicks has been studied quite extensively but not very intensively. I will not digress on the available literature (see Michon, 1967 for a review), but instead I refer briefly to two major shortcomings of the earlier work on timing in key tapping.

First, there appears to be no common frame of descriptive terms which would enable authors to express their experimental data and conclusions in a uniform

---

[2] By *kappa effect* we mean the influence of spatial relations between stimuli (distance) on subjective duration, discovered by Cohen, Hansel and Sylvester (1953). The *tau effect* was first described by Helson and King (1931) as the effect of inter-stimulus intervals on perceived distance.

way. Needless to say, the absence of such a frame of reference makes it very difficult to combine experimental data from different projects.

Second, the research on key tapping and synchronization has largely overlooked the sequential relation between the successive intervals, while the presence of such interactions is in fact beyond doubt.

The model to be presented here is based on an extensive experimental study, described in greater detail in an earlier study (Michon, 1967). It offers a remedy for the first problem and at the same time inherently deals with the second problem.

The Experimental Situation

Participants in my experiments found themselves in the following situation. They had to listen to a sequence of clicks at a level of 45 dB above threshold, presented through headphones, and were instructed to tap a lightweight Morse key in subjective synchrony with the sequence heard. The intervals presented were not necessarily of equal duration. Hence subjects were actually faced with the problem of predicting the duration of the next interval on the basis of the previous information available to them, thereby minimizing their errors of synchronization or, in other words, the time difference between a click and the corresponding tap. In the experiments several types of input sequences were used, among others the step function, ramp functions and sine wave functions (Figure 1).

The representation of the interval durations in Figure 1 is different from the usual representation of events on a time line. Instead, the duration of the i^{th} interval is indicated by a vertical bar at point i of the abscissa. This Time/Order representation has more than just a pictorial import: it actually transforms the data into a regular time series, thus enabling us to analyze the properties of these data by means of techniques like discrete time systems analysis which would otherwise not be applicable.

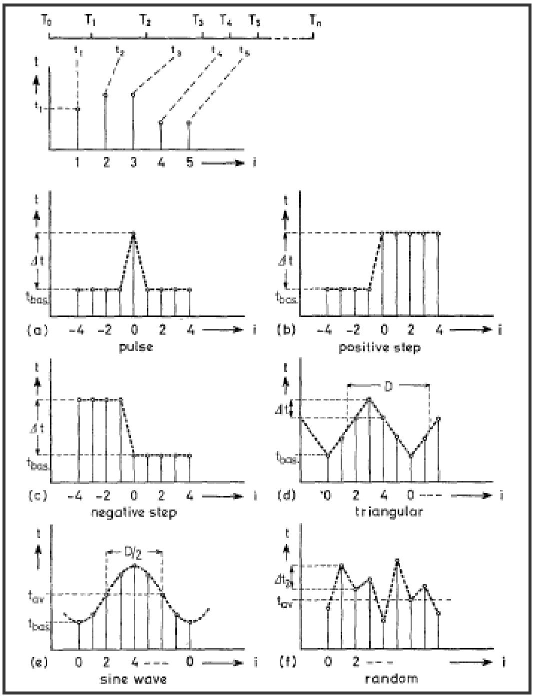

Figure 1. Examples of modulated interval sequences presented to subjects. Representation in duration/order diagram: the rank number *(i)* of the intervals is plotted along the abscissa, their duration *(t)* along the ordinate.

The Model

In order to meet the requirements of the instruction, the subject must be able to retain information about previous intervals and information about synchronization errors, that is, the time difference between tap and click. On the basis of this information he has to estimate the period between his last tap and the one to be produced, i.e. his next interval estimate.[3]

Phenomenologically these two components appear to be quite distinct entities. So long as the input sequence is not entirely random, the subject needs a strategy which results in an acceptable prediction, a correct prediction

[3] In this article I use the masculine pronouns because all subjects in the experiments under concern were of the male gender.

being a subjectively accurate temporal coincidence of a click and its corresponding tap.

On the basis of preliminary experimental results and parsimony requirements I started from the following simple assumptions:

(1) the subject is able to store correctly: the duration of the length of the previous input interval (t_{i-1}) and the size of the immediately preceding synchronization error (ε_{i-1});

(2) the ith interval produced $(\hat{t}_i)$ is simply the sum of these two components:

$$\hat{t}_i = t_{i-1} + \varepsilon_{i-1} \tag{1}$$

An example of the extent to which this model adequately describes the performance of subjects in response to sinusoidally modulated conditions, is presented in Figure 2. In these experiments the length of successive intervals varied according to a simple sine wave, or a weighted combination of two or three simple sine wave functions.

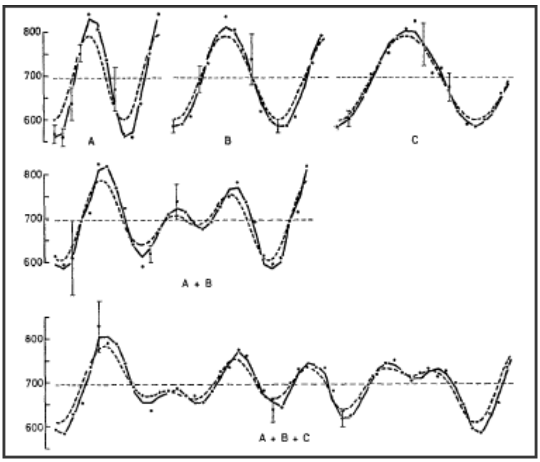

Figure 2. Responses (dots) of a subject to simple and composite sinusoidal inputs. The dashed line represents the input function, the solid polygons are the response functions as predicted by the basic model. Time in milliseconds.

A closer look at Figure 2, as illustrated in Figure 3, shows a consistent lag of the experimental data behind the predicted values. This suggests that the prevailing synchronization error is not fully compensated at the next interval. Instead compensation is spread out over a number of successive intervals. In fact the value of the i^{th} output interval can be expressed as:

$$\hat{t}_i = t_{i-1} + \sum_{k=1}^{\infty} f_k(a,b)\varepsilon_{i-k} \tag{2}$$

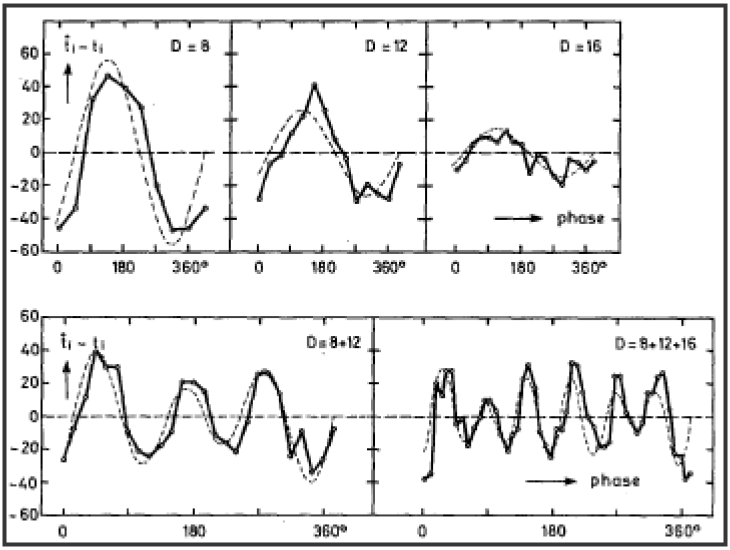

Figure 3. Averaged responses (three subjects) to simple and composite sinusoidal inputs (solid polygons). The thin dashed curves are the responses as predicted by the basic model.

This equation indicates that the internal representation of the synchronization error actually stored in memory, is a weighted sum of a number of previous values of ε. For $k \rightarrow \infty$ these weights (f_k), which contain two parameters a and b, will tend to zero. The weights are not independent, as in multiple regression models, but derive in a deterministic way from the systems analytic approach, the z-transform analysis. The Appendix at the end of this article provides further details (Michon, 1967; Michon and Van der Valk, 1967).

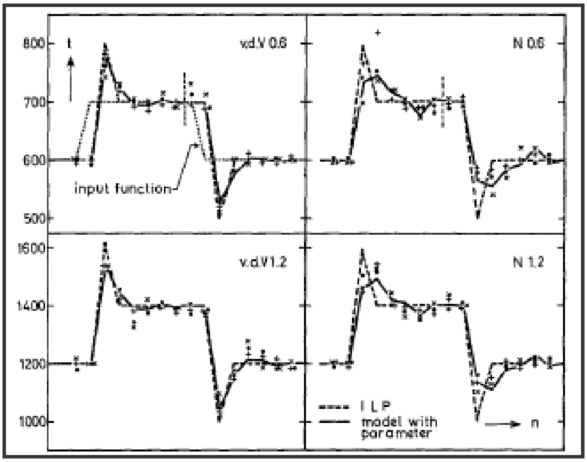

Figure 4. Average responses to step function inputs. Data of two subjects for two durations (0.6 and 1.2 sec) and three step widths (Δt = 8% (×), 16% (•) and 32% (+). The dashed line represents the prediction by the basic model, the solid line shows the model response with parameters fitted.

An example of the fit that can be obtained in this way is shown in Figure 4, which presents data of a positive and negative step function of two subjects. In this and other cases, the fit obtained by suitable estimates of the parameters a (a personal parameter) and *b* (a time dependent parameter) is very good as long as we deal with responses averaged over a number of trials. The estimated values of a and *b* for the two subjects represented in Figure 4 show quite marked and characteristic distinctions. The variance explained amounts to 75 to 95% of the non-random variance present in the results, depending on subjects and conditions.

The response curves shown in Figs. 2 and 4 are average responses. The constituent single trial response curves are affected by 'noise' which is adequately described by the exponential function

$$\delta\hat{t} = k\hat{t}_{av}^{1.5} + \mu \,, \tag{3}$$

with $k \approx 0.05$ and μ represents the intrinsic noise of the motor system (Figure 5).

This relation strongly suggests that the 'internal representation' of an interval is subject to progressive decay while it is being held in an apparently imperfect memory.

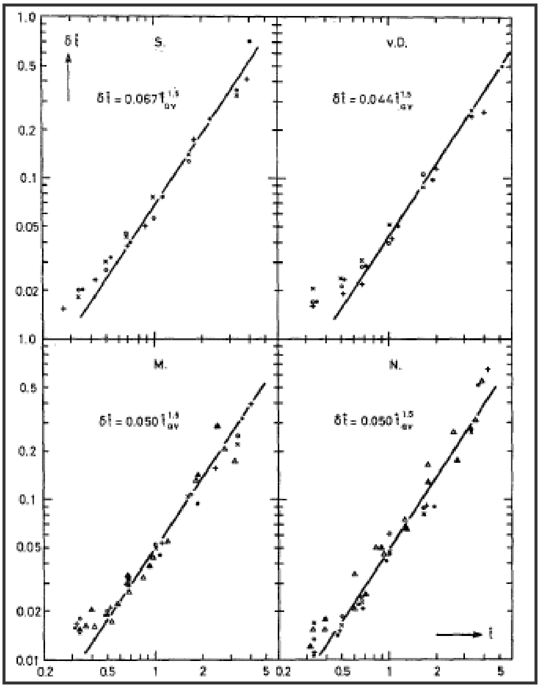

Figure 5. Variability of the response to input sequences as a function of the average length of the output intervals. Time in seconds. Each point is based on 200 trials.

An Information Processing Model

How can we establish a meaningful relation between these findings ? The most elegant way of doing this is to implement the model as a computer program, in such a way that the information processing routines are reasonable in view of what is known about the time sense. In and of themselves mathematical models are psychologically void of sense. This is to say that it is not sufficient to program the 'time sense' simply in terms of its difference equation [Eq. (2)] but that we also need to be explicit about functions like memory (Creelman, 1962) and order discrimination (Hirsh and Sherrick, 1961) which are known to play a role in time perception. Let me start therefore by trying to summarize the state of knowledge with regards to the 'time sense'. This state has been characterized more than once as chaotic. The main reason of the prevailing confusion, as I have argued earlier (Michon, 1965, 1967), is that almost all authors tried to point out the specific physiological or psychological mechanisms responsible for our sense of *protensity* (Titchener), such as respiration, a-rhythm, brain traces, etc.

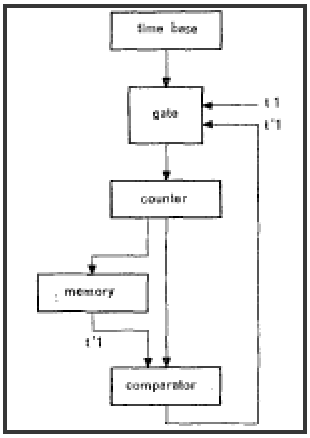

Figure 6. Formal lay-out of 'time sense' underlying most historical theories (Reproduction experiment).

I do not intend to expose the arguments in more detail at present. It appears to me however that the theory of time perception can be summarized in terms of the very simple lay-out shown in Figure 6.

This basic arrangement consists of a pulse generator which provides the time base against which a stimulus duration is evaluated. The stimulus boundaries trigger a gating mechanism, thus allowing the time base pulses to proceed to an integrating mechanism during the stimulus interval.

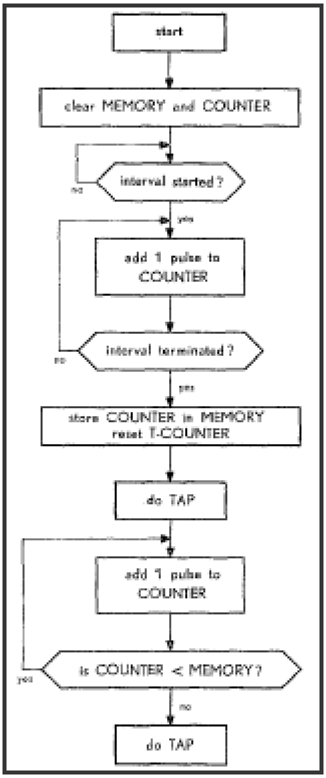

Figure 7. Logic diagram corresponding to the functional diagram of Figure 5 (Reproduction experiment).

Since time experiments involve a succession of at least two intervals we need to store the first interval while the second is in progress, so that a comparison can be made after the second interval has ended. The logic diagram, Figure 7, shows the actual steps carried out in the operation of this functional model. The crucial aspects of this situation are:

(1) the properties of the memory in which temporal information is stored, which has to be a compound memory as we concluded from our experiments;

(2) the way in which the time difference (synchronization error) is evaluated.

The final model, accounting not only for reproduction performance but also specifically for synchronization is shown in the diagram Figure 8. The program has been implemented in Fortran II on a PDP-7 computer and takes into account both of the afore mentioned aspects. It contains a separate counter for the input interval t_i and the synchronization error r_i (which can be positive or negative depending on the order of tap and click).

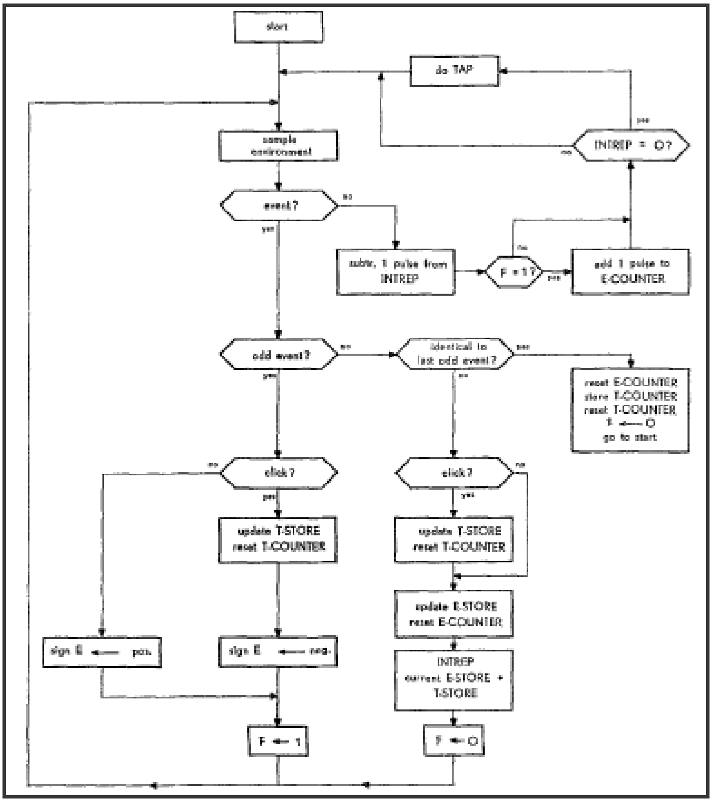

Figure 8. Logic diagram of ‘time sense’ for synchronization experiments and for generalization.

According to this model, the subject is sampling his environment in search of ‘events’, which can be either *click* or *tap* or a coincidence of both. If no event is detected during the sampling period Δt (which may be conceived of as the equivalent of one ‘time base’ pulse) the counters for t_i and—depending on the state of an ‘error flag’ the current synchronization error, are updated and a decision is taken whether or not a tap is to be produced in this sampling instant. If no tap is given, the next sample of information is taken. When an event occurs, it is determined whether it is odd or even in the series of all events since the start of the experimental run. This is achieved by means of binary ‘error-flag’ which changes its value with each event (tap or click).

The error counter is incremented only between an odd and the next even event, unless both are clicks or taps. At each even event the current content of the error counter is stored and the counter is reset.

This distinction between even and odd events is seemingly not given phenomenologically. However, in actual performance subjects are pairing events: a click and a tap are taken as belonging together, irrespective of the order in which they appear, as long as their distance is small in comparison to t_i. The distinction between odd and even which is made by the program is simply a formalization of this implicit pairing. The remainder of the diagram

Figure 8 will be self-explanatory; it simply determines if and when the various memory stores are updated and reset. In the present model the three memories: T-STORE, E-STORE and INTREP have properties that were more or less stringently dictated by the experiments.

The T-STORE simply contains an internal representation of $t_i = n\Delta t$. The E-STORE contains a weighted error score [equation (2)] and INTREP is a decaying sum of these two, in accordance with the exponential relation [equation (3)].

Although in this way the model is completely circumscribed, it should be stressed that after all it is immaterial to the mathematical basis of the model how the contributions of the T-store and the E-store are weighted in the equations that describe the performance of the system. The choice of weights is essentially a matter of psychological interpretation. The adopted solution rests, however, on several experimental results described elsewhere (Michon, 1967), and is therefore not an entirely arbitrary choice on the part of the investigator.

Extension and Conclusion

The temporal structure of behaviour reflects a problem in human information transmission. Information about durations is to be stored and retrieved at specific instants, and appears to be subject to decay like any other kind of information, symbolic or spatial. The question of the interaction between temporal and non-temporal information is still far from answered, but we may draw several conclusions from a couple of preliminary experiments. In these experiments subjects not only had to synchronize by tapping one single key, but were at the time responding differentially to visual signals randomly taken from a set of up to 8 alternatives (Michon, 1967). In one set of conditions the extra information was integrated with the temporal information. The subject was synchronizing while tapping differentially on a set of response keys. In other conditions the non-temporal information was functionally separated from the temporal information. In that case the reactions were given independent of the ongoing single key-tapping performance.

Two main conclusions were drawn from these experiments:

(1) If it is possible to integrate the two types of information, the response of the 'time sense' will be hardly affected at all. Only in the second condition will the extra load have a detrimental effect.

(2) However, it appears that in the latter case, it is not so much the systematic (average) response which is affected, but the short-term random (noise) variations about the average response. In other words: the system parameters of the timing mechanism will not change as a function of extra information load. This suggests a relative independence of the 'time sense', and a time-sharing between functionally 'unconnected' information channels.

Apart from the fact that these results may be of importance for our insight in the time constants of short term memory and sensory switching, they also open up vistas for the practical problem of the so called ‘mental load’ (or information processing load) in industrial performance. An exposition of the work in this important area was laid down in several earlier studies (Michon, 1966a, b).

Appendix

Preliminary observations have led us to conceive of the timing mechanism as a simple, causal, linear predicting system (Michon and Van der Valk, 1967). In essence it should be considered as a purely formal model, although it may be embodied in a physical or physiological system.

The series of stimulus intervals with which the subject is required to synchronize and which may be stationary or modulated in a systematic or random fashion, is considered as a series of successive states of a discrete function $x(n)$. The response sequence produced by the synchronizing subject is likewise considered as the series of states of a second function $y(n)$. We want to establish the relation between any input and its output, taking into account that the momentary response may be influenced by previous input intervals. If there is such an influence the after-effect of the response to an interval will last for one or more intervals beyond the interval to which it is tie response, and it can be described by a function $h(n)$. This function is called the ‘impulse response’ and describes the system’s behaviour in response to a single interval of unit duration, x_0.

Provided the system is linear, $h(n)$ is independent of the interval to which it is the response, except for a multiplication constant: if the input interval is x_0 then $y(n) = x_0 h(n)$; see e.g., Schwarz and Friedland (1965, p. 12). If the first interval is followed by others, the response at any interval will be the sum of the partial responses to all preceding intervals up to and including the last, such that

$$y_n = \sum_{k=0}^{n} x_{n-k} h_k \quad (4)$$

Direct calculation of this ‘convolution sum’, usually expressed as

$$y(n) = x(n) * h(n) \quad (5)$$

is very cumbersome, but can be simplified by the use of generating functions, that is, by transformations defined by

$$F_n(z) = \sum_{n=0}^{\infty} f(n) z^{-n} \quad (6)$$

(Feller, 1957, chapter 11; Schwarz & Friedland, 1965, chapter 8). The generating function is in fact a power series in some ordering variable, of which the coefficients of successive terms are the values of successive states of the original discrete function. It follows that the algebraic multiplication of the generating functions of two number sequences will yield a third generating function the terms of which have coefficients of the general form shown in equation (4). The convolution expressed in equation (5) can therefore be determined from

$$Y_n(z) = X_n(z) \cdot H_n(z) \quad (7)$$

If, like in the present situation, *y(n)* and *x(n)* are known and the transfer function is to be determined, the results can be obtained by dividing $Y_n(z)$ through $X_n(z)$ and transforming the resulting $H_n(z)$ back, to obtain *h(n)* (Schwartz & Friedland, 1965, p. 246-247).

The model we developed is based on the most simple linear predictive system that can be conceived of. The ideal linear predictor (ILP) simply predicts the value of inputs from direct extrapolation of the two foregoing input intervals (Truxal 1955, p. 534). The behaviour of this system is therefore completely described by the difference equation

$$y_n = x_{n-1} + (x_{n-1} - x_{n-2}) = 2x_{n-1} - x_{n-2} \tag{8}$$

The generating function of the transfer function follows directly from the transform of equation (8)

$$H_n(z) = 2z^{-1} - z^{-2} = \frac{2z-1}{z^2}. \tag{9}$$

The behaviour of this very simple system in response to a step function is illustrated by the dashed lines in Figure 1. This figure leaves no doubt however, that subjects usually do not behave in as ideal a way as the model requires. The question is if the observed results can be accounted for by introduction of parameters and if it is possible to describe the behaviour of different subjects under
different conditions in terms of these parameters rather than by altering the basic properties of the model.
Since a predictive system like the ILP implies feed back of information, we may write $H_n(z)$ in terms of a relation with feedback (Truxal, 1955, p. 519):

$$H_n(z) = \frac{G_n(z)}{1+G_n(z)}. \tag{10}$$

From equations (6) and (7) it follows that

$$G_n(z) = \frac{2z-1}{(z-1)^2}. \tag{11}$$

The parameters we may introduce in equation (11) represent factors which determine how strongly the present output interval will depend on current and previous information fed into the system, *a* being an amplification factor and *b* representing an exponential factor which determines the rate at which the correct synchronization relation is re-established when correct synchrony is lost:

$$G_n'(z) = \frac{a(2z-1)}{(z-b)(z-1)}. \qquad (12)$$

For ease of computation we make the substitutions $\alpha = 1 - a$ and $\beta = 1 - b$, whence it follows that:

$$H_n'(z) = \frac{(1-\alpha)(2z-1)}{z^2-(2\alpha+\beta)z+(\alpha+\beta)}. \qquad (13)$$

By long division or partial fraction expansion we obtain an expression for y_n in terms of the input function $x(n\text{-}k)$

$$\begin{aligned} y_n &= (2-2\alpha)x_{n-1} + \\ &+(-1+5\alpha+2\beta-4\alpha^2-2\alpha\beta)x_{n-2} + \\ &+(-4\alpha-3\beta+12\alpha^2+11\alpha\beta+2\beta^2)x_{n-3} + \\ &+(\alpha+\beta-13\alpha^2-17\alpha\beta-5\beta^2)x_{n-4} + \\ &+(6\alpha^2+10\alpha\beta+4\beta^2)x_{n-5} + \\ &+(-\alpha^2-2\alpha\beta-\beta^2)x_{n-6} \end{aligned} \qquad (14)$$

Since $a \approx 1$ and $b \approx 1$ it follows that $\alpha \approx 0$ and likewise $\beta \approx 0$, and that in general all terms of third and higher degree in this expansion can be neglected in, as in equation (14). In the experiments that were carried out in this study I found $\alpha_{av} = 0.17$ (range 0.04 – 0.32) and $\beta_{av} = 0.11$ (range 0.00 – 0.24), which brings all quadratic terms to within the range 0.00 – 0.10. Note that if α and β are exactly equal to 0, the model reduces to the ideal linear predictor as in equation (8).

References

Cohen, J., Hansel, C. E. M., & Sylvester, J. D. (1953). A new phenomenon in time judgment. *Nature, 172*, 901-903

Creelman, C. D. (1962). Human discrimination of auditory duration. *Journal of the Acoustical Society of America, 34*, 582-593.

Feller, W. (1957). *An introduction to probability theory and its applications* (2nd edition). New York: Wiley.

Helson, H., & King, S. M. (1931). The tau-effect: An example of psychological relativity. *Journal of Experimental Psychology, 14*, 202-217.

Hirsh, I. J., & Sherrick, C. E., Jr. (1961). Perceived order in different sense modalities. *Journal of Experimental Psychology, 62*, 423-432.

Michon, J. A. (1966a).: The problem of perceptual load. In J. J. Vos (Ed.), *Studies in perception dedicated to M. A. Bouman*. Soesterberg, The Netherlands: Institute for Perception RVO-TNO.

Michon, J. A. (1966b). Tapping regularity as a measure of perceptual motor load. *Ergonomics 9*, 401-412.

Michon, J. A. (1967) *Timing in temporal tracking.* Soesterberg, Netherlands: Institute for Perception RVO-TNO.

Michon, J. A., & Van der Valk, N. J. L. (1967). A dynamic model of timing behavior. *Acta Psychologica, 27*, 204-212.

Rasch, R. (1981). *Aspects of the perception and performance of polyphonic music*. PhD Dissertation, Faculty of Social Sciences, University of Groningen, The Netherlands.

Seashore, R. H. (1951). Work and motor performance. In S. S. Stevens (Ed.), Handbook of experimental psychology. New York: Wiley.

Schwarz, R. J., & Friedland, B. (1965). *Linear systems*. New York: McGraw-Hill.

Truxal, J. G. (1955). *Automatic feedback control system synthesis*. New York: McGraw-Hill.

Link to author's website:
http://www.jamichon.nl