Biol. Cybern. 56, 305–311 (1987)

Biological Cybernetics

Synchronizing Human Movement with an External Clock Source

D. Hary and G. P. Moore

Department of Biomedical Engineering, University of Southern California, Los Angeles, CA 90089-1451, USA

Abstract. Temporal information processing was studied in humans attempting to tap a key in synchrony with a metronome whose base period was subjected to subliminal random changes. Statistical measures of the sequential timing of metronome and key-tap events were compared with similar time series generated by computer-simulated models of synchronization strategies. From this analysis, synchronization appears to be a sensory-dependent second-order recursive process indexed and stabilized by a combination of internal and external resetting events, with at least two independent sources of timing error.

Introduction

The ability of humans to tap in synchrony or near synchrony with a periodic time source such as a metronome has been studied for over a century (Shaffer 1982). Skillful subjects generate a sequence of motor acts whose average intervals are equal to the mean interval of the metronome, and place that sequence of movements so that their terminal points nearly coincide, temporally, with the external source. Thus, they effectively track the sequence of metronome times. Yet, unlike the usual stimulus-response situation, the "response" (motor act) in this task generally is not delayed relative to the "stimulus" (the nearly coincident metronome click sound) and often precedes it. Indeed, this task is predictive, requiring subjects to initiate their tap well before the anticipated target sound.

We presume that the duration of the motor act required to synchronize with the metronome is initially estimated by listening to the source. Subsequently, information derived from the timing of taps in relation to the source is used to alter the duration of the motor act and its temporal placement relative to the metronome. Evidently this information is restricted by uncertainties in the perception of synchronicity, and confounded by uncontrollable timing variability within the motor system.

A number of models have been proposed to account for human metronome tracking behavior (Fraisse and Voillaume 1971; Michon 1967; Michon and Van der Valk 1967; Voillaume 1971). These models either implicitly or explicitly share the postulate, which we retain, that some event serves as a marking event, acting operationally to reset a cyclic process to a local zero time. The "reset" event marks the beginning of a period which culminates in the discrete event of tapping. The duration of this process, which we call the "reference interval", includes both motor and non-motor delays, and is adjusted so that the tapping events nearly coincide with the metronome events. Thus, temporal tracking strategies entail two essential components: a resetting event, and a procedure to establish a reference interval for the duration of the response process, which follows this event.

Previously hypothesized tracking strategies identify the subject's own taps as the exclusive resetting events; the cyclic process is initiated and terminated by the tap itself, with the subject establishing the duration of the intervening periods between taps (Fraisse and Voillaume 1971; Michon 1967; Michon and Van der Valk 1967; Voillaume 1971). We have shown elsewhere (Hary 1987), that if this tap-reset process proceeds without correction, the errors from previous cycles accumulate and the variance of the synchronization error (defined as the interval between the metronome and the tap event intended to synchronize with it) increases monotonically with time. If the subject thus effectively ignores the reference time afforded by the metronome, the probability that the synchronization error will eventually exceed the metronome interval increases asymptotically to one. This is a special case of a Markovian auto-regressive process that "wanders" with increasing variance (Kendall 1976). Even a subject who has established a reference interval exactly equal to the metronome interval will, as a result of his inability to reproduce that interval with absolute precision, accumulate a synchronization error that grows without bound. Such instability, manifested by a drifting of the tap in relation to the metronome, is not observed empirically.

In a tap-reset strategy that includes an error correction scheme, a subject can either adjust the reference interval, or keep a constant reference interval and simply feed forward the current error to regulate the time of the next tap (Fraisse and Voillaume 1971; Voillaume 1971). The former process is equivalent to a second-order auto-regressive, or Yule (Kendall 1976) process, but it is, at best, only marginally stable, and in all practical situations produces unstable and divergent tracking behavior (Hary 1987). Our previous data (Hary and Moore 1985) argue against any type of simple tap-reset temporal tracking strategy: In response to a 10 ms step change in the metronome interval, the tap interval sequence undergoes gradual, stable, and sustained changes, which eventually return the synchronization error to its initial distribution.

An obvious alternative is that the metronome is the resetting event. In the simplest case, the metronome establishes a "local" time reference, and the next tap occurs after a delay process whose duration is not adjusted by the error it generates, but may include a random component. This random component is independent of, and uncorrelated with, the metronome. The initial duration of the delay is established, perhaps, during the initial learning period. Although stable in operation, this model was rejected because it contradicts the observed correlation between metronome and delay intervals, and between successive delay intervals (Hary and Moore 1985).

Our aim in the present investigation was to identify the systematic process by which humans achieve synchronicity in this timing task. We conclude that humans employ the exogenous metronome source in two ways: firstly, as an occasional independent marker to re-initialize the tracking process and thus provide an absolute and independent time reference; and secondly, in conjunction with the tap, to estimate deviations from synchronicity which can then be used to adjust, recursively, the basic period of the response.

Methods

Subjects who participated in the study were skilled musicians (instructors and graduate students from the USC School of Performing Arts). All subjects were informed of the purpose and procedures of the study and signed an informed consent.

Subjects were seated comfortably with the arm of their preferred hand resting on a table in proximity to a Morse telegraph key. A loudspeaker (Radio Shack Minimus-7), which presented audible clicks serving as metronome sounds, was placed 0.5 m in front of the subject. Electronically generated pulses of 0.4 ms duration were delivered to the loudspeaker via an audio amplifier (Heathkit AA-14). The subject could adjust the loudness of the click.

The delay of the metronome click in reaching the subject was 1.8 ms. This was measured by placing a microphone (Teac MC-10 moving coil cardoid microphone) 0.5 m in front of the loudspeaker. An oscilloscope (Tektronix 5013) was used to measure the metronome pulses before amplification and the signal picked up by the microphone. Time delay was measured from the beginning of the generator pulse to the point of highest amplitude in the microphone signal. A delay of 0.4 ms was found when the microphone was placed as close as possible to the loudspeaker.

Metronome pulses were generated by an integrate-and-reset circuit in which a voltage increased linearly with time until reaching a controlled variable threshold level. In one procedure small Gaussian random and uncorrelated interval perturbations were produced by varying the threshold voltage with a noise generator (Grason-Stadler model 901B). In another procedure, small square-wave changes in the threshold voltage were used to vary the metronome period between two durations. In both cases, the variability introduced into the metronome interval distribution was considerably smaller than the natural variability of the subject. In a third procedure, the metronome period was kept constant. The subjects were not informed about the three metronome sequences, and in no case did a subject report noticing any differences between these sequences.

The transition times of the metronome pulses and the pulses generated by the down-stroke of the Morse key were digitized to the nearest 0.1 ms by a computer (PDP-11/23). The Morse key pulses were processed, before computer acquisition, by a special circuit to prevent multiple triggering of the digitizer due to key bounces. This circuit ensured that at least 250 ms passed between successive tap events.

Subjects were instructed to tap the key "on the beat" so that the tap sound resulting from the down-stroke of the Morse key would coincide, as closely as possible, with the click of the metronome. The sound of the upstroke of the Morse key was muffled such that its peak amplitude, as measured by a microphone placed 0.5 m in front of the Morse key, could not be distinguished from other background noise. Each subject was free to choose the pattern of movement to execute the key tap. Subjects could also adjust the resistive spring tension and the traverse distance of the Morse key lever to their comfort. The Morse key mechanical sound was clearly distinguishable from the loudspeaker clicks.

Practice trials consisted of a few minutes of presentation of the strictly periodic metronome at an interval different from the one presented during the data runs. When the subject was familiar with the procedure, a sequence of 120–140 metronome clicks with a mean period of 700 ms and a particular perturbation pattern (random, strictly periodic, or step) was presented. A complete test session consisted of a randomized series of these metronome patterns. Rest periods of at least one minute were introduced between runs.

The first 25 metronome events with their associated taps were discarded from the data. This ensured that any initial transients that might arise while the subjects acquired the basic metronome period were excluded from the analysis. Figure 1 shows the event data from one test run where the metronome period was subjected to small random Gaussian variations. This figure serves also to define the metronome, tap, delay and error intervals.

The dependence of the subject-generated intervals on the metronome intervals was quantified by means of correlation coefficient functions (auto- and cross-correlograms) between the time series of event intervals (e.g., Figs. 2, 4) (Chatfield 1984; Kendall 1976),

$$C_x(k)=\sum_{i=1}^{N-k}(X_i-\mu_x)(X_{i+k}-\mu_x)/(N-k),\quad k=0,1,\ldots,30, \tag{1}$$

$$\mu_x=\sum_{i=1}^{N}X_i/N, \tag{2}$$

$$r_x(k)=C_x(k)/C_x(0), \tag{3}$$

$$C_{xy}(k)=\begin{cases}\displaystyle\sum_{i=1}^{N-k}(X_i-\mu_x)(Y_{i+k}-\mu_y)/(N-k), & k=0,1,\ldots,30,\\ \displaystyle\sum_{i=1-k}^{N}(X_i-\mu_x)(Y_{i+k}-\mu_y)/(N+k), & k=-1,2,\ldots,-30,\end{cases} \tag{4}$$

$$r_{xy}(k)=C_{xy}(k)/\sqrt{C_x(0)C_y(0)}, \tag{5}$$

$$CI(k)=2/\sqrt{(N-k)}, \tag{6}$$

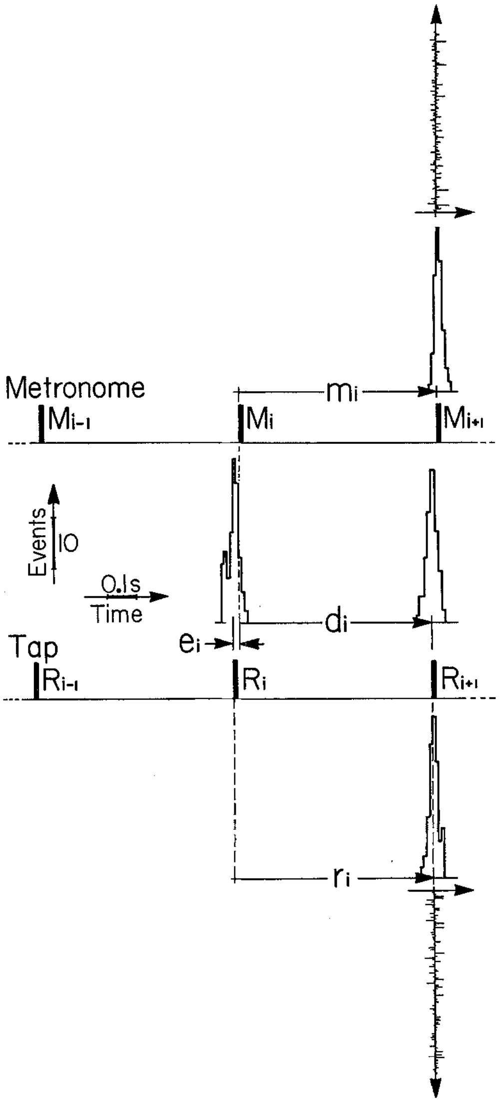

Fig. 1. Data from a subject attempting to synchronize with a metronome whose interval duration includes small random additive and uncorrelated variation. On the horizontal time axis we define event times (upper case) and time intervals (lower case). Plotted also are the interval histograms for the metronome (*m*) (704 ± 12 ms, mean ± standard deviation), synchronization error (*e*) (−22 ± 21 ms), delay (*d*) (682 ± 19 ms), and tap (*r*) (704 ± 15 ms) intervals. The actual 107 metronome and tap intervals acquired in this test are plotted along the vertical (event or index) axis as deviations from the mean

where $C_x(k)$ is the sample auto-covariance function, $r_x(k)$ is the sample auto-correlation coefficient function, $C_{xy}(k)$ is the sample cross-covariance function, and $r_{xy}(k)$ is the sample cross-correlation coefficient function. μ is the sample mean and $C(0)$ is the sample variance for the population of N time intervals in the time series $\{X_i\}$ and $\{Y_i\}$. $CI(k)$ is the approximate 95% confidence interval for testing the hypothesis that the correlation coefficient is not significantly different from zero (Chatfield 1984). Thus, the correlogram consists of a sequence of correlation coefficients whose values lie between plus and minus one. The coefficients are arranged graphically according to an index variable k, ranging here from -30 to $+30$, which indicates the number of intervals separating the two variables being correlated. For variables in adjacent intervals, k is plus or minus one.

Models of synchronization strategies were simulated by computer programs written in FORTRAN IV (Digital Equipment). These strategies were cast as difference equations as well as indexed block diagrams (e.g., Fig. 3). The data from the simulation runs were subjected to the same analysis procedures as the laboratory data.

The general procedure we followed in validating the simulated models was to compare, first, the global structures of the simulated and experimentally derived auto- and cross-correlograms of the metronome, tap, delay, and synchronization-error intervals. Obvious departures of the simulated data from the general structure of the experimental data resulted either in a change in model parameters, or, when the global features could not be obtained for any choice of parameters, in the outright rejection of that model. Next, we looked for certain specific features in the correlograms that were critical to the testing of a model; for example, the algebraic sign of a particular coefficient. When the restricted sample size of the laboratory data left some uncertainty about the values of the coefficients, we used additional methods for making more accurate determinations of the correlation coefficients (Akaike 1969; Anderson 1974). For lack of alternative evidence, all hypothesized random variations were simulated using uncorrelated, normally-distributed random numbers with the FORTRAN IV linear congruential function (RAN) using the Box-Muller method (Box and Muller 1958).

Result

Data were obtained from 5 musicians for all of whom the mean synchronization error was negative, indicating, as can be seen from the error histogram in Fig. 1, that the tap preceded the metronome. The lead was observed in all subjects and all tests and ranged from 5 to 90 ms (for 33 tests), with a mean of 24 ms and a standard error of the mean of 2.7 ms. Consequently, the mean delay interval was consistently smaller than the metronome interval.

The random and independent variations in metronome intervals ensure the absence of any significant correlation between its interval durations, giving its auto-correlogram an impulse-like structure (Fig. 2d). Significant correlations were found between successive delay intervals (Fig. 2a), and between metronome and delay intervals (Fig. 2c). Because the metronome interval correlogram lacks any structure, the presence of a temporal microstructure in the subject-generated tap sequence, revealed in these cross-correlograms, is not derived from any intrinsic structure in the metronome sequence itself, but rather can only be a consequence of the underlying strategy of synchronization executed by the subject, i.e. the causal, functional, and dynamic dependency of the tap intervals on the metronome.

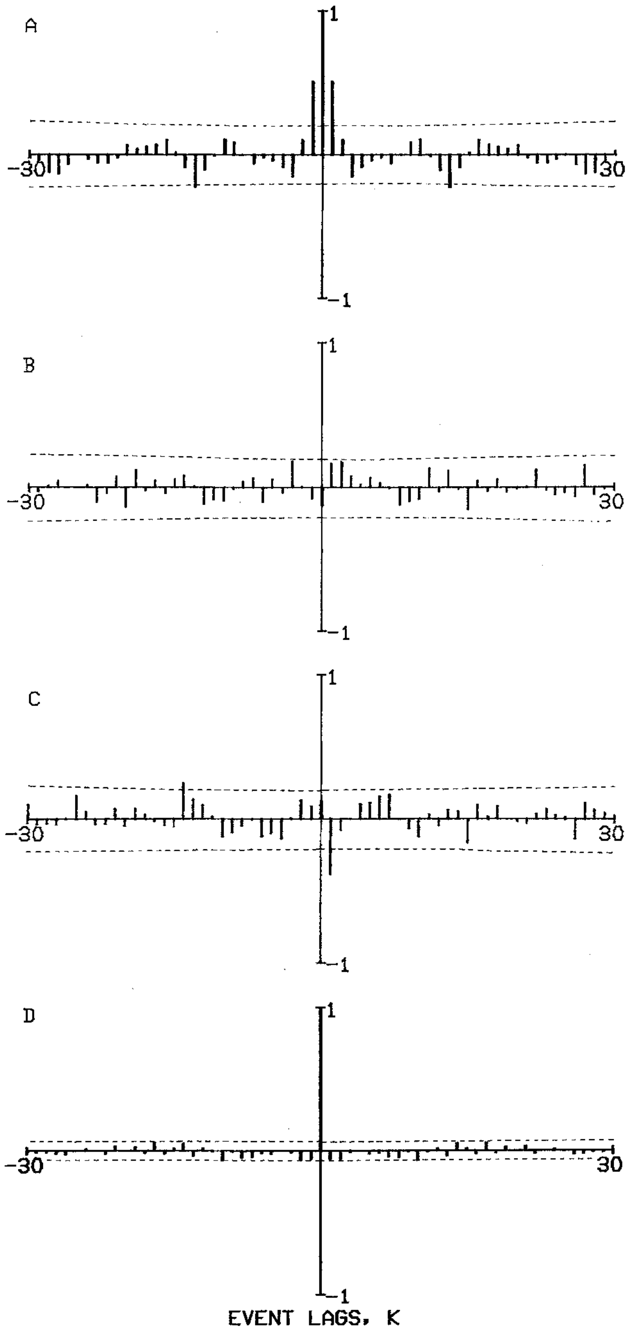

Fig. 2a–d. Auto- and cross-correlograms derived from the data of Fig. 1. **a** Auto-correlogram of the delay intervals. **b** Cross-correlogram between the metronome and the tap intervals. **c** Cross-correlogram between the metronome and delay intervals. **d** The auto-correlogram of 1000 metronome intervals. The dashed line represents the 95% confidence interval for the correlation coefficient

Metronome-reset strategies without error correction cannot, by definition, produce correlations between metronome and delay intervals, or between

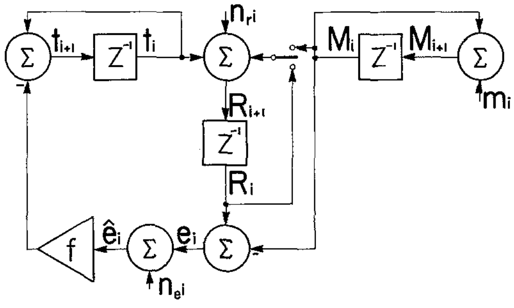

Fig. 3. A block diagram of a tracking strategy that has a variable reset source (symbolized by a switch), and a reference interval (t) which determines, except for motor variability, the delay of the next tap (R_{i+1}) following the reset event. The switch allows a given decision rule to use the time of either the metronome (M_i) or the tap (R_i) as a reference for the next tap. A fraction (f) of the estimated synchronization error ($\hat{e}$) is used to correct the reference interval. The system is affected by two sources of internal noise: noise in the motor system (n_r), which determines the actual tap time, and noise (n_e) in the process which estimates the synchronization error (e). The Z^{-1} blocks are operators that set their current output equal to their previous input, an operation equivalent to decrementing the index variable (i) by one. Thus, for example, the metronome circuit block yields the equation $M_{i+1} = M_i + m_i$ where m_i is the metronome interval. For convenience, we chose the initial condition for t as equal to the mean metronome interval

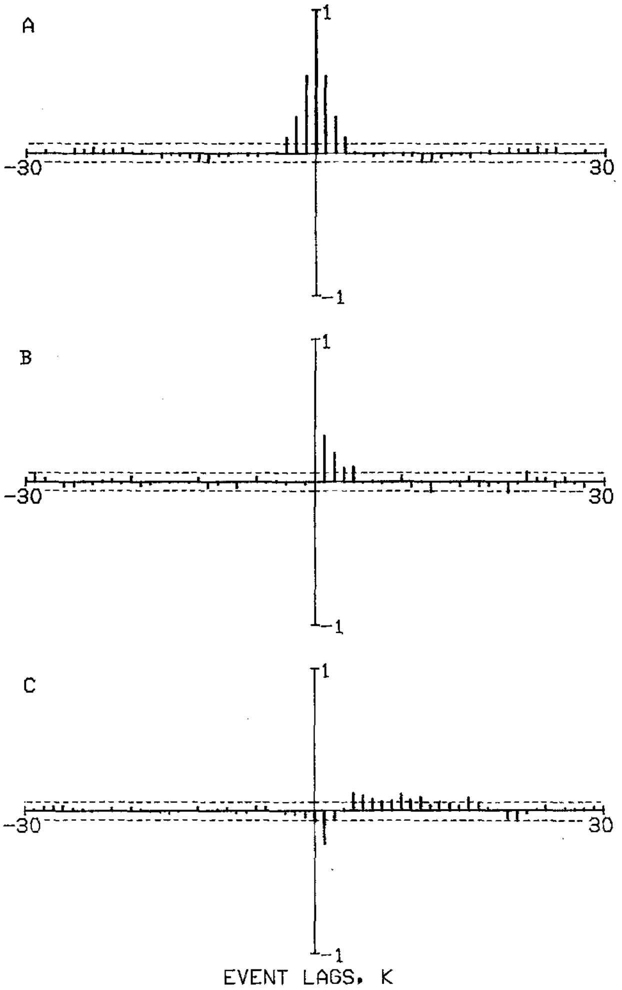

Fig. 4a–c. Correlograms of data from a model that tracks a "random" metronome using an error-correcting, random-reset strategy. **a** The delay auto-correlogram. **b** Cross-correlogram between the metronome and the tap interval. **c** Cross-correlogram between the metronome and the delay interval (Cf. Fig. 2a–c). Parameters were adjusted to fit the correlations of Fig. 2 and a mean lead of 15 ms. These parameters are: mean metronome period of 700 ms with 10 ms standard deviation; random motor variability (n_r) with 8 ms standard deviation; feedback fraction (f) of 0.05 of the synchronization error estimated in the preceding cycle; and synchronization error estimation bias of 15 ms and standard deviation of 40 ms. Gaussian distributions were used for all the random variables. The correlogram of the metronome sequence was indistinguishable from that of Fig. 1d

successive delay intervals, when the metronome sequence is random and uncorrelated. Moreover, stable metronome-reset strategies with error correction do not produce the *negative* first-order correlations between the metronome and delay intervals such as those seen in the laboratory data (Fig. 2): In a metronome-reset regime, a metronome interval larger than the mean would delay the mean time of the next tap; this lag in the tap is equal, in the mean, to the change in metronome interval if the subject is oblivious to the interval change, but would be larger than the metronome change if the subject responded to the increase in the metronome period with an increase in the internal reference interval. This leads to a non-negative correlation between the metronome and the delay intervals. Because our subjects always[1] exhibited negative first-order correlation between metronome and delay, we are forced to reject the possibility of pure metronome-reset strategies, with or without error correction.

Because our previous analysis (Hary 1987) showed response-reset strategies to be unstable, we considered the possibility of "mixed-reset" strategies. Such strategies, using both tap and metronome events as resets, might involve a random switching between

---

[1] In two other subjects we did observe auto-correlation coefficients in the delay variable and cross-correlations between the metronome and delay that were not significantly different from zero. That suggested that these subjects had established, during the first few cycles of the test, an internal reference interval which was used without further significant modification for the remainder of the task. Their performance, in other words, could be accounted for by a metronome-reset strategy with a fixed reference interval and variable motor delay. Metronome-reset models that ignore the error are still unconditionally stable in that they do not lead to a divergent synchronization error. Both subjects later modified their strategy and eventually displayed negative first-order correlations between the metronome and delay time series

classes of reset events, or a deterministic process in which the reset event would be, for example, either the first or the last of the two events in each cycle. The presumption of a mixed reset vastly complicates the mathematical analysis of any model, but poses little additional difficulty for computer simulations of its behavior.

Our proposed mixed-reset model of metronome tracking is summarized formally, as it was programmed, in Fig. 3. A switch connecting either the metronome or the tap to the resetting junction symbolizes the mixed-reset mode of operation. The feedback loop connecting the synchronization error (e) to the reference interval (t) symbolizes the path by which the experience of the tracker appears to modify his behavior. The exogenous input to the model is the current metronome event time (M_i) which serves to establish the synchronization error, and sometimes also to establish the reset time. Time, in this temporal information processing strategy, is the dependent variable; it is indexed and ordered by the discrete independent variable (i) which is the synchronization cycle index counter. The index counter is decremented by the two endogenous shift operators (Z^{-1}) which give the model its second-order dynamic properties.

With the use of mixed reset and the feedback adjustments of the reference interval (t) based on the synchronization error, we generated simulated data whose correlations (Fig. 4) closely resembled the structure of the empirical correlograms shown in Fig. 2. The simulations fit best when approximately equal numbers of metronome and tap resets were used.

Discussion

If our model of synchronization is valid, it sheds light on a puzzling feature of our own and previous metronome tracking data which consistently show that the subject's tap, on the average, precedes the metronome. The structure of the model shown in Fig. 3 implies that any persistent error of this kind would be corrected by feedback unless it originated in the feedback loop itself, e.g., in the process that estimates the synchronization error. A persistent bias in this process would not be corrected, and therefore we have introduced a net error into the feedback loop in order to match the observed lead of our subjects.

This lead, of course, is in the laboratory frame of reference which has some arbitrary elements in the definition of the "time" of the metronome and tap events. It is possible that in the central nervous system there is a critical node at which the average synchronization error is, in fact, either zero or perceived to be zero even when, in the laboratory frame of reference, it is negative. For example, there may be a locus of convergence in the brain for two sensory modalities with different conduction delays, a point at which temporal information about the metronome event is compared with information about the tap event. The process at this locus might include several time-consuming sub-processes: detection and identification of the metronome event; detection and identification of the tap event (which would require both kinesthetic and auditory inputs); a decision about the order of the detected events; and a decision about the time difference between them (Mitrani et al. 1986). But these are speculations, and until we know more about it, the puzzling finding of a significant lead in all subjects must be regarded as unexplained.

This possibility makes a random-reset strategy (with approximately equal probabilities for each reset event) nearly equivalent to using first-event or last-event reset, and if the reference interval correction is based on the synchronization error of the preceding cycle, the error is always available at the moment of reset regardless of the identity of the reset event. Indeed, a somewhat better fit to the empirical correlations was obtained in this way.

In addition to the bias associated with the process of estimating the synchronization error, we have also introduced noise. In order to account for the statistical character of our data we had already followed the practice of others in assuming a random variability in the motor response. But our simulations showed that randomness in the duration of the motor component of the response alone is insufficient to account for the observed correlations, and that an additional independent source of variability was necessary. It seemed plausible to associate that source with the presumably "sensory" process that estimates the synchronization error. In our model this error estimation process (with a standard deviation of 30–40 ms and a bias of 10–50 ms) is used to regulate a centrally stored reference interval, but, given the conservative feedback coefficient, the subject adjusts this stored interval by only a millisecond or so in a typical cycle. The noise contribution to the reference interval is of comparable magnitude.

We estimate that even in the presence of a fixed prediction interval, the motor response independently contributes a timing variability with an 8–15 ms standard deviation. The effect of this randomness on the subsequent synchronization error is indistinguishable from that introduced by the randomness in the metronome interval. Due to the incorporation of subliminal variations in the metronome period, both the motor and metronome randomness are unknown to the subject and presumably treated identically by the nervous system.

The tracking data shown in Fig. 1 exhibited a variability in both the tap interval and the synchronization error. According to our model, roughly a third of the variability in the tap interval and synchronization error (Fig. 1) is due to uncontrollable variability in the response process. The remainder is primarily due to variability in the time of the reset event. Noise in the estimation of the synchronization error is relatively well buffered by the low feedback coefficient. Thus, noise in the overall (motor) response system would be the primary contributor to variability in the tracking of a strictly periodic metronome using a metronome reset strategy or a mixed-reset strategy without a lead. In a mixed-reset strategy, the existence of a lead further increases the variability in the performance-defined intervals.

A paradoxical feature of the strategy summarized in Fig. 3 is that the postulated reference interval, the variable modified by the experience of the tracker, is not even asymptotically equal to the mean metronome interval. It could only be equal, with metronome reset, if there were no lead in the time of the tap relative to the metronome. Given the observed lead, the reference interval could only be equal to the metronome period if all resets were tap-generated. With our postulated mixture of reset events, the reference period must be shorter than the metronome period.

Our model of temporal information processing in the nervous system, which accounts for both the quantitative and qualitative features of the random-metronome tracking data, and for our previous data obtained from tracking subliminal step changes in the metronome period, is not as simple nor as intuitive as we had expected for such a seemingly trivial task. Yet it appears to be employed in an entirely unconscious way by our subjects, who have, in some way not yet known, managed to acquire it and thereby to avoid the inaccuracies and instabilities that, according to our analysis, are inherent in simpler tracking strategies. Humans, according to this model, can measure, process and store temporal information and use it to direct their motor acts to achieve synchronicity with a relative precision of 2–3%. It remains to be seen whether this basic model is adequate for more complex synchronization tasks.

Acknowledgements. This research was partially supported by NIH grant RR-01861 and by BRSG S07RR07012-19 grant awarded by the Biomedical Research Support Grant program, Division of Research Resources, NIH.

References

Akaike H (1969) Fitting autoregressive models for prediction. Ann Inst Statist Math 21:243–247

Anderson M (1974) On the calculation of filter coefficients for maximum entropy spectral analysis. Geophysics 39:69–72

Box GEP, Muller ME (1958) A note on the generation of random normal deviates. Ann Math Stat 29:610–611

Chatfield C (1984) The analysis of time series. Chapman and Hall, London New York

Fraisse P, Voillaume C (1971) Les repères du sujet dans la synchronisation et dans la pseudo-synchronisation. Annèe Psychol 71:359–369

Harry D, Moore GP (1985) Temporal tracking and synchronization strategies. Human Neurobiol 4:73–76

Hary D (1987) On the stability and performance of human synchronization strategies (submitted for publication)

Kendall M (1984) Time series. Hafner Press, New York

Michon JA (1967) Timing in Temporal Tracking. Institute for Perception RVO-TNO, Soesterberg, The Netherlands

Michon JA, Van der Valk NJL (1967) A dynamic model of timing behavior. Acta Psychol 27:204–212

Mitrani L, Shekerdjiiski S, Yakimoff N (1986) Mechanisms and asymmetries in visual perception of simultaneity and temporal order. Biol Cybern 54:159–165

Shaffer LH (1982) Rhythm and timing in skill. Psychol Rev 89:109–122

Voillaume C (1971) Modèles pour l'étude de la régulation de mouvements cadences. Annèe Psychol 71:347–358

Received: September 22, 1986

Dr. David Hary
Dr. George P. Moore
University of Southern California
Department of Biomedical Engineering
University Park
Los Angeles, CA 90089-1451
USA