Psychological Review

VOLUME 89 NUMBER 2 MARCH 1982

Rhythm and Timing in Skill

L. H. Shaffer
University of Exeter, Exeter, England

The concept of rhythm points to diverse phenomena in skilled performance. A theoretical frame sufficiently general to deal with these is presented, together with data from a variety of skills. It is argued that a motor system, responsible for movement production, can produce movements that realize given time-scales and hence can act as a timekeeper. In contexts requiring temporal coordination, a more abstract timekeeper, or clock, can provide temporal references for a motor system or several motor subsystems; this enables complex timing such as is found in musical performance. Skilled movement is teleological in the sense that it aims at targets in space and time. Its timing is based on an internal schedule of motor events—the targets of movement—not movement onsets: The motor system arranges the latter on considerations of movement economy or expressiveness. Tapping, handwriting, typing, playing music, and speech are examined as instances of skills having a near-miss periodicity.

"Skilled performance is rhythmical." Like many apparent truisms this one is weakened by ambiguity, in this case in the meaning of *rhythm*. If we ask what is common between the rhythms of speech, playing tennis, and playing music, we find that it refers to different properties of time series, relevant in varying degrees to the different skills. Among these are properties of temporal pattern, periodicity, stress, expression, and quality of movement. The purpose here is to develop a theoretical frame for rhythm that can encompass this variety of phenomena.

We can illustrate the different senses of rhythm by taking music as a paradigm of serial-structured forms. Excluding some of folk music and a handful of modern free-form compositions, the timing in a piece of music is based on the fixed interval of a recurrent beat. This is specified in written music; the beat is easy to identify in a competent performance; and a listener can dance or clap in time with it. The beat usually supports two other rhythmic structures: one is meter, a recurrent group of beats in which certain beats are more salient, or stronger, than others; the other is a temporal pattern obtained by joining and subdividing beat intervals to create a series of note and pause durations. Temporal patterns are constructed by producing rhythmic figures and then transforming them in ways that simplify, ornament, expand, contract, or invert the original version, embedding the results in a variety of contexts. Thus the generative grammar of these patterns is typically more complex than those described by Simon (1972) and Greeno and Simon (1974). The term *rhythm* has often been used indifferently to refer to the beat, the meter, or the temporal pattern of music.

The strong beats in the meter provide focal points toward which musically important events tend to gravitate (Lerdahl & Jackendoff, 1977). Hence they tend to receive

---

This research was supported by Grant HR 4671 from the Social Science Research Council. The paper arose from discussions with Carol Fowler, Scott Kelso, Peter MacNeilage, Don Norman, Dave Rumelhart, Dick Schmidt, Saul Sternberg, Ewart Thomas, and many others. It is just possible that none of them agree with any of the ideas presented here.

Requests for reprints should be sent to L. H. Shaffer, Department of Psychology, University of Exeter, Exeter EX4 4QJ, England.

stress in performance so that a listener can interpret the meter and clap more vigorously on the strong beats. The stress rhythms of music, however, seldom remain within the meter, and rhythmic interest is obtained by occasionally displacing stress from a strong to a weak beat (Forte, 1979).

Jackendoff and Lerdahl (Note 1) have examined the possibility of a deep parallel between music and speech by showing that the abstract structures of salience in music and speech produce similar graphs. Their structural analysis of prosodic stress follows the work of Liberman and Prince (1977) and Selkirk (1980). What is of interest here is that the stress rhythms of speech may not normally be associated with a meter: Stress groups may not form isochronous intervals (Lehiste, 1977), and the temporal patterns of speech, summarized by Klatt (1976), may not contain a periodic beat. Speech can, of course, acquire a regular beat in certain contexts; it is possible to count in regular time, recite verse, and perform song-speech as in the music of Schoenberg. Conversely, musicians can depart from regular timing in playing music, and in Romantic music they may play with something approaching the rhythmic freedom of speech. Thus, if there is a difference between music and speech it is normative, and we expect to find that the ability to impose a beat as a superordinate timing constraint is, to some extent, independent of the performance medium.

To a musician an important sense of rhythm is the expressive use of timing in performance. This entails a judicious use of rubato, or tempo variation, to give the beat an elastic quality, and of playing off the beat, leading or lagging it, when it is provided by an accompaniment. A musician who plays squarely on an inflexible beat definitely has not "got rhythm."

One more sense of rhythm, and perhaps the only one that applies to ball games and other episodic activities, is the quality of movement, which in the skilled performer gives the appearance of being unhurried, fluent, and avoiding abrupt accelerations. Thus, a theory of rhythm has to distinguish between the timing of motor events, such as making contact with a ball, striking a note on a keyboard, or producing phonation of a speech sound, and the kinematics of the movements that lead to them. Motor events are the successive goals of movement. They can be represented in models of timing as occurring at discrete moments in time, but a theory of rhythm has to take account of the underlying continuity of movement.

The production of rhythm in general is made possible by the preparation of movements and their coordination (Lashley, 1951; Leonard, 1953; Poulton, 1957; Shaffer, 1981). Without preparation performance is a series of discrete reactions to external events. The development of motor skill can be traced as the progress from reactive movement to movement fluency, coupled with a flexibility in tailoring action to the details of an infinite variety of contingencies.

The properties of fluency and flexibility in skilled performance lead us to a theory of motor programming, whose major components are a performance grammar and a control system. A performance grammar is a collection of generative rules for creating syntactic strings relevant to an infinite variety of actions in a certain domain, together with rules for translating symbols in these strings first into the commands for movements and then into the movements themselves. Liberman (1970) described such a grammar in showing that the linguist's grammar must be supplemented by rules for translating phonological strings into speech. A motor program is a set of grammatical representations of intended action constructed, by a control system, as a hierarchy of abstractions, terminating in motor output. In serial performance these representations undergo continual renewal as the action unfolds.

Grammars are usually considered as systems for generating strings of symbols or categorial elements. A performance grammar contains a level of rules that map symbols into continuous variables, such as force, rate, and direction. If a motor command is a symbol specifying the targets of a movement (MacNeilage, 1970), then the rules at the terminal level are computational algorithms determining the trajectories of limbs and articulators.

Movement Timing

A suitable starting point for a theory of motor rhythm is the observation that skilled movement can be continuous over large time spans. Although it is convenient to talk of a succession of movements, it is more accurate to describe performance as a sequence of motor events produced by movements that evolve fluently from one another. Also, movements in different limbs and articulators can overlap in time in the production of successive motor events. If the motor events are fixed in time as well as space by requirements external to the movements themselves, as in playing music, for instance, then movements have to be structured around these fixed points.

The proposition that a movement has a temporal goal as its reference point is fundamental to the present thesis. It is markedly different in emphasis from an assumption that movements are triggered at the times of motor events (Kozhevnikov & Chistovich, Note 2; Shaffer, 1976, 1981; Wing & Kristofferson, 1973). The proposition carries an important corollary: The motor system responsible for constructing movements can translate a given time interval, or equivalent information, such as that derived from visual stimulus flow (Lee, 1980), into a movement trajectory having that duration. Hence, the motor system can act as a timekeeper, realizing a time scale in a movement trajectory. The data from some recent studies are relevant to these proposals.

A minimum condition for a motor system to act as a timekeeper is that the time course of a movement trajectory be reproducible. Cooke (1980) has shown, for step tracking between two fixed points, that the phase-plane plots of individual movements, relating limb velocity to position, superimpose on one another with very little variance. They also have the smooth unimodal shape of a ballistic movement. He also obtained a linear relationship between peak velocity and the amplitude of simple reaching movements, when there was no specified distance or time for the movement. Movement time was not invariant; but the result raises the possibility that timing can be a determining parameter of movement; and this can be examined in phase-plane plots of step tracking paced by a metronome, with duration and amplitude as controlled variables.

This was indirectly studied by Kelso, Southard, and Goodman (1979) by looking at concurrent movements made with both hands to targets differing in size and distance from the resting position. In making single movements the times to initiate and complete a movement varied with target size or movement amplitude. In the combined condition, with the instruction to complete both movements as quickly as possible, where one had a large close target and the other a small distant target, the hands tended to start and arrive together, having to produce different velocity trajectories to do so. Also, the velocity curves were smooth and reached their peaks at the same time. This indicates that timing parameters were involved in the movements and that it was easier to prepare the combined movements if the parameters took the same values for each limb. In this case the limb making a large movement to a small target preempted the timing for both limbs.

Timing as a goal of performance can become more important in serial than in single movements. Schmidt (1980) reports a study in which a lever had to be moved to a specified endpoint in a given amount of time. On occasional trials and without warning the mass of the lever was changed. For a simple movement to target, the endpoint of movement was unaffected by the change but its timing altered; in making a serial movement, to the target and back to the start position, movement time was unaffected, but the endpoint of movement altered. The priority of attaining the spatial and temporal components of the target thus differed between the conditions.

A motor event can usually be given a multivariate description, but not all the variables may have been set in programming the movement. Playing a note on the piano in the context of a piece of music entails a movement to a certain location on the keyboard, timing the movement of striking the key, and controlling both the intensity and manner of strike, the last affecting the quality of legato or staccato by sustaining pressure on the key or rebounding from it. All these variables are musically important and so should be programmed. In a less stringent context, some variables may be allowed to inherit values from earlier movements or take default values to simplify the computation of movement trajectory. The evidence from Schmidt's experiment shows that both location and timing variables were programmed, but only one or the other, depending on the movement, was protected from input distortion.

We can suppose that the temporal rhythm of a performance, if it is reproducible, is the realization of a schedule in a motor program. The variable specified in the schedule need not itself be time; it can be any variable that has a definite timing concomitant within a mode of computation of the motor system. Stress level or movement amplitude can serve if they do not already have to vary orthogonally with timing.

Given a schedule, the motor system can produce a movement to the next temporal target, taking a previous target as a reference point for the movement trajectory. In repetitive tapping with one finger, the previous tap provides the reference point for the next one. When tapping with two fingers in alternation, the movement for the next tap can start on the penultimate tap, and there is a choice of taking timing reference from that moment or from the moment of the previous tap on the alternate finger. In typing (Gentner, Grudin, & Conway, Note 3) and piano playing, using all 10 fingers in irregular sequence, a finger can move towards its next key at any time following its last use. However the moment of onset of movement to a spatial target may not be the reference point from which it is timed to reach the target. The reference point can occur fairly late in the trajectory if there is extensive coarticulation of movements, in which case it should be possible to detect a change of acceleration in the trajectory as a movement changes from "ballpark" timing to reference timing.

Subsequences of movements may become organized as a group if they occur frequently or have a short time scale. We can thus think of a compound movement-trajectory, which includes the movements for several motor events, launched from a common reference point and articulated on an internal schedule. In this case the motor program can specify the onset of the first event and leave the motor system to compute the internal schedule.

In summary, we have outlined a theory of motor skill in which movements have the teleological property of anticipating targets in space and time and can be fluently coordinated to realize a given schedule. It is assumed that the motor system producing the movements has the computational resources to act as a timekeeper by controlling the time scale of movement trajectories and can construct trajectories producing more than one event.

It has been supposed that motor commands in an output buffer are serially activated on an internal schedule (Kozhevnikov & Chistovich, Note 2; MacKay, 1971; Shaffer, 1976). This is reasonable if motor timing is related to movement onsets but not if it is related to movement targets. To enable anticipatory movement a representation of event sequence would have to be transformed to one ordered in terms of movement onset, with the undesirable consequence that effective timing control would pass to the inessential variable (movement onset) rather than to the essential variable (event onset).

With the new proposal the motor system has to be given a perspective on the output string of the motor program to enable it to anticipate movement requirements. It can achieve this by scanning the string in advance of production to obtain the relevant information. Such a scan can serve several purposes: (a) It allows the motor system to anticipate movement targets and to add auxiliary information relevant to the background conditions for movement, such as the regulation of breathing in speech (Ladefoged, 1967) or the postural adjustment preparatory to raising a limb (Belenkii, Gurfinkel, & Pal'tsev, 1967); (b) it allows the motor system to group motor commands to construct compound movement trajectories; and (c) it allows the motor system to interpret expressive information ranging over a sequence of movements.

Apart from matters of timing, the scan

concept can account for certain properties of speech errors. The displaced phonemes and phonetic features in spoonerisms tend to move to similar phonetic contexts (Fromkin, 1971), or such errors are triggered by the conflicting rhythms of phoneme or feature alternation in tongue twisters (Clark & Clark, 1977; Schourup, 1973). A prior scan of the output string, picking up context information, can confuse this with order information and create an ordering ambiguity in the translation from commands to movement (cf. Shattuck-Hufnagel, 1979).

The motivations for the assumptions made here should become clearer in the following sections, when we discuss particular skills.

Clock Reference

The function of timekeeping may be distributed in the neuromotor system rather than be the property of a single subsystem. A time interval may be realized in a movement trajectory or constructed at a more abstract level. A timekeeper at the latter level can be called a clock: It can serve to provide temporal reference points for movements and hence coordinate movements in independent motor subsystems. The case for a clock can be made from the performance of tapping in time with a metronome, either tapping in synchrony with it or tapping the continuation of its series when it stops (Michon, Note 4; Stevens, 1886; Wing & Kristofferson, 1973).

Synchronic tapping is achieved if a movement can be found having the same period as the metronome and then, starting in synchrony with the metronome, iterating on this movement. To obtain synchrony in the first place, the performer must construct an internal model of the beat interval and compute a movement that leads to a tap on the next beat. Here are the ingredients of distributed timing: an internal model that by recycling can act as a clock and a motor system that can construct a movement having a given period.

Let us examine the possibility of performing synchronic tapping without the use of a clock. Like Michon we shall for this purpose take it for granted that performers can start in time with the metronome. They produce a series of tapping movements, each aiming at the duration of the metronome period, but having a random error of production. In effect they act as a stochastic metronome in parallel with the external one. If tapping cycles are serially indexed, then over a series of taps the asynchrony between metronome beat and its indexed tap will describe a random walk about zero, its starting point. But asynchrony is additive and so its variance increases linearly with the number of cycles; hence in a prolonged performance the asynchrony can become indefinitely large, though its expected value is zero.

As a way of improving synchrony, the performer may adopt one of two tactics. The first uses feedback information of the previous asynchrony to regulate the next movement interval. In the nth tapping cycle, a candidate interval, I'_n, drawn from a fixed distribution, is modified by adding a proportion, ϕ, of the asynchrony on the previous tap, a_{n-1}. The actual tap interval is then

$$I_n = I'_n + \phi a_{n-1}, \quad -1 < \phi < 0.$$

It can be shown that as the feedback coefficient, ϕ, approaches -1 the variance of the tapping interval increases. Conversely, the variance of asynchrony decreases to an asymptote, the variance of a single tap interval. Thus, synchrony error can be traded against variability of the tapping interval.

Alternatively the performer may time each tap from the previous metronome beat, hence eliminating the possibility of a cumulative drift out of synchrony. Then, if the metronome period is C and a movement interval referred to a beat is I', the interval between taps n and $n + 1$ is

$$I_n = C + I'_n - I'_{n-1}.$$

If the successive I' are independent variables, the variance of I is twice the variance of I'; also the serial covariance at a lag of one (i.e., between adjacent intervals) is minus the variance of I', whereas covariances at greater lags are zero (Wing & Kristofferson, 1973).

These two tactics are equivalent if $\phi = -1$ in the first tactic, because this degree of feedback effectively cancels the previous asynchrony. Michon (Note 4) commented on the equivalence but failed to recognize the distinct tactics. They lead to the same outcome but they entail very different computations, because with one tactic there is no need to compute and then subtract the previous synchrony error. Note that these tactics depend on an intimate coupling between metronome beats and tapping movements, such that information derived from a sensory input can be used to affect the time course of a movement already in progress. Thus it seems that they put performance in the domain of reactive rather than prepared movement and so deny the possibility of movement fluency.

The situation changes if the performer can utilize an internal clock, superordinate to the motor system. McGill (1962) developed a clock model to characterize the time series of intervals that fluctuate randomly but have an underlying periodicity, free of temporal drift. He assumed a determinate clock generating pulses with a fixed period, C, each pulse triggering a response with a random, independent delay, d. The interval, I, between successive responses is then a random variable whose nth value is

$$I_n = C + d_n - d_{n-1}.$$

This model gives a good account of the time characteristics of the pulse train in a nerve fiber when it fires. However, the assumption of a determinate clock is too strong for psychological theories. Accordingly, Wing and Kristofferson (1973), in a description of repetitive tapping, weakened the assumption to one of a stochastic clock, but retained the form of the model, which led them to represent the tapping movement as a response triggered with random delay by a clock pulse. This model has been successful in describing properties of the time series of motor events in continuation tapping, that is, tapping the continuation of a periodic metronome series, and it has led to interesting extensions of the task and the model (Vorberg & Hambuch, 1978; Wing, 1977, 1980b).

If we wish to retain a constructive role of timing for the motor system, this can be achieved in a variant of McGill's model that gives it the form of the second tactic for synchronic tapping, but substitutes an internal stochastic clock for the metronome. Because C is now a random variable, the earlier expression becomes

$$I_n = C_n + I'_n - I'_{n-1}.$$

Assuming independence between the construct variables, the variance–covariance properties of I can be simply described in terms of the variances of C and I':

$$\mathrm{Cov}_k I = \begin{cases} \mathrm{Var}\, C + 2\mathrm{Var}\, I', & k = 0 \\ -\mathrm{Var}\, I', & k = 1 \\ 0, & k > 1, \end{cases}$$

where k is the serial lag between covariate intervals so that covariance at a lag of zero is just the variance. This model gives the same description of timing among motor events as the Wing–Kristofferson (1973) model, but it interprets the motor component differently.

Before evaluating these proposals we should reflect that the only point of using a clock to control movement timing is that it can minimize temporal drift between separate streams of events that have to be coordinated. The clock serves this purpose if and only if it can generate time intervals with a lower order of variance than the motor system can.

The advantage of having a clock can be seen in the context of musicians playing together. In musical groups one musician usually acts as leader, or timekeeper, for the others, and so we can consider the performance of a duet without losing generality. Typically each player plays a part with its own temporal rhythm so that between them they produce notes at different points in time. Their rhythms are based on a common beat and meter, but these are only occasionally made explicit in the performances. Hence, the second player can derive timing information from the leader only on an irregular basis. Even if the beat and meter were made explicit more regularly in the leader's performance, attempting to follow its stochastic vagaries would distort the rhythm of the second player. Thus, ensemble playing appears to depend on the players being able to generate their timekeeping in parallel, each using an internal clock. Of course, even the clocks will drift apart over time, but if the variances of clock intervals are small enough, the drift can remain within the bounds of perceptual acceptability over long stretches of the music, and feedback correction, at the level of bringing one clock into synchrony with the other, can be deferred to suitable boundaries in the music.

The proposed clock model can equally serve as a model of synchronic or continuation tapping. All that differs between these conditions is that in synchronic tapping the performer has the opportunity to detect drift between his clock and the metronome and so can apply occasional correction. The advantage of the proposed model over that of Wing–Kristofferson is that the proposed model opens the theory to the domain of movement kinematics and allows us to consider how movements as well as motor events are serially structured. It also has the advantage of showing how expressive timing can be introduced in the tapping task. If a performer can use a clock to model the metronome beat and can separately time movements in relation to clock pulses, then he has the option to tap ahead of or behind the beat. Musicians frequently move in and out of an accompanying beat in a controlled way, and pianists can go one step further by letting one hand play in and out of time with the other (Shaffer, 1981).

Vorberg and Hambuch (1978) have considered theoretical extensions of clock timing in which a hierarchy of clocks is used to generate temporal structures, such as may occur in musical performance. This will be taken up below; the existing evidence suggests that the timing hierarchy of a clock and a motor system are sufficient to account for musical performance (Shaffer, 1981).

Near-Miss Periodicity

Many serial skills seem to have the property of being nearly periodic, at some level of performance unit, with some skills being nearer than others. In skills that allow rapid performance, perhaps the strongest motive for periodicity is that it creates an iterative processing cycle that can facilitate the control of motor output. It can also help to regulate the accelerations in movement and so reduce the energy dissipated in performance; and at another level it may simplify the computation of successive movement trajectories if a basic timing parameter can be propagated over a sequence of computations. In communicative acts it can provide a boundary cue for segmenting the message signal (Martin, 1972). In music it is a structural property of rhythm, and it allows temporal coordination among players.

These skills can be classified on two factors. The first raises the distinction of whether an underlying periodicity is based on a clock or on a nominal specification given to the motor system by the motor program. In either case it is assumed that a timekeeper, clock or motor system, has a way to construct a required time interval. The second factor considers whether departure from periodicity is merely stochastic, that is, random error, or includes a structured component; if the latter whether this component is introduced at the syntactic or the logistic level of constructing motor output. The logistic level is concerned with movement production, and so the last distinction is related to that of basic rhythms and their expressive, or quasi-expressive, modulations.

Handwriting

Handwriting is a good example of a skill having a near-miss periodicity. Hollerbach (1979) describes a robot arm whose basic production is drawing a sine wave on a writing surface. Handwriting is obtained by modulating the amplitude of the vertical component, producing the oscillation, and independently modulating the left-to-right displacement. With suitable choices of modulation parameters, the robot arm begins to mimic the different handwriting of human individuals. Wing (1978, 1980a) has reported some preliminary results on the time course of human handwriting. In cursive writing the letters *e* and *l* are similar loops differing only in amplitude: Their production times within a word are not the same, but they are relatively more alike than their amplitudes. Thus there appears to be an underlying periodicity modulated by factors affecting letter shape. An analysis of timing across letters showed that there was no consistent negative covariation between adjacent letters, which suggests that the underlying periodicity is nominally specified rather than clock oriented. These studies did not consider a calligraphic component, modulating pressure on the writing surface to affect line thickness, which may in turn affect timing.

Typing

Genest (1956) studied typists at different levels of proficiency and found that the more skilled the typist the more nearly her typing was periodic. In an intensive study of a typist whose speed was above 100 words per minute, Shaffer (1973) observed a regularity in her timing that led to the conjecture that it was clock oriented. But it was also observed that words typed on different occasions preserved temporal profiles over the sequence of inter-key latencies. These properties of timing were separately examined by constructing special texts for typing (Shaffer, 1978).

One kind of text was designed to test a clock model of timing. For the purpose of using serial covariance techniques, it was desirable to make the timing of movement transitions as uniform as possible: accordingly the texts contained only words that entailed strict alternation between hands in the normal conventions of typing (leading to phrases like "authentic divisor," "fiendish paucity," etc.). The typist was faster and her timing even less variable with these texts; also the serial covariance between successive inter-key intervals was negative, as a clock model requires, and the correlation sometimes approached the theoretical limit of −½ for a determinate clock.

The other kind of text was designed to examine whether timing profiles were reducible to effects of movement transition or involved factors beyond the immediate transition. In particular, we were interested in the possible anticipations of right context. Accordingly, texts were constructed from sets of words sharing the first three letters, distributing words from different sets over a text, and in some texts alternating these words with *the* to fix the left context. The latencies, that is, intervals to onset, in typing the first three letters of words were the critical variables. It was found that the pattern of latency in these letters was affected by the continuation of the letter sequence, demonstrating extensive anticipation.

Rumelhart and Norman (Note 5) have run a computer simulation of skilled typing, which attempts to account for anticipatory timing as a secondary consequence of anticipatory movement. In the model every motor command in an output buffer receives a partial activation, the amount decreasing on average with its ordinal remoteness from the current event. A movement to press a key with a certain finger can implicate the hand so that the position of the hand at any moment is the vector resultant of a number of weighted movement tendencies. Not only may the hands move but uncommitted fingers may also move towards their next keys so that there is extensive coarticulation. The latency of typing the next letter depends on the current position of the hand and relevant fingers; because hand position is affected by movement tendencies towards letters to be typed, these can influence the latency of the next keypress.

This model has many interesting properties, but a counterexample from Shaffer (1978) shows that it gives an inadequate account of latency pattern. The following contrast shows the average latencies, in milliseconds, of typing the sequences *whig* and *whim*.

whig	109	121	153	103
whim	110	155	95	164

The preceding word in each case was *the*, and so we see the anticipatory effect of *g* or *m* on the sequence *whi*. On the keyboard *w* and *g* are on the left hand, *h*, *i*, and *m* are on the middle, top, and bottom rows, respectively, of the right hand. In *whig* hand position should make *i* slightly slower than *h*; in *whim* the opposing pulls of *i* and *m* should make *h* slightly faster than in *whig*; the downward pull of *m* should make *i* much slower than in *whig*; and the remaining transition to *m* should make this very slow. The relative latencies of *h* and *i* between the two words go markedly against the predictions.

This and other statistically significant contrasts from the data suggest that the typist used definite tactics to negotiate transition sequences. The inverted slowing between *h* and *i* preceding *m* suggests that these within-hand letters were typed as part of a group, coordinated in a compound movement trajectory. We have still to discover the rules governing such groupings, but the consistency of word-timing profiles leads us to suppose that they form a logistic component of a performance grammar.

How should the two results on timing be theoretically combined? Let us suppose that in fast typing a clock provides a periodicity, and departures from periodicity arise from the play of logistic rules over groups of letters. Then either the logistic rules arrange a timing pattern that moves away from and returns to the sequence of invariant clock pulses, in the manner of a musician making an excursion from the beat, or the clock rate is itself modulated by those rules. Shaffer (1981) suggested the latter to draw out a resemblance between the latency patterns in typing and musical rubato. An alternative assumption is that normal typing may not use a clock, but clock timing can be adopted as an exotic option if it is required or if the text induces a high degree of regularity in the movement sequence, as on "alternation" texts.

The idea that timing patterns move away from and return to an unvarying clock pulse is difficult to test in its general form, and going back to the data we find no obvious support for it. The idea of a flexible clock has been tested using cross-covariance analysis (Vorberg & Hambuch, 1978). This technique, applicable to sequences with unequal intervals, correlates pairs of intervals over the replications of a sequence. A more synoptic test is obtained by using analysis of variance (Shaffer, 1980, 1981), in which case negative covariation leads to *F* ratios in the left tail of the *F* distribution (Wainer, 1973). In effect it generalizes a test for covariance used by Kozhevnikov and Chistovich (Note 2).

Taking the data used to examine temporal pattern, 29 analyses of variance were carried out on letter latencies over whole words. Only half the *F* ratios were less than 1 and their mean was 1.03. This fails to support a clock model. Some of the *F* ratios were nevertheless so small, the smallest being .14, that the temporal patterns were reexamined in these extreme cases. They indicated something other than random fluctuation on an underlying pattern: It appears that the typist sometimes vacillated between two patterns, changing a slow–fast pattern into a fast–slow pattern for a pair of letters. Sometimes the two letters were not even adjacent.

We are led to conclude that normal fast typing has only a nominally specified periodicity. This is modulated by contingencies of movement transition, and for some of these contingencies there are logistic rules grouping movements into compound trajectories.

Music Performance

Because music has a beat and a meter, there is reason to expect its performance to be controlled by clock timing. As a preliminary to studying this, Vorberg and Hambuch (1978) have looked at metrical tapping, using the data to test models of concatenated or hierarchic clock structures.

A variety of timing structures can be obtained by assembling in different ways two basic methods of timing a rhythmic group of two beat intervals. One way is to generate the intervals in concatenation; the other is to generate the group interval and its first beat interval, which thus determines the second interval. These concatenated and hierarchic groups have different variance–covariance properties, which are inherited by the larger structures that assemble them.

Clock intervals, *C*, and their derived beat intervals, *B*, and also their variances and covariances, are related in the following ways (Vorberg & Hambuch, 1978). Let C_i be an independent clock interval ($i = 1$ or 2) and B_i be a derived beat interval. Then in a concatenation group

$$B_i = C_i,$$

$$\text{Var } B_i = \text{Var } C_i,$$

$$\text{Cov } B_i, B_{i+1} = 0.$$

In a hierarchic group C_2 generates the group interval and C_1 the interval between the first and second beats. Hence,

$$B_1 = C_1, \quad B_2 = C_2 - C_1;$$

$$\text{Var } B_1 = \text{Var } C_1,$$

$$\text{Var } B_2 = \text{Var } C_1 + \text{Var } C_2;$$

$$\text{Cov } B_1B_2 = -\text{Var } C_1, \quad \text{Cov } B_2B_1 = 0.$$

The last expression is of a covariance across a group boundary. These expressions describe abstract timing structures; the timing properties of observable motor events can be obtained by adding assumptions about the relationship between a beat and its corresponding motor event.

These ideas of structure were used to analyze two piano performances of a Bach fugue, obtained in succession from the same pianist (Shaffer, 1981). The fugue was in four voices, and so the pianist had to play four melodic lines at once. Each melody was a series of notes and rests varying in duration, and these durations at any moment were independent across voices. Playing the four voices in coordination while maintaining their rhythmic and melodic independence seems feasible only if the performer can use a clock, or clock structure, to generate an abstract meter and play the notes in relation to this. A covariance analysis performed on the data supported this conjecture. It appears that a clock was used to construct the half-beat intervals of the music, which coincided with the most frequent note duration: The cross-covariances between adjacent intervals were negative, and the pattern of covariance indicated that the intervals were generated in concatenation rather than in a hierarchy. On the other hand, the covariance between the durations of short notes that divided a half-beat interval favored a hierarchic structure, supporting the idea that the motor system constructed a compound movement trajectory to produce short notes.

Another feature was that the timing of successive metrical units of the music varied over the performance. This variation was replicated in the two performances and therefore represented expressive rather than random variation. Here we must suppose that expressive information in a motor program was used to modulate the clock rate. In the nature of our assumption that the motor system uses clock pulses as temporal reference points, it is necessary that both levels of timekeeper, clock and motor system, should receive the expressive information so that they can negotiate changes in tempo without becoming radically out of phase.

Also reported with the Bach fugue were studies of performances of a Chopin study and a Bartok dance, played by different pianists. These illustrate the mastery of different kinds of rhythmic complexity in musical performance. The Chopin study was an exercise in playing the polyrhythm of three against four, the right hand playing three equal notes in a beat interval while the left hand played four. In the performance it was played with a free use of rubato and with one hand moving off and back onto the beat. Yet it was possible to show by mapping the timing of one hand onto the other that under the two kinds of expressive deformation of timing the two hands preserved a proper arithmetic division of a common beat interval into three and four parts.

The computational demands of such a performance indicate the use of a hierarchy of timekeepers: a flexible clock to time the beats and timekeepers in the motor subsystems, one for each hand, to construct the appropriate subdivisions of the beat and allow one hand to move off the beat.

The rhythmic interest of the Bartok dance is that it has a meter of three beats unequal in length. made up of eight equal notes grouped three, three, and two. Much of the time one hand had to play chords on the beats while the other played the notes. Again using covariance methods to analyze the performance, it was found that the hand playing notes used a clock, and therefore carried the beat, while the hand playing chords was allowed to move off the beat.

In short, there is evidence in piano performance of a clock whose rate can be expressively varied. A motor system can produce compound movement trajectories that subdivide clock intervals and can undershoot or overshoot clock references in a controlled way.

We have incidentally shown that skilled pianists have greater hand independence and can make more complex use of timing than was observed by Kelso et al. (1979). In general, with prolonged practice the skilled performer goes progressively beyond the processing constraints that limit the unskilled person. Hence, we have argued that many problems of skilled performance cannot be studied within the conventional paradigms of introducing subjects to arbitrary laboratory tasks (Shaffer, 1981). There are, for instance, no counterparts of expressive timing in unskilled performance. The importance of studying speech in skills research is that almost every adult is a relatively skilled exponent of this complex activity.

Speech

In English speech there are two units that may plausibly have an underlying periodicity: the stress group and the syllable, both of which are contentious entities in linguistic theory. The major proponents of the idea that speech production uses an abstract clock were Kozhevnikov and Chistovich (Note 2), who thought of it as a rhythm generator. Fowler (1980, Note 6) has criticized such a theory on the grounds that it underestimates the timekeeping ability of the motor system.

Kozhevnikov and Chistovich found negative covariation between the durations of successive phoneme segments and between successive syllables. The results are replicable but there are reasons for resisting the inference of a clock (Ohala, Note 7). One of these arises from the difficulty of locating the boundaries of speech units from a physical record. The inference of a clock depends on being able to show, through a covariance analysis of time series, that there is an underlying time interval more stable than the one measured. If the observed interval contains measurement error, this can induce a negative covariance as an artifact and so create the illusion of hierarchic timing.

Another source of doubt arises from the method of obtaining data. Kozhevnikov and Chistovich had subjects repeat the sentence "Tanya topila banyu" 100 times in succession. In such a situation the speaker may fall into a rhythmic mode of speech that assumes clock timing. Thus, as in the typing of alternation texts, the task may induce clock timing, but it does not follow that this is the normal mode of speech.

The problem of inadvertently inducing clock timing in speech can be partly overcome by distributing the repetitions of target phrases in time or by embedding them in larger texts. The problem of measuring intervals separated by fuzzy boundaries is more intractable and may restrict the scope of covariance analysis as a test of clock timing. Rather than attempt to refine methods for testing a clock concept it may be better to finesse the problem and question whether the concept is relevant to a skill in which a clock serves no useful purpose. If timing is a primary variable in speech production, its specification can be interpreted directly by the motor system. We have argued that the basic role of an abstract clock is to provide a basis for coordination, and this is not a requirement of normal speech; people are not required to speak in unison, except in formal chants and church litanies.

If people repeatedly hear a recording of a spoken sentence and are then asked to tap in time with the rhythm, they readily do so, and they tend to interpret the instruction by tapping in time with the onsets of stressed vowels (Allen, 1972). Further, if they are asked to listen to recorded speech to detect irregularities in its rhythm, this task is easier if the speech has been distorted to affect the interval between the onset of stressed vowels (Huggins, 1972). If the distortion prolongs one segment and shortens another, then the two changes may go unnoticed if they leave the stress interval relatively unchanged. Thus we learn that stress rhythm is the perceptually dominant rhythm in speech, at least in languages such as English, and that the interval between stressed syllables may carry definite timing expectations.

A hypothesis, arising from Pike (1945), that stress groups in English are temporally regular, is easily falsified by measurement, which shows that the stress intervals in a single discourse may vary fivefold in magnitude (Lehiste, 1977). A weaker hypothesis—that stress timing is regulated in the direction of isochrony—is too weak to account for Huggins's result. An alternative strong hypothesis is that there are linguistic rules structuring degrees of stress over an utterance, and what the listener detects in the distorted recording is a violation of stress patterning determined by this structure. Liberman and Prince (1977) have shown how a relative-prominence rule, acting recursively on a hierarchy of binary syntactic constituents, can construct such a stress pattern over the sequence of syllables in a phrase or sentence. Their theory may need modifying to take account of the prosodic freedom of sentence stress assignment in making semantic and pragmatic distinctions (Bolinger, 1972; Cutler & Isard, 1980). The important point is that speaker and listener can orient to the syllables of major stress, using stress pattern rather than duration as the primary variable of speech timing. Duration has a more definite role in the pragmatic use of pausing, used to its greatest effect by actors and comedians, and as a cue to the syntactic structure of an utterance (Cooper, 1980; Huggins, 1978; Lindblom & Rapp, Note 8).

Stress is a variable acting on syllables rather than on phonemes. It affects the timing or intensity of all the phonemes in a syllable, although, as with speech rate, it affects the vowel more than the consonants (Klatt, 1976). Turning to the timing of phonemic segments, a case can be made that the rules governing segmental timing are based on the syllable or its demisyllables.

Segment duration is context dependent. Some of the dependencies can be stated by phonotactic rules, for instance, the effect on vowel duration of whether the following consonant is voiced or unvoiced; others require rules going beyond immediate adjacency, such as those affecting consonant duration in consonant clusters (Klatt, 1976). There is no loss in generative power in stating these as rules acting on units such as demisyllables. Fudge (1969) identifies the major division of a syllable into its initial consonant cluster and its rhyme, containing the nuclear vowel and a consonantal coda. Fujimura (1979) obtains a gain in descriptive power by separating the postvocalic consonants into those belonging to the core syllable and their affixes. The gain in stating phonological rules on such units is that it becomes relatively easy for the motor system controlling speech to compute compound movement trajectories (articulatory gestures) for syllabic groups containing one or more syllables.

The question arises whether the syllable has a nominally specified periodicity. In support of periodicity is the observation that adding consonants to increase the size of a syllable has the effect of compressing segmental durations (Lindblom & Rapp, Note 8). The effect does not produce an invariant syllable duration but it may indicate a regulation of duration. Another consideration is that speech appears to have a definite rate: In order to perceive stress variation and the slowing at syntactic boundaries as prosodic variables rather than as fluctuations in rate, there may need to be some underlying periodicity that these variables modulate. However we have not taken account, for instance, of the way in which pitch contour may affect the perception of stress or rate.

Two further points on speech timing are briefly discussed here. They are raised by Fowler (1980, Note 6), but the treatment here is different. The relative timing of vowels and consonants changes with stress level or speech rate. Fowler attempts to account for this by supposing that vowels and consonants are produced by parallel but coordinated articulatory subsystems. This does not achieve the desired result, however, because it predicts either that one stream of segments, vowels or consonants, should change without the other (being parallel) or that both streams should change proportionally (being coordinated). In the present theory a syllable is produced by a compound articulatory gesture. Its trajectory is computed to convey an internal rhythm of sound appropriate to the language or dialect when the syllable is uttered in citation form: Changing stress level or speech rate can affect the trajectory and in doing so alter the internal rhythm. Thus we are supposing that within-syllable rhythm is coded in terms of movement and that duration itself is not a primary variable even at this level of speech. Recalling the earlier argument about schedules, this is not inconsistent with the idea that speech has a nominally specified periodicity.

Alteration of internal rhythm in the syllable is not the only symptom of stress or rate change. Speaking at a faster rate or reducing stress also affects vowel color, moving it towards the neutral schwa. In the limit the vowel is deleted so that *governor* becomes *guvnor*. Under the pressures of rapid speech and casual style, several syllables may be compressed in an articulatory trajectory, as in "watcha gonna do?"

Fowler allows syllable boundaries to be defined by vowel succession and attempts to account for both timing and coarticulation phenomena in terms of the superposition of consonantal gestures on the continuous gesture of vowel production. This fails to account for the coarticulation of consonants and for coarticulation across syllable boundaries; it does not consider the timing of postvocalic consonants or show why syllable duration is affected by the size of the consonant clusters. In this paper we have separated the factors affecting movement timing, and hence coarticulation, from those governing event timing; the main purpose being to allow movements to anticipate their temporal goals. Some aspects of coarticulation have immediate acoustic consequences (Kent & Minifie, 1977), and so their timing must be computed in a compound articulatory gesture. For example, the velum must be lowered in preparation for a nasal stop consonant, and the timing of its lowering and return has to limit the spread of nasality to adjacent phonetic segments. On the other hand, many forms of coarticulation have no important acoustic consequences and their timing can be arranged to satisfy factors such as movement economics. Lip rounding in preparation for a vowel may have no effect on preceding consonants, and it has been observed to anticipate the vowel by as many as six intervening segments (Benguerel & Cowan, 1974): This goes beyond economy requirements, but it may be serving as a visually expressive gesture.

Reference Notes

1. Jackendoff, R., & Lerdahl, F. *A deep parallel between music and language*. Unpublished manuscript, Indiana University Linguistics Club, 1980.
2. Kozhevnikov, V. A., & Chistovich, L. A. Speech: Articulation and perception (Publication No. 30543). Washington, D.C.: U.S. Department of Commerce, 1965.
3. Gentner, D. R., Grudin, J., & Conway, E. *Finger movements in transcription typing* (Occasional Paper No. 4). La Jolla, Calif.: Center for Human Information Processing, 1980.
4. Michon, J. A. *Timing in temporal tracking*. Soesterberg, The Netherlands: Institute for Perception, RVO-TNO, 1967.
5. Rumelhart, D. E., & Norman, D. A. *Simulating a skilled typist: A study of skilled cognitive-motor performance* (Report No. 102). La Jolla, Calif.: Center for Human Information Processing, 1981.
6. Fowler, C. A. *Timing control in speech production*. Unpublished manuscript, Indiana University Linguistics Club, 1977.
7. Ohala, J. J. *Aspects of the control and production of speech* (UCLA Working Papers in Phonetics, No. 15). Unpublished manuscript, University of California, Los Angeles, 1970.
8. Lindblom, B., & Rapp, K. *Some temporal regularities of spoken Swedish* (Publication No. 21). University of Stockholm, Sweden, Institute of Linguistics, 1973.

References

Allen, G. D. The location of rhythmic beats in English: An experimental study 1. *Language and Speech*, 1972, *15*, 72–100.

Belenkii, V. Y., Gurfinkel, V. S., & Pal'tsev, Y. S. Elements of control of voluntary movements. *Biophysics*, 1967, *12*, 154–161.

Benguerel, A. P., & Cowan, H. A. Coarticulation of upper lip protrusion in French. *Phonetica*, 1974, *30*, 41–55.

Bolinger, D. Accent is predictable (if you're a mindreader). *Language*, 1972, *48*, 633–644.

Clark, H. H., & Clark, E. V. *Psychology and language*. New York: Harcourt Brace Jovanovich, 1977.

Cooke, J. D. The organization of simple skilled movements. In G. E. Stelmach & J. Requin (Eds.), *Tutorials in motor behavior*. Amsterdam: North-Holland, 1980.

Cooper, W. E. Syntactic-to-phonetic coding. In B. Butterworth (Ed.), *Language production 1*. New York: Academic Press, 1980.

Cutler, A., & Isard, S. D. The production of prosody. In B. Butterworth (Ed.), *Language production 1*. New York: Academic Press, 1980.

Forte, A. *Tonal harmony in concept and practice* (3rd ed.). New York: Holt, Rinehart & Winston, 1979.

Fowler, C. A. Coarticulation and theories of extrinsic timing. *Journal of Phonetics*, 1980, *8*, 113–133.

Fromkin, V. A. The non-anomalous nature of anomalous utterances. *Language*, 1971, *47*, 27–52.

Fudge, E. C. Syllables. *Journal of Linguistics*, 1969, *5*, 253–286.

Fujimura, O. An analysis of English syllables as cores and affixes. *Zeitschrift fur Phonetik, Sprachwissenschaft und Kommunikationsforschung*, 1979, *32*, 471–476.

Genest, M. L'analyse temporelle du travail dactylographique. *Bulletin du Centre d'Etudes et Recherches Psychotechniques*, 1956, *5*, 183–191.

Greeno, J. G., & Simon, H. A. Processes for sequence production. *Psychological Review*, 1974, *81*, 187–198.

Hollerbach, J. A competence model for handwriting. *Visible Language*, 1979, *13*, 252–264.

Huggins, A. W. F. On the perception of temporal phenomena in speech. *Journal of the Acoustical Society of America*, 1972, *51*, 1279–1290.

Huggins, A. W. F. Speech timing and intelligibility. In

J. Requin (Ed.), *Attention and performance VII.* Hillsdale, N.J.: Erlbaum, 1978.

Kelso, J. A. S., Southard, D. L., & Goodman, D. On the coordination of two-handed movements. *Journal of Experimental Psychology: Human Perception and Performance*, 1979, *5*, 229–238.

Kent, R. D., & Minifie, F. D. Coarticulation in recent speech production models. *Journal of Phonetics*, 1977, *5*, 115–133.

Klatt, D. H. Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *Journal of the Acoustical Society of America*, 1976, *59*, 1208–1221.

Ladefoged, P. *Three areas of experimental phonetics.* New York: Oxford University Press, 1967.

Lashley, K. S. The problem of serial order in behavior. In L. A. Jeffress (Ed.), *Cerebral mechanisms in behavior.* New York: Wiley, 1951.

Lee, D. N. Visuo-motor coordination in space-time. In G. E. Stelmach & J. Requin (Eds.), *Tutorials in motor behavior.* Amsterdam: North-Holland, 1980.

Lehiste, I. Isochrony reconsidered. *Journal of Phonetics*, 1977, *5*, 253–263.

Leonard, J. A. Advance information in sensori-motor skills. *Quarterly Journal of Experimental Psychology*, 1953, *5*, 141–149.

Lerdahl, F., & Jackendoff, R. Toward a formal theory of tonal music. *Journal of Music Theory*, 1977, *21*, 111–171.

Liberman, A. M. The grammars of speech and language. *Cognitive Psychology*, 1970, *1*, 301–323.

Liberman, M., & Prince, A. On stress and linguistic rhythm. *Linguistic Inquiry*, 1977, *8*, 249–336.

MacKay, D. G. Stress pre-entry in motor systems. *American Journal of Psychology*, 1971, *84*, 35–51.

MacNeilage, P. F. Motor control of serial ordering of speech. *Psychological Review*, 1970, *77*, 182–196.

Martin, J. G. Rhythmic (hierarchical) versus serial structure in speech and other behavior. *Psychological Review*, 1972, *79*, 487–509.

McGill, W. J. Random fluctuations of response rate. *Psychometrika*, 1962, *27*, 3–17.

Pike, K. L. *The intonation of american english.* Ann Arbor: University of Michigan Press, 1945.

Poulton, E. C. On prediction in skilled movements. *Psychological Bulletin*, 1957, *54*, 467–478.

Schmidt, R. A. On the theoretical status of time in motor-program representations. In G. E. Stelmach & J. Requin (Eds.), *Tutorials in motor behavior.* Amsterdam: North-Holland, 1980.

Schourup, L. Unique New York unique New York unique New York. *Chicago Linguistic Society. Papers from the Ninth Regional Meeting*, 1973, 587–596.

Selkirk, E. O. The role of prosodic categories in English word stress. *Linguistic Inquiry*, 1980, *11*, 563–605.

Shaffer, L. H. Latency mechanisms in transcription. In S. Kornblum (Ed.), *Attention and performance IV.* New York: Academic Press, 1973.

Shaffer, L. H. Intention and performance. *Psychological Review*, 1976, *83*, 375–393.

Shaffer, L. H. Timing in the motor programming of typing. *Quarterly Journal of Experimental Psychology*, 1978, *30*, 333–345.

Shaffer, L. H. Analysing piano performance. In G. E. Stelmach & Requin (Eds.), *Tutorials in motor behavior.* Amsterdam: North-Holland, 1980.

Shaffer, L. H. Performances of Chopin, Bach and Bartok: Studies in motor programming. *Cognitive Psychology*, 1981, *13*, 327–376.

Shattuck-Hufnagel, S. Speech errors as evidence for a serial-ordering mechanism in sentence production. In W. E. Cooper & E. C. Walker (Eds.), *Sentence processing.* Hillsdale, N.J.: Erlbaum, 1979.

Simon, H. A. Complexity and the representation of patterned sequences of symbols. *Psychological Review*, 1972, *79*, 369–382.

Stevens, L. T. On the time sense. *Mind*, 1886, *11*, 393–404.

Vorberg, D., & Hambuch, R. On the temporal control of rhythmic performance. In J. Requin (Ed.), *Attention and performance VII.* Hillsdale, N.J.: Erlbaum, 1978.

Wainer, H. The other tail. *British Journal of Mathematical and Statistical Psychology*, 1973, *26*, 182–187.

Wing, A. M. Effects of type of movement on the temporal precision of response sequences. *British Journal of Mathematical and Statistical Psychology*, 1977, *30*, 60–72.

Wing, A. M. Response timing in handwriting. In G. E. Stelmach (Ed.), *Information processing in motor control and learning.* New York: Academic Press, 1978.

Wing, A. M. The height of handwriting. *Acta Psychologica*, 1980, *46*, 141–151. (a)

Wing, A. M. The long and short of timing in response sequences. In G. E. Stelmach & J. Requin (Eds.), *Tutorials in motor behavior.* Amsterdam: North-Holland, 1980. (b)

Wing, A. M., & Kristofferson, A. B. Response delays and the timing of discrete motor responses. *Perception & Psychophysics*, 1973, *14*, 5–12.

Received April 13, 1981
Revision received September 1, 1981 ■