Biol. Cybern. 76, 339–347 (1997)

Biological Cybernetics


# Spectral properties of human cognition and skill

**Jeff Pressing, Garry Jolley-Rogers**

Department of Psychology, University of Melbourne, Parkville, Victoria 3052, Australia

Received: 4 October 1995 / Accepted in revised form: 25 February 1997

**Abstract.** Many interactive human skills are based on real-time error detection and correction. Here we investigate the spectral properties of such skills, focusing on a synchronization task. A simple autoregressive error correction model, based on separate 'motor' and 'cognitive' sources, provides an excellent fit to experimental spectral data. The model can also apply to recurrent processes not based on error correction, allowing commentary on previous claims of $1/f$-type noise in human cognition. A comparison of expert and non-expert subjects suggests that performance skill is not only based on reduced variance and bias, but also on the construction of richer mental models of error correction.

## 1 Introduction

Recent research has suggested that human cognition and skill, as found in sequential time and distance estimation tasks (Gilden et al. 1995), speech (Voss and Clarke 1975), bimanual coordination (Schmidt et al. 1992) and music (Voss and Clarke 1978; Voss 1988) may display $1/f$-type power density spectra. Although strong methodological criticisms of assertions of $1/f$ noise in music and speech have been previously mounted (Nettheim 1992; Pressing 1994), spectral shapes of this form are widely found in many physical systems, and the search for a common mechanism is long-standing in the physical sciences (Voss and Clarke 1978; Miller et al. 1993; Gilden et al.). So far, no cognitive or motor mechanisms that explicitly produce such spectral shapes have been proposed. Here we investigate the effectiveness of a general model of human skill based on temporal error compensation and show that it provides an excellent fit to spectral data from simple synchronized tapping tasks, without the need to postulate $1/f$ noise sources. The model can also be extended to non-synchronized timing tasks, which allows a direct discussion of possible factors creating $1/f$ type effects.

*Correspondence to:* J. Pressing (Fax: + 61-3-9347-6618, e-mail: jlp @psych.unimell.edu.au)

## 2 The autoregressive tapping model

Consider a person who taps as synchronously as possible to a steady repeating auditory tone. If the asynchrony between the tap and the tone at time $n$ is defined as $A_n$, the intertap interval $I_n$ is

$$I_n = P + A_{n+1} - A_n \tag{1}$$

where $P$ is the intertone interval.

Another expression for $I_n$ is given by the basic two-tiered Wing-Kristofferson model (Wing and Kristofferson 1973a, b), which describes the timing of human rhythmic production in the absence of error correction:

$$I_n = C_n + (M_{n+1} - M_n) \tag{2}$$

This is based on an internal cognitive clock process generating regular intervals $C_n$, whose endpoints trigger motor system actions characterized by delay intervals $M_n$. The central feature of this process description is that if the $C_n$ and $M_n$ are treated as uncorrelated noise processes, (2) yields zero autocorrelations for lags greater than 1. This equation, and its extensions to more complex motor patterns, have been shown to provide an excellent quantitative account of the covariance structure of human performance (Wing and Kristofferson 1973a, b; Wing 1980; Vorberg and Hambuch 1984; Wing et al. 1987; Jagacinksi 1988; Turvey et al. 1989; Pressing et al. 1996).

This formulation can be extended to include error correction by supposing that (2) is modified by a correction of the clock interval $C_n$ based on the most recent asynchronies between tone and tap. If only the last asynchrony affects the process, then an additional linear correction term of type $-\alpha A_n$ appears, as has been considered by previous authors in various forms (Hary and Moore 1987; Schulze 1992; Mates 1994a, b; Vorberg and Wing 1996), albeit purely in the time domain. In the experimental work here lag 2 effects appear, and so we propose a more general error correction expression of the form $-\alpha A_n - \beta A_{n-1}$. Given this expression, (2) becomes

$$I_n = C_n - \alpha A_n - \beta A_{n-1} + (M_{n+1} - M_n) \tag{3}$$

Combining this with (1) yields

$$A_{n+1} = (1-\alpha)A_n - \beta A_{n-1} + (C_n - P) + (M_{n+1} - M_n) \quad (4)$$

It is convenient here to measure asynchronies relative to their mean positions, which are normally close to zero[1].

This is a second-order autoregressive equation in the asynchronies, and differs from a standard ARMA(2, 1) (autoregressive moving average) equation only in the nature of the noise process represented by the last four terms, which here is autocorrelated at lag 1 and therefore not white. From this equation, under a condition of (second-order) stationarity, the autocovariance functions at lag $k$, notated $\gamma_A(k)$, can be calculated. Taking the variance of both sides of (4) allows computation of the series variance as

$$\gamma_A(0) = \sigma_A^2 = \frac{\sigma_C^2 + 2\alpha\sigma_M^2 - 2\beta(1-\alpha)\gamma_A(1)}{1-(1-\alpha)^2-\beta^2} \quad (5)$$

where $\sigma_C^2$ and $\sigma_M^2$ are the clock and motor variances, respectively. In this derivation we have used the property that, defining

$$I_n^* = C_n + (M_{n+1} - M_n) \quad (6)$$

as the interonset interval in the absence of error correction,

$$\mathrm{Cov}(I_n^*, A_{n-l}) = -\sigma_M^2 \delta_{l0} \quad (7)$$

for all $l \geqslant 0$, where $\delta_{l0}$ is the Kronecker delta. Equation (7) has been previously given by Vorberg and Wing 1995; and relies on the fact that, by its definition, $I_n^*$ is a random variable unaffected by times earlier than $n$.

By taking the covariance of (4) with $A_n$ we similarly obtain

$$\gamma_A(1) = \frac{(1-\alpha)\gamma_A(0) - \sigma_M^2}{1+\beta}. \quad (8)$$

For $k \geqslant 2$, the autocovariance functions satisfy the Yule-Walker equations.

$$\gamma_A(k) = (1-\alpha)\gamma_A(k-1) - \beta\gamma_A(k-2), \quad k \geqslant 2 \quad (9)$$

which may readily be derived by taking the covariance of each side of (4) with $A_{n-k+1}$, in light of (7).

The general solution of these equations is

$$\gamma_A(k) = \lambda 2^{-k}\left[\left(1-\alpha+\sqrt{(1-\alpha)^2-4\beta}\right)^k + \left(1-\alpha-\sqrt{(1-\alpha)^2-4\beta}\right)^k\right] \quad (10)$$

valid for $k \geqslant 2$, where

$$\lambda = \frac{(1-\alpha)\gamma_A(1) - \beta\gamma_A(0)}{(1-\alpha)^2 - 2\beta} \quad (11)$$

The Yule-Walker equations can be rearranged to provide estimates of both alpha and beta directly from the time series measurements:

$$\alpha = 1 - \frac{\gamma_A(k-1)\gamma_A(k) - \gamma_A(k-2)\gamma_A(k+1)}{\gamma_A^2(k-1) - \gamma_A(k-2)\gamma_A(k)} \quad (12)$$

$$\beta = \frac{\gamma_A^2(k) - \gamma_A(k-1)\gamma_A(k+1)}{\gamma_A^2(k-1) - \gamma_A(k-2)\gamma_A(k)} \quad (13)$$

valid for $k \geqslant 2$.

After this estimation process, (5) and (8) can be solved directly for the motor and cognitive variances:

$$\sigma_M^2 = (1-\alpha)\gamma_A(0) - (1+\beta)\gamma_A(1)$$

$$\sigma_C^2 = (\alpha^2 - \beta^2)\gamma_A(0) + 2(\alpha+\beta)\gamma_A(1) \quad (14)$$

From this knowledge of the autocovariance functions, we can determine a covariogram.

## 3 Spectral properties of the model

Spectral properties of our model may be derived from the Wiener-Khintchine Theorem (Gottman 1981),

$$f_A(\omega) = \gamma_A(0) + 2\sum_{k=1}^{\infty} \gamma_A(k)\cos\omega k \quad (15)$$

Here $f_A(\omega)$ is the power spectral density function for the asynchronies and $\omega$ is the angular frequency.

Substitution of (5), (8), (10) and (11) in (15) and simplification yields (see Appendix A)

$$f_A(\omega) = \frac{\sigma_C^2 + 2\sigma_M^2(1-\cos\omega)}{1+(1-\alpha)^2+\beta^2-2(1-\alpha)(1+\beta)\cos\omega + 2\beta\cos 2\omega} \quad \text{(AR2 case)} \quad (16)$$

This is our general functional form for human error correction processes that are linear and second order. We denote this model as AR2. Considered as a spectral model, it has no free parameters, as $\alpha$, $\beta$, $\sigma_M^2$ and $\sigma_C^2$ are all computable from the measured autocovariance functions, as given above.

The numerical fit of the spectrum will depend strongly on values of the first- and second-order error correction parameters $\alpha$ and $\beta$. Investigation of (12) and (13) reveals that for the case $\beta = 0$ (termed AR1), they contain the indeterminacy 0/0, and hence estimation is strongly affected by error and noise effects. Thus, a different estimation procedure for $\alpha$ is preferable for this case, even

[1] The error correction terms are assumed to be based on a null position that is approximated by the mean. By subtracting the mean from each side of the equation, the entire equation can then be put in the form of deviation from mean values.

though formally AR1 is a subcase of AR2. For the AR1 case, we find

$$\gamma_A(k) = \begin{cases} \sigma_A^2 & k = 0 \\ (1-\alpha)^{k-1}\gamma_A(1) & k > 0 \end{cases} \tag{17}$$

$$\text{with } \sigma_A^2 = \frac{\sigma_C^2 + 2\alpha\sigma_M^2}{1-(1-\alpha)^2} \quad \text{and} \tag{18}$$

$$\gamma_A(1) = (1-\alpha)\gamma_A(0) - \sigma_M^2 = \frac{(1-\alpha)\sigma_C^2 - \alpha^2\sigma_M^2}{1-(1-\alpha)^2}. \tag{19}$$

The autocovariance function is a geometric function of lag $k$, and asynchronies of non-adjacent time intervals are correlated. The series is stationary provided $0 < \alpha < 2$. These last results for the AR1 case have been previously derived by Vorberg and Wing (1996).

For $\beta = 0$, we can obtain the AR1 spectral density function from (16):

$$f_A(\omega) = \frac{\sigma_C^2 + 2\sigma_M^2(1-\cos\omega)}{\alpha^2 + 2(1-\alpha)(1-\cos\omega)} \quad \text{(AR1 case)} \tag{20}$$

We can estimate $\alpha$ from the Yule-Walker equations with $\beta = 0$ as

$$\alpha = \alpha(k) = 1 - \left[\frac{\gamma_A(k)}{\gamma_A(1)}\right]^{\frac{1}{k-1}} = 1 - \left[\frac{\gamma_A(k)}{\gamma_A(k-1)}\right] \tag{21}$$

for all $k \geqslant 2$.

When both $\alpha \to 0$ and $\beta \to 0$ (no error correction), (20) reduces to

$$f_A(\omega) = \sigma_M^2 + \frac{\sigma_C^2}{2(1-\cos\omega)} \quad \text{(no error correction)} \tag{22}$$

All three forms for $f_A(\omega)$ can be converted to equivalent predictions about the intertap interval spectral density function $f_I(\omega)$ by use of the linear filtering theorem (Gottman 1981), which from (1) yields:

$$f_I(\omega) = 2(1-\cos\omega) f_A(\omega) \tag{23}$$

In particular, in the case of no error correction, as in non-synchronized ('spontaneous') tapping, the interonset spectral density function will be [from (22) and (23)]

$$f_I(\omega) = 2(1-\cos\omega)\sigma_M^2 + \sigma_C^2 \tag{24}$$

This form displays no low frequency peak and shows that the strict Wing-Kristofferson formulation is inconsistent with a $1/f$ noise form.

## 4 Experimental and analytical procedures

The validity of these models in the spectral domain was investigated by experimental study of synchronous tapping to a repeating auditory tone by two subjects of contrasting experience. Subject A had no previous background in musical performance or the tapping task, whereas subject B is an extensively trained professional percussionist and pianist. Both were self-ascribed right-handers.

The subjects tapped one or both of two keys on a flat table, one with each hand. The keys were connected to an IBM 486 computer equipped with special recording and data analysis software written by Young Ho Kim. Asynchrony data were recorded to an accuracy of 1 ms. The experimental set-up is otherwise identical to that described in Pressing, Summers and Magill (1996).

Synchronization tones were delivered to the right ear. Both subjects performed 8 runs at each of 5 pulse lengths: 250, 375, 500, 750 and 1000 ms. Subject B was also able to perform runs at shorter pulse lengths of 175, 150, 125, and 100 ms. All tapping was performed with the right hand alone, except for runs with $P \leqslant 150$ ms (subject B), which used hand alternation. The length of the runs was $n = 950$ for runs of pulse length $\leqslant 375$ ms, and $n = 600$ for the others. The order of runs was interleaved to minimize learning effects.

Autocovariance properties of the series were calculated after quadratic detrending to help ensure stationarity. Detrending – that is, the use of deviations from a curve of best fit – is commonly used in long time series (Gottman, 1981), and we found here that the effects of quadratic detrending were very minor, causing spectrally only a modest change in the single lowest frequency Fourier bin relative to undetrended data, with no significant implications for overall fit. The spectral density function for each detrended run was then computed using a Hanning-windowed periodogram,

$$D_j = \sum_{n=0}^{N-1} A_n W_n e^{2\pi i n j/N} \tag{25}$$

with $W_n = \frac{1}{2}[1 - \cos(2\pi n/N)]$, normalized to preserve total power relative to the unwindowed periodogram. The optimality of these last procedures was verified by simulation trials using autoregressive models with run lengths and parameter values matching the experimental values.

In the cases where hand alternation was used, which was necessary to allow the expert performer to achieve the fastest speeds, the asynchronies were measured relative to the separate means for each hand. Differences between the hands' mean asynchronies were small, and this procedure yielded a small but consistent improvement in the correspondence between theory and experiment.

## 5 Results and discussion

The covariograms showed an excellent correspondence between theory and data. A sample covariogram for a single run is shown in Fig. 1. However, the main focus here is on the spectral domain, to which we now turn.

Spectral results were consistent for a given subject among the eight runs of each specific pulse size, and hence they were combined to provide optimal parameter estimation. $\alpha$ values, $\beta$ values, spectral density functions and variances were linearly averaged; autocovariances

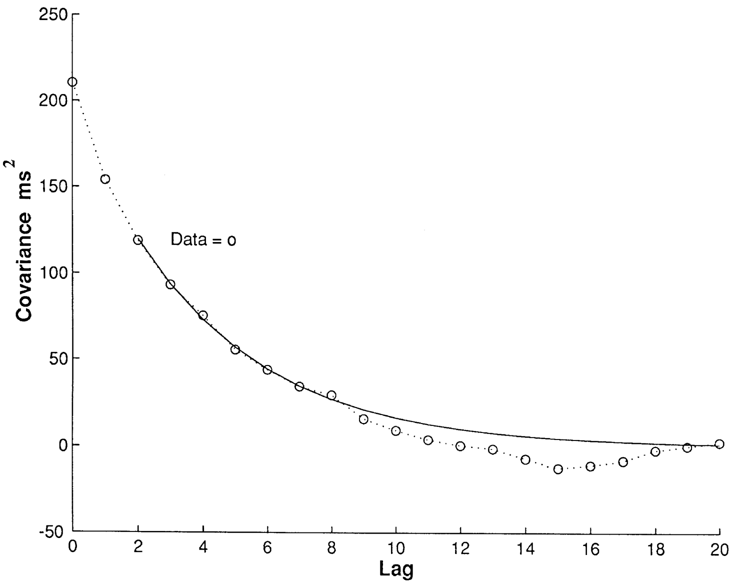


**Fig. 1.** A covariogram for a typical run of 950 taps. Pulse size = 250 ms, subject A. *Dotted line* connects the data points and the smooth curve shows predictions of the AR1 model for $\alpha = 0.220$

were averaged by use of the Fisher z-function technique via conversion to autocorrelations (Pressing et al. 1996). The spectral plots for the two subjects for all pulse sizes are shown in Fig. 2.

For comparison, the best straight line log-log fits (of the form $f_A(\omega) = K\omega^{-L}$, corresponding to $1/f$-type noise) were also fit to the data. (It should be noted that these data do not directly test claims for $1/f$ noise made by Gilden et al. (1995) as their data were recorded under conditions of non-synchronous tapping, unlike here). Table 1 summarizes the model parameters and relative fits to the data.

The results provide strong evidence in favour of the AR1 model for pulse sizes $\geqslant 175$ ms, and for the AR2 model in the range $\leqslant 150$ ms, with a transition between model forms in the range 150–175 ms. In the $\geqslant 175$ ms range, AR1 fits were clearly better than the AR2 fits, because for $\beta \approx 0$ accurate parameter estimation in AR2 tends to be swamped by system noise, as described earlier. Below 175 ms, the AR2 form shows lower residuals than AR1 and displays the high frequency peak found in the data, which the AR1 form does not. There was no suggestion of $1/f$ behavior at low frequencies.

These results are consistent with previous experimental work suggesting a qualitative change in cognitive processing for pulse sizes below 250–300 ms (Wing et al. 1987; Peters 1989). They are also understandable on the grounds that the fastest paces of $\leqslant 150$ ms are less than or of the order of typical estimates for a minimum auditory reaction time, usually cited as about 140 ms (Luce 1986), as well as typical estimates for decision-based error correction times (110–200 ms, depending on experimental contest: Vince 1948; Gibbs 1965). It might be considered that the hand alternation necessary to produce these fastest speeds is at least partially responsible for sequential within-hand (i.e., lag 2) effects. However, this is quite unlikely, as extensive experimental trials with alternating hands for pulse sizes above 200 ms for both subjects showed no convincing evidence for the AR2 model. This work will be reported elsewhere (Pressing, manuscript submitted).

This result, that expertise enables richer mental modeling (via an increase in autoregression order), has parallels in other motor tasks. Notably, in tracking and steering tasks, one primary result of practice is a shift from the use of purely low-order spatial information like position to the inclusion of higher-order control information, such as velocity and acceleration (Kelley 1968; McCormick 1970). The use of higher-order information is found to be context-dependent, as here, according to the progression-regression hypothesis (McCormick 1970; Jagacinski and Hah 1988).

As seen in Table 1, AR1 $\alpha$ uniformly increases with pulse length, that is, greater processing time between taps allows more rapid and reliable error correction. $\alpha$ can be considered a measure of information accumulation in working memory, a process which presumably entails transmission of peripheral sensory information to the central nervous system and its conversion to a representation useful for controlling action. Except for the slowest speed, $\alpha$ values for the expert performer are substantially larger at each pulse length than for the non-expert, corresponding to more effective error compensation due to expertise. Motor variance is approximately independent of pulse size and subject, as found by other workers (Wing and Kristofferson 1973b; Wing 1980), whereas clock variance increases with pulse size and is much larger for the novice than the expert, as expected.

Figure 3 shows a plot of error parameter values. For the AR2 cases, with $\alpha$ labelled $\alpha_2$, consistently $\alpha_2 > 0$, $\beta < 0$, and $\alpha_2 + \beta \approx 0.10 - 0.20$. From the form of (4), it can be seen that $\alpha_2 + \beta$ can be interpreted as a total or net error correction parameter. Since values of both AR1 $\alpha$ and AR2 $(\alpha_2 + \beta)$ plateau to this same range for both subjects for shorter pulses, this range may represent a general error correction threshold found as processing demands increase. The fact that $\beta$ is negative we interpret to mean that $\beta$ corrects for errors of correction in the first-order process that are increasingly likely at fast speeds.

Equations (16) or (20) can be used to dissect the spectrum into 'cognitive' and 'motor' components via the corresponding two terms in the numerator. The motor component makes its largest contribution for higher frequencies, and the cognitive component is dominant for lower frequencies (Fig. 4), producing a characteristic cognitive noise plateau.

The work here achieves a good match between data and theory without the postulation of any $1/f$ noise sources, and in accord with this, no low-frequency peak appears in our asynchrony spectral data. However, Gilden et al. (1995) found a $1/f$-type low-frequency peak in the power spectrum of interonset intervals for non-synchronized tapping experiments, and interpreted this as being due to the cognitive noise component of the Wing-Kristofferson model. Although both the experimental conditions (synchronized/non-synchronized) and the variables used as a basis for power spectrum calculation (asynchronies/interonset times) are different in these two cases, these results actually require reconciliation, since cognitive noise of the Wing-Kristofferson type appears in explanations of both results, and our data would be expected to show some low-frequency evidence of $1/f$ noise in their interonset spectrum if cognitive clock noise were the source of it, and they do not.

We believe the reason is as follows. The Wing-Kristofferson form was generated on the basis of the empirical finding that for a stationary time series of spontaneous tapping, the interonset autocovariance at lag 1 is negative, whereas for higher lags it is effectively

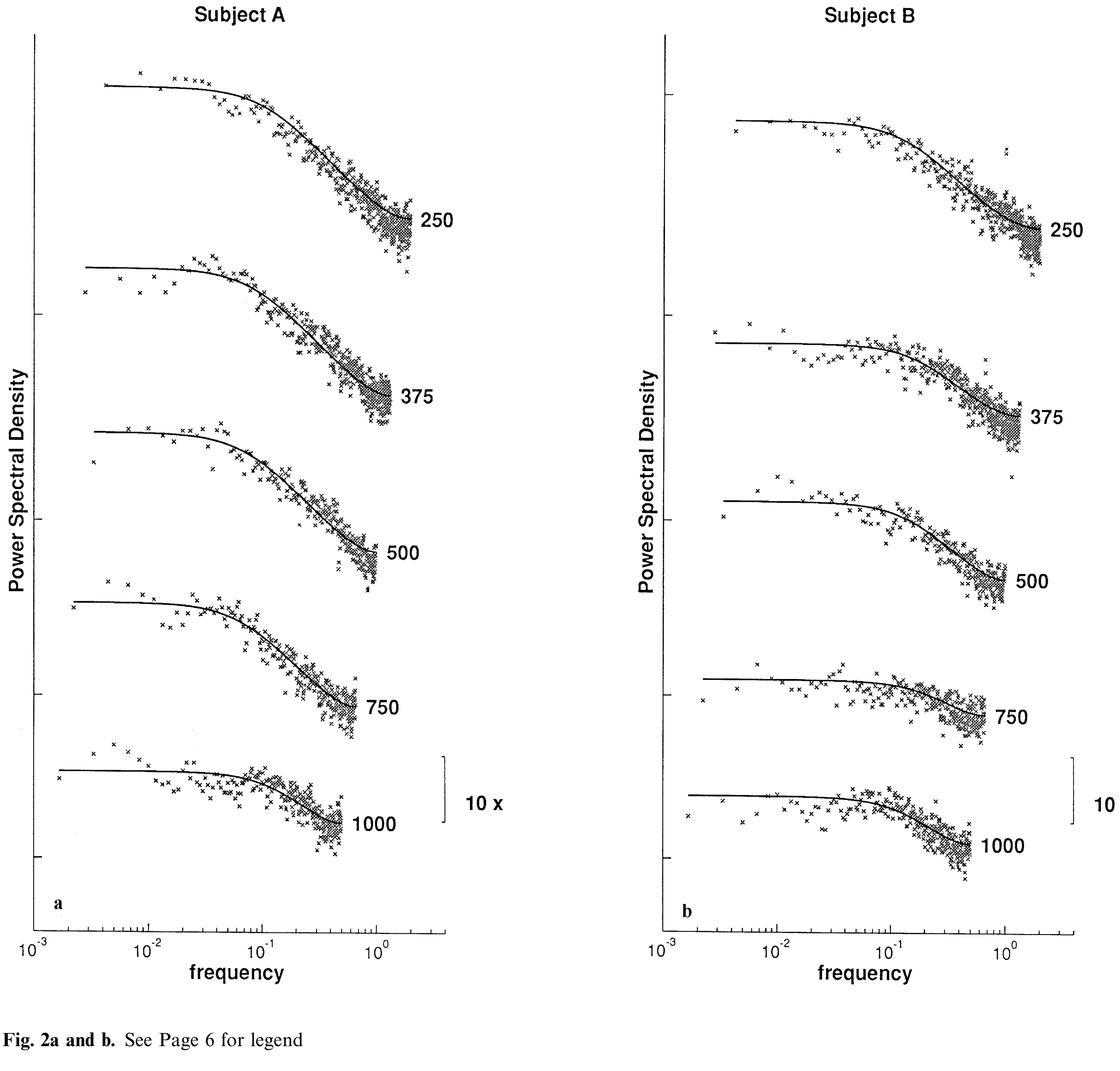


**Fig. 2a and b.** See Page 6 for legend

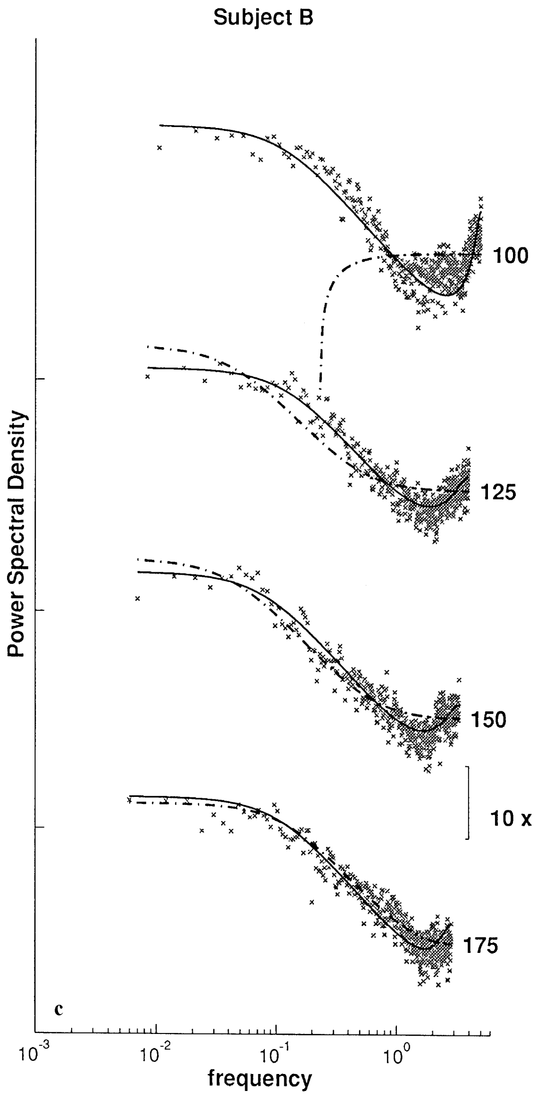


**Fig. 2a–c.** Plots of spectral density vs frequency ($\omega/2\pi$) of asynchronies for synchronized tapping experiments for two right-handed subjects. Subject A is not a musician. Subject B is a highly trained professional musician, who was able to perform runs at faster speeds. Pulse size $P$ is indicated for each case in milliseconds. Data points are an average of 8 runs in each case, totalling ca 5000–7500 taps at each pulse size. *Curves* represent model predictions. Successive cases have been shifted up by scale factors for clarity of display. For **a**, **b**, the curves show a first-order autoregressive (AR1) spectrum, as given in (20), using parameters directly calculated from the time series data. For $P \leqslant 175$ ms (**c**), both AR1 and AR2 predictions are shown: AR2 = —, AR1 = –·–, with the second-order autoregressive (AR2) spectrum calculated from (16). The superior fit of the AR2 model for fast pulse speeds is particularly evident for high frequencies

zero. The covariance properties of 'clock' and 'motor' noise sources, namely those of complete stochastic independence (white noise), are thus essential to the production of the correct time domain behaviour.

Given these conditions, the spectrum of the Wing-Kristofferson process follows ineluctably from the Wiener-Khintchine theorem as (24), or equivalently in the Appendix (29), which shows no $1/f$ behavior. $1/f$ behavior can only occur if higher autocovariance functions are non-zero, and they are not. Hence, the source of $1/f$ noise cannot be the Wing-Kristofferson cognitive clock process, contradicting Gilden et al.'s (1995) interpretation.

There seems to be a simple way out of this problem. We note that the Wing-Kristofferson form does not hold for a non-stationary series, although the issue has not always been emphasized, because tapping runs have typically been short (Wing and Kristofferson, 1973a, b). Likewise, the Wiener-Khintchine theorem does not apply in the absence of second-order stationarity. We believe therefore that the source of Gilden et al.'s $1/f$ noise for low frequencies may be series non-stationarity, due to medium- to long-term fluctuations in speed of performance. Significantly, Gilden et al. (1995) did not perform data detrending, and so their results are in fact not expected to be stationary, supporting this interpretation. Further support comes from the fact that Gilden et al.'s (1995) $1/f$ form fit well primarily for longer duration pulse intervals, which presumably generated longer runs. Given this, we suggest that the human source for such fluctuations may rest with sustained attention processes, one form of which is measured in classical vigilance tasks. Sustained attention has been identified as one of four central factors making up attentional processes (Cohen 1992).

We examined these ideas in a preliminary replication effort. About 2000 additional trials were performed with the expert performer at each of two pulse sizes: 250 ms and 750 ms, using a non-synchronous ('spontaneous') protocol. In this work the performer's task was to keep a steady beat at a pretrained speed without a reference tone. Other experimental conditions were identical to those described earlier.

With no detrending, the interonset series were markedly non-stationary, and we found a clear low-frequency rise in the interonset power spectra at the longer pulse size, as in Gilden et al. (1995), but only a very weak rise at the faster speed. The low-frequency rises were markedly reduced by (cubic to quintic polynomial) detrending to improve stationarity, supporting the idea that the genesis of the low-frequency rise is medium- to long-term fluctuations in speed. Such fluctuations do not occur in synchronized tapping because the presence of a recurring reference source damps them out. The fact that Gilden et al. (1995) found that fluctuations of this kind also apply to distance estimations but not to an unlinked series of reaction time experiments supports the view that the source of $1/f$ spectral noise is not specifically linked to a cognitive clock process.

Additional features characteristic of expertise also appeared in these non-synchronized runs. There were small sharp peaks at the fast speed that corresponded to subharmonics of the basic pulse. These were equivalently visible as peaks in the correlogram at intervals of four taps $\gamma(4)$, $\gamma(8)$, $\gamma(12)$. . . and corresponded to cognitive groupings based on 4/4 musical meter applied to the temporal patterns, resulting from our expert subject's musical training. These peaks were not apparent at the slower speed.

**Table 1.** Parameters and relative fits of the models to the data

| Subject | Pulse size (ms) | Model type and parameter values | Log-log slope | $\sigma_M^2$ (ms$^2$) | $\sigma_C^2$ (ms$^2$) | Fraction of cases with: Better AR fit than '$1/f$' fit | Better AR1 fit than AR2 fit |
|---|---|---|---|---|---|---|---|
| A | 1000[a] | AR1: $\alpha = 0.563 \pm 0.058$ | $-0.584 \pm 0.048$ | $8.42 \pm 79.9$ | $943.17 \pm 158.1$ | 8/8** | 7/8* |
| | 750[b] | AR1: $\alpha = 0.274 \pm 0.025$ | $-1.010 \pm 0.038$ | $8.70 \pm 31.5$ | $589.73 \pm 82.1$ | 8/8** | 7/8* |
| | 500 | AR1: $\alpha = 0.209 \pm 0.026$ | $-1.253 \pm 0.023$ | $15.98 \pm 41.0$ | $466.09 \pm 74.5$ | 8/8** | 8/8** |
| | 375[b] | AR1: $\alpha = 0.180 \pm 0.013$ | $-1.277 \pm 0.031$ | $15.56 \pm 17.0$ | $228.50 \pm 22.6$ | 8/8** | 7/8* |
| | 250 | AR1: $\alpha = 0.155 \pm 0.027$ | $-1.273 \pm 0.052$ | $9.32 \pm 3.15$ | $69.16 \pm 9.16$ | 8/8** | 8/8** |
| B | 1000 | AR1: $\alpha = 0.559 \pm 0.049$ | $-0.529 \pm 0.048$ | $27.75 \pm 23.90$ | $384.8 \pm 45.5$ | 8/8** | 8/8** |
| | 750 | AR1: $\alpha = 0.654 \pm 0.043$ | $-0.383 \pm 0.060$ | $13.96 \pm 13.49$ | $231.1 \pm 23.8$ | 8/8** | 8/8** |
| | 500 | AR1: $\alpha = 0.379 \pm 0.069$ | $-0.862 \pm 0.049$ | $11.41 \pm 7.80$ | $125.1 \pm 15.2$ | 8/8** | 7/8* |
| | 375 | AR1: $\alpha = 0.353 \pm 0.036$ | $-0.841 \pm 0.033$ | $13.75 \pm 6.95$ | $64.70 \pm 7.94$ | 8/8** | 8/8** |
| | 250 | AR1: $\alpha = 0.194 \pm 0.015$ | $-1.090 \pm 0.052$ | $10.76 \pm 2.82$ | $35.47 \pm 2.98$ | 8/8** | 8/8** |
| | 175 | AR1: $\alpha = 0.144 \pm 0.025$ | $-1.168 \pm 0.074$ | $13.40 \pm 6.30$ | $27.04 \pm 4.29$ | 8/8** | 8/8** |
| | | AR2: $\alpha_2 = 0.644 \pm 0.159$, $\beta = -0.477 \pm 0.177$ | | $-6.84 \pm 12.07$ | $44.12 \pm 30.96$ | 7/8* | |
| | 150 | AR2: $\alpha_2 = 0.497 \pm 0.076$, $\beta = -0.393 \pm 0.070$ | $-0.875 \pm 0.027$ | $7.3 \pm 4.3$ | $36.3 \pm 4.5$ | 7/8* | 2/8 |
| | 125 | AR2: $\alpha_2 = 0.529 \pm 0.051$, $\beta = -0.397 \pm 0.047$ | $-0.770 \pm 0.017$ | $12.43 \pm 3.05$ | $40.76 \pm 4.55$ | 7/8* | 2/8 |
| | 100 | AR2: $\alpha_2 = 0.845 \pm 0.065$, $\beta = -0.721 \pm 0.063$ | $-0.792 \pm 0.074$ | $1.456 \pm 6.704$ | $49.9 \pm 3.80$ | 8/8** | 1/8* |

$^{**}P < 0.005$ (one-tailed), $^{*}P < 0.05$ (one-tailed)

[a] Additional constraint in $\alpha$ estimation: $0 < \alpha < 1$

[b] Seven runs used in parameter calculation. One excluded on the basis of parameter deviations $> 4$ standard errors

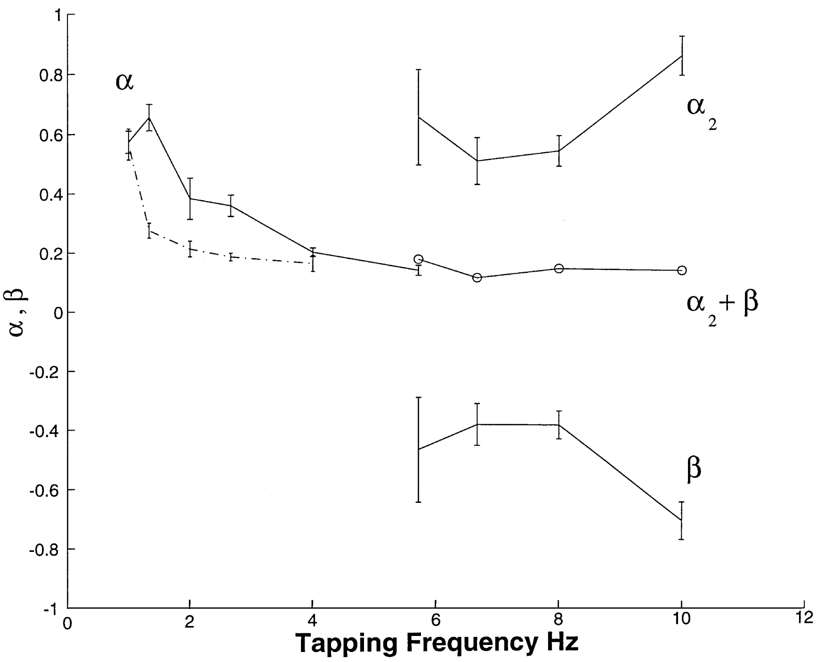


**Fig. 3.** Error correction parameters for all runs. AR1 parameter $\alpha$ increases with pulse size. Net error correction at faster speeds appears to be limited by information-processing constraints to a parameter range of ca 0.10–0.20, as measured by AR1 $\alpha$ and AR2 ($\alpha_2 + \beta$)

## 6 Conclusions

The model here is a simple one, and it does not address possible effects due to separate error correction processes for phase and period (Mates 1994a, b) or non-linear effects (e.g. order discrimination thresholds – typically about 20 ms, or higher polynomial terms (Hirsch and Sherrick 1961; Mates 1994b). However, its exact solvability allows a closed form for the spectral distribution function, and in view of the excellent fit of data and model in the absence of free parameters, these possible effects are apparently small or mutually compensatory. We suggest that, in particular, the impact of the order discrimination threshold is limited by the frequently

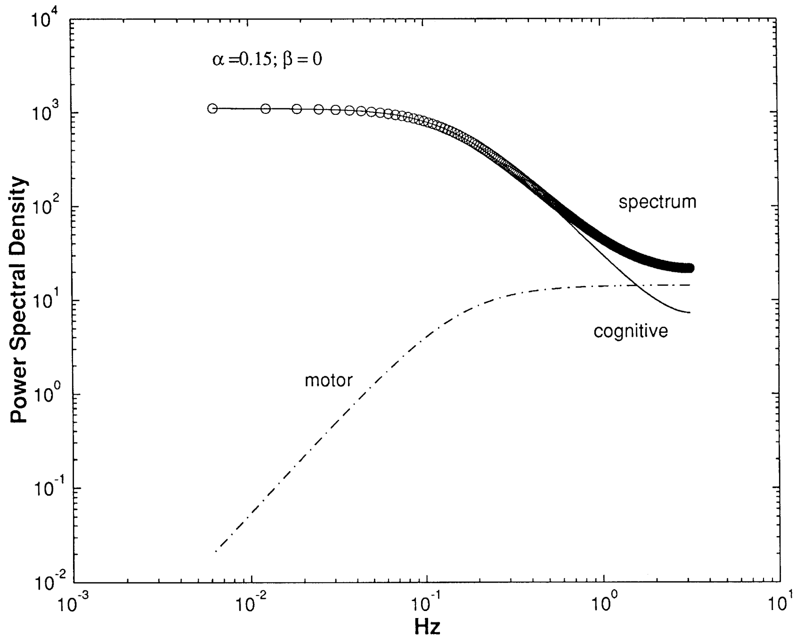


**Fig. 4.** Decomposition of the spectrum into its cognitive and motor components

observed tendency, found here as well, for tapping to consistently anticipate the stimulus (Fraisse 1982; Mates 1994a).

The model's success supports the idea that linear autoregressive error correction processes are dominant in synchronized tapping and exhibit characteristic spectral shapes. Expertise is distinguished in a consistent fashion, by higher-than-non-expert error correction parameters and lower values of clock variance. The model presented here has also been extended to many other types of rhythmic patterns, and linear error correction processes have been found also to offer a good description of those experimental results (Pressing, manuscript submitted). The reverse-sigmoidal spectral shapes found here are also reflected in those cases. In other work, we have shown that linear error correction processes can be treated as special cases of a more general non-linear formulation (Pressing, manuscript submitted).

Overall, the spectra of cognition and skilled human performance show at least four types of behaviour. First, when memory-less production with independent trials is involved, the data display no autocorrelation, and the spectrum is white (flat), as for example with iterated reaction time trials (Gilden et al. 1995). Second, when error correction variables based on successive synchronization to an external time template are treated, the error spectrum is reverse sigmoidal, normally arising from a first-order autoregressive process, unless task demands of coordination and speed require greater accuracy, when a second-order process may be used. Third, estimation based on an internal reference and a sequential assessment process can produce $1/f$-like behaviour at low frequencies, due perhaps to medium- to long-term time fluctuations in attention. Finally, musical training can produce specific subharmonic peaks due to learned principles of cognitive grouping (phrasing).

*Acknowledgement.* This research was supported in part by an ARC grant to the first author.

## Appendix A

The method of direct substitution described in the paper correctly produces (16). However, a simpler way to compute the spectrum corresponding to the autocorrelation functions of (10) is to use both the Wiener-Khintchine theorem and the linear filtering theorem as follows. We can use (4) in the main paper to define a series

$$E_n = (C_n - P) + (M_{n+1} - M_n) = I_n^* - P$$
$$= A_{n+1} - (1-\alpha)A_n + \beta A_{n-1} \tag{26}$$

for the second-order autoregressive case, where $I_n^*$, as given by (6), is just the intertap in the absence of error correction, that is, the basic Wing-Kristofferson model. Since $E_n$ only differs from $I_n^*$ by an additive constant, their power spectra are identical. What is $f_{I^*}(\omega)$? This is computed as follows. Under the standard assumption of independence of the C's and M's, Wing and Kristofferson (1973a) computed the following values for the autocovariance functions:

$$\begin{aligned} \gamma_{I^*}(0) &= \sigma_C^2 + 2\sigma_M^2 && (j=0) \\ \gamma_{I^*}(1) &= -\sigma_M^2 && (j=1) \\ \gamma_{I^*}(j) &= 0 && (j>1) \end{aligned} \tag{27}$$

From the Wiener-Khintchine theorem,

$$f_{I^*}(\omega) = \gamma_{I^*}(0) + 2\sum_{k=1}^{\infty} \gamma_{I^*}(k)\cos\omega k$$
$$= \gamma_{I^*}(0) + 2\gamma_{I^*}(1)\cos\omega$$
$$= \sigma_c^2 + 2\sigma_M^2(1-\cos\omega) \tag{28}$$

Hence

$$f_E(\omega) = f_{I^*}(\omega) = \sigma_c^2 + 2\sigma_M^2(1-\cos\omega) \tag{29}$$

and since the linear filtering theorem in this case yields [from (26)]

$$f_E(\omega) = |1-(1-\alpha)e^{i\omega} + \beta e^{2i\omega}|^2 f_A(\omega) \tag{30}$$

rearrangement results in

$$f_A(\omega) = \frac{\sigma_c^2 + 2\sigma_M^2(1-\cos\omega)}{1+(1-\alpha)^2+\beta^2-2(1-\alpha)(1+\beta)\cos\omega+2\beta\cos 2\omega} \tag{31}$$

as was to be shown (16).

## References


Cohen R (1992) The neuropsychology of attention. Plenum Press, New York

Fraisse P (1982) Rhythm and tempo. In: Deutsch D (ed) Psychology of music. Academic Press, New York, pp 149–180

Gibbs CB (1965) Probability learning in step-input tracking. Br J Psychol 56:233–242

Gilden DL, Thornton T, Mallon MW (1995) $1/f$ noise in human cognition? Science 267:1837–1839

Gottman J (1981) Time series analysis. Cambridge University Press, Cambridge

Hary D, Moore GP (1987) On the performance and stability of human metronome-synchronization strategies. Br J Math Stat Psychol 40:109–124

Hirsch IJ, Sherrick CE (1961) Perceived order in different sense modalities J Exp Psychol 26:423–432

Jagacinski RJ, Hah S (1988) Progression-regression effects in tracking repeated experiments. J Exp Psychol Hum Percep Perform 14:77–88

Jagacinksi RJ, Marshburn E, Klapp ST, Jones MR (1988) Tests of parallel versus integrated structure in polyrhythmic tapping. J Motor Behav 20:416–442

Kelley CR (1968) Manual and automatic control. John Wiley, New York

Luce RD (1986) Response times. Oxford University Press, Oxford

Mates J (1994a) A model of synchronization of motor acts to a stimulus sequence. I. Biol Cybern 70:463–473

Mates J (1994b) A model of synchronization of motor acts to a stimulus sequence. II. Biol Cybern 70:475–484

McCormick EJ (1970) Human factors engineering. McGraw-Hill, New York

Miller SL, Miller WM, McWhorter PJ (1993) External dynamics: a unifying physical explanation of fractals, $1/f$ noise, and activated process. J Appl Phys 73:2617

Nettheim NN (1992) On the spectral analysis of melody. Interface 21:135–148

Peters M (1989) The relationship between variability of intertap intervals and interval duration. Psychol Res 51:38–42

Pressing J (1994) Novelty progress and research method in computer music composition. Proc 1994 Int Comp Music Conf, ICMA, San Francisco, pp 27–30

Pressing J, Summers J, Magill J (1996) Cognitive multiplicity in polyrhythmic performance. J Exp Psychol Hum Percept Perform 22:1127–1148

Schmidt RC, Treffner PJ, Shaw BK, Turvey MT (1992) Dynamical aspects of learning an interlimb rhythmic movement pattern. J Motor Behav 24:67–83

Schulze H-H (1992) The error correction model for the tracking of a random metronome: statistical properties and an empirical test. In: Macar F, Pouthas V, Friedman WJ (eds) Time, action, and cognition. Kluwer, Dordrecht, pp 275–286

Turvey MT, Schmidt RC, Rosenblum LD (1989) 'Clock' and 'motor' components in absolute coordination of rhythmic movements. Neuroscience 33:1–10

Vince MA (1948) Corrective movements in a pursuit task. Q J Exp Psychol 1:85–103

Vorberg D, Hambuch R (1984) Timing of two-handed rhythmic performance. In: Gibbon J, Allan L (eds) Timing and time perception. New York Academy of Sciences, New York, pp 390–406

Vorberg D, Wing A (1996) Modeling variability and dependence in timing. In: Heuer H, Keele SW (eds) Handbook of perception and action, Vol 2. Motor skills. Academic Press, London, pp 181–262

Voss RF (1988) Fractals in nature: from characterization to stimulation. In: Peitgen H-O, Saupe D (eds) The science of fractal images. Springer-Verlag, Berlin Heidelberg, New York, pp 21–70

Voss RF, Clarke J (1975) $1/f$ noise in music and speech. Nature 258:317

Voss RF, Clarke J (1978) $1/f$ noise in music: music from $1/f$ noise. J Acoust Soc Am 63:258–263

Wing AM (1980) The long and short of timing in response sequences. In: Stelmach GE, Réquin J (eds) Tutorials in motor behavior. North Holland, Amsterdam

Wing AM, Kristofferson AB (1973a) The timing of interresponse intervals. Percept Psychophys 13:455–460

Wing AM, Kristofferson AB (1973b) Response delays and the timing of discrete motor responses. Percept Psychophys 14:5–12

Wing AM, Church RM, Gentner DR (1987) Variability in the timing of responses during repetitive tapping with alternate hands. Psychol Res 51:28–37