# Microtiming deviations in groove

Andrew McGuiness

Thesis submitted for the degree of Master of Electronic Arts by
Research of the Australian National University.

April 2005

# DECLARATION

**The thesis and computational model comprising this submission are my original work. All sources of data and other information are acknowledged in the list of references.**

……………………………………………..

Andrew McGuiness

April 2005

# PREFACE

## *Physical creativity*

I was originally drawn to study fine motor control in musical performance because of a phenomenon I have experienced when performing at the piano. I found that, in situations where a musical piece (of Western notated art music, in this case) was sufficiently well known and any technical difficulties sufficiently overcome, the most finely-tuned and expressive execution of short passages was always unplanned and in fact not the execution which I had practised. I discovered that I could at times allow my hands to take over the performance, and that the result would be more subtle and musically successful than I had consciously anticipated. In light of the theory described in this thesis, it seems possible that those short passages were in fact under the control of open loop motor programs. The schemata for the motor programs appear to have been dynamically generated, and to have taken into account both the musical structure of the piece, as laid out in the notation, and ephemeral details of the current performance.

According to the theory outlined in this thesis, a series of open loop motor programs, triggered at intervals within the macroperiod, control timing at finer temporal levels than the tactus. The short passages I played seemed outside of cognitive control and yet took account of the musical context at different levels. If they were in fact under open loop motor program control, it would seem that the mechanism for generating the schemata for the motor programs is capable of dynamically integrating input and making creative decisions – a kind of *physical creativity*.

I gratefully acknowledge the support of the MARCS Auditory Laboratories; and of the Centre for New Media Studies at the Australian National University.

# List of figures

# List of equations

# List of tables

# List of Appendices

# Extended Abstract

## Overview

Much work has and is being done on the topic of timing variation in Western notated art music: see Appendix A to this thesis and, representatively, (Bresin, 1998,2000; Bresin, De Poli, & Vidolin, 1992; Canazza, De Poli, Di Federico, & Drioli, 1999; Canazza, De Poli, Drioli, Roda, & Zamperini, 2000; Canazza, De Poli, Roda, & Vidolin, 1997; Canazza, Roda, & Orio, 1999; Cannazza, De Poli, Di Sanzo, & Vidolin, 1998; Clynes, 1983; Clynes, 1986a,b; Clynes, 1987; Clynes, 1992; Clynes, 1995; Clynes & Walker, 1982; Desain & Honing, 1991,1992,1993,1996; Epstein, 1988; Friberg, 1995; Friberg & Bresin, 1997; Friberg, Bresin, Fryden, & Sundberg, 1998; Gabrielsson, 1982; Honing, 1992; Honing, 2001; Mazzola & Zahorka, 1994a,b; Palmer, 1989,1996; Palmer, 1997; Palmer & Pfordresher, 2003; Palmer & van de Sande, 1995; Repp, 1990,1992,1998; Shaffer, 1980,1981; Shaffer, 1985; Shaffer, Clarke, & Todd, 1985; Sundberg, Friberg, & Fryden, 1991; Timmers, Ashley, Desain, & Heijink, 2000; Todd, 1985,1989; Widmer, 1994,1996,2000,2001). However, less investigation has been undertaken into microtiming deviations in repetitive musics with globally stable tempo. Such musics are here referred to inclusively as *groove musics* – some examples are: traditional African drumming; funk; and Latin music.

This thesis presents a theory (the *Covert Clock Theory*) and proposes a model of the production of musical groove. The computer model presented rests on arguments regarding what constitutes groove, both in a musical sense and - beyond the scope of musicology - in terms of how and why humans produce and respond to groove.

Through the examination of human perception and production of rhythm, I propose to develop a model, based on the Covert Clock Theory, of the generation of microtiming deviations characteristic of groove. Ideally, the model should be intuitively controllable by a musician setting a small number of parameters, and should generate deviation patterns which are musically useful and have similar characteristics to those generated by human musicians.

### What is groove microtiming?

Groove musics can be characterised as synchronic and repetitive. A macroperiod is defined within the music and material is repeated, with variations, in successive macroperiods. Systematic, consistent and musically significant variations from the temporal grid of strictly proportional timing have been found to exist in these musics, with deviations typically in the range ~5ms to ~50ms. Some limited musicological studies (see eg. (Bilmes, 1993b; Cholakis, 1999b; Freeman & Lacey, 2001; Freeman & Lacey, 2002)) of microtiming deviations in groove have been undertaken but these do not attempt to account for microtiming in terms of underlying timing processes. This thesis represents an attempt to integrate the musicology of microtiming with psychological and human movement theories of rhythmic timing control, in order to develop a theory of the process by which performers generate microtiming deviations. The theory and results have implications for the psychology of rhythm, for music technology, for musicology and music performance, and for motor planning theory in general.

## A new theory of rhythmic timing and entrainment

Theories and models of human rhythmic entrainment, musical timing and motor control are critically reviewed in light of the experimental and musicological literature. Evidence is presented for different psychological or perceptual processes for *a)* timing control within the range of rhythmic synchronisation, and *b)* time intervals below that range.

A new theory of production is presented in which a system of interlocking interval timers or clocks provides trigger signals to motor programs at the start of each isochronous interval. Interval timers with resting periods close to the period of the driving stimulus (which may simply be an isochronous pulse train) alter their period to entrain to the driving stimulus. Entrainment may be in any harmonic relation to the stimulus – clock triggers and stimulus pulses may be in a one-to-one relation, or a 1:2, 1:3, 1:4, 2:1, 3:1, or 4:1 relation. The ratios of clock period to stimulus period are effectively limited by the range of synchronisation, which is approximately between 200ms and 1500ms. Within this range, nested harmonic clocks may occur, with the longest period perhaps four times the shortest period. Harmonic relations between clock periods are more stable because they lead to most reinforcement between clocks, but non-harmonic ratios, although less stable, are also possible.

Interval timers with onsets which are temporally close interact to reset phase so that their onsets tend to occur together, without the involvement of any period correction process. In this way, clocks with harmonically related periods help to stabilise each other to produce more precise timing. No phase information is available from interval timers between triggers.

Timing control at intervals smaller than the current synchronisation period (ie. between the triggers of the fastest active interval timer) is achieved by open-loop motor programs triggered by the interval timer. Timing at this lowest level might (hypothetically) be achieved by means of a very fast (perhaps 1000Hz) oscillation and a counter mechanism, which are not available to cognitive control. Revisions to timing in the motor program schema are made following execution of the motor program, on the basis of sensory feedback.

The delay for revisions to take effect following execution is equal to the time needed for planning, sensory feedback, and any other delays attributable to the neuromuscular system. At least some of these delays contribute to the necessity for *functional anticipation* by the timer trigger of the sensorimotor goal. *Functional anticipation* is defined as the time between the timer trigger and the completion of the action. For instance, a sensorimotor goal of eg. producing a drum attack to be heard at a defined time within the rhythmic stream, requires a functional anticipation at least equal to the time needed to convey the command to the muscles and for the physical completion of the action.

Phase correction is modelled in the implementation of the Covert Clock Theory as a separate process from period correction - although referencing the same asynchronies between sensory feedback and expected sensory feedback following the interval timer trigger. Phase correction is thus a motor implementation process. In the theory, phase correction involves an adjustment to the amount of *functional anticipation.* The phase correction process functions in the same temporal range as microtiming deviations (ie. below ~50ms) and may be invoked in response either to subliminal timing perturbations of the driving stimulus, or to asynchronies resulting from interval timer errors or motor implementation errors.

## Original contribution

### The *Covert Clock* Theory

A new theory and computational model of microtiming deviations in groove music are presented. The *Covert Clock Theory* of microtiming deviations posits the existence of an internal clock or interval timer at a cross-rhythm to the tactus clock – that is, with a period at a non-harmonic ratio to the tactus period. Such cross rhythms are made explicit in eg. African traditional drumming, where a cross-rhythm of three beats against four beats of the tactus is often established.

The Covert Clock Theory suggests that the covert clock and the tactus interact so that a given clock onset may occur earlier or later, but without the actual period of the clock being adjusted. Early or delayed triggering of the beginning of the interval timer cycle will of course result in some inter-trigger periods being shorter or longer than the clock period; but a key point of the theory is that the clock retains its previously established period.

The coincidence of the main and covert clocks at every four beats of the tactus and every three beats of the covert clock (in the example given) defines the start of the macroperiod. The macroperiod has duration within the capacity of echoic memory – the ‘psychological present’.

The *Covert Clock Theory* can be seen as a means of extending the duration of motor programs. Other devices found in traditional African drumming are also discussed as fulfilling the same function.

## The *VROOGE* model

The *VROOGE* model of microtiming deviations in groove is a software implementation of the microtiming process described by the Covert Clock Theory*.* (The name *VROOGE* is simply an anagram of ‘groove’.) The VROOGE model entrains to an external isochronous pulse train through period correction of a main tactus interval timer. Weighted referencing by the motor timing process to both the tactus and the covert clock results in changes to the local tempo of motor program implementation, without altering the period of the underlying clock. This facet of the model corresponds to expansion or contraction of motor program duration, as described in the Covert Clock Theory. Local tempo changes in motor program implementation from one tactus trigger to another result in progressively increasing or decreasing deviations through the duration of the tactus. The deviation amount is in fact the integral of the tempo change relative to the base tempo given by the (unchanged) tactus clock.

The *VROOGE* model is evaluated through analysis of microtiming data of examples of groove music and comparison with model –generated data. Predictions of the theory are tested against the data and some agreement is found.

The VROOGE model is successful in providing a means of generating new musically meaningful microtiming deviation patterns by making a low number of parameter settings, to which the model responds in a musically intuitive way.

# Chapter 1 Psychological and musicological approaches to groove microtiming

The first section of the thesis comprises a review of the literature dealing with the psychology of musical rhythm and meter, to provide the basis for developing a theory of microtiming deviations. As with the rest of the thesis, the focus will be on groove music – in this section, in its relation to the perception and production of rhythm and meter. Chapter 3 provides relevant musicological background, while Chapter 4 relates the musicological material of Chapter 3 to psychological theory.

The musicological characteristics of groove are dealt with more fully below. For the present, it is enough to note that groove music has a focus on rhythm, expressed via repetitive patterns usually of the same length, within the framework of a tempo which is apparently invariant. The musicological section will mainly discuss various relevant aspects of traditional African drumming practice (from several different cultures), although the main thrust of the theory is applicable to the wider range of musics which could be called groove, including some of the musics of the African diaspora (such as Cuban music and funk).

The *Covert Clock Theory* of microtiming deviations is presented in Chapter 5; while Chapter 6 tests results of the *Covert Clock* theoretical model against microtiming data from a range of musics. Chapter 7 reviews existing computational models of entrainment, and such models of microtiming which are relevant for groove music. Chapter 8 presents the VROOGE model of microtiming deviations, based in the *Covert Clock* theory described in Chapter 5.

Appendix A reviews models of microtiming which are not relevant to groove music. These are mainly related to stylistic structural aspects of Western notated art music.

We turn now to a definition of terms used in the thesis.

## 1.1 Defining terms

Before attempting to define groove music, and what separates it from, say, 18th century classical music, it is necessary to clarify the meaning of commonly used terms relating to rhythm. Words such as *beat*, *meter*, *measure*, *accent,* and even *rhythm*, are used at different times with a variety of overlapping meanings.

### 1.1.1 Pulse, beat, tactus, tatum,

Snyder (Snyder, 2000) describes a *beat* as a single point in time, with no duration but only a temporal position. It is an imaginary temporal reference point which need not be represented by a musical event, one of a series of such points which compose a *pulse*. This raises the question, “How do we know it is there?” Beats are the components of a series of identical (unaccented) and isochronous points in time – so that once a pulse is established, each beat does not need to be explicitly marked by a musical event for us to retain a sense of an ongoing stream of isochronous beats.

Fraisse (Fraisse, 1982) noted that people can easily match a motor act with a regular series of sounds, which indicates a strong link between rhythm perception and production (Krumhansl, 2000). An important feature is the ability to tap to more complex rhythmic sequences (Fraisse, 1982) - that is, to find a regular pulse which fits a musical sequence, although the music contains musical events which fall on beat subdivisions, and may not include a musical event on every tactus pulse. Clearly, there is a choice of pulse tempos from which to choose in this situation, and the pulse level at which most people tap may be termed the *tactus,* following Lerdahl and Jackendoff’s use of the term. (Lerdahl & Jackendoff, 1981) The tactus can be subdivided into equal or quasi-equal parts (eg in Western music notation, eighth-note subdivisions of a quarter-note tactus may be made) and the term pulse is then used to refer to the stream of tactus subdivisions which are most often sounded in the music.

Tactus subdivision parts may themselves be subdivided. The smallest subdivisions of the tactus have usefully been dubbed *tatums* (for **t**emporal **atom**) by Bilmes (Bilmes, 1993b), in tribute to stride pianist Art Tatum.

### 1.1.2 Meter

*Meter* refers to a cyclic grouping of either tactus beats or tatums, or both. In the last case, tactus beats are subdivided as well as being organised into repeating higher-level groups. Binary subdivisions at one or more levels are common, as are ternary subdivisions. Compound meters, comprising eg. binary groupings or subdivisions at one level and ternary groupings or subdivisions at another level are found in different cultures.

Meters which apparently include tactus beats of different length, formed by grouping different numbers of subdivisions, are found in eastern Europe –

asymmetrical or so-called ‘limping’ meters. For example, groupings of 12 subdivisions in a 3-2-2,3-2 pattern, 11 subdivisions in a 3-2-2-2-2- pattern, 7 subdivisions in a 3-2-2 pattern, and 9 subdivisions in a 2-2-2-3 pattern are found variously in areas of Greece, Turkey, Albania and the former Yugoslavia. (Friedberg, 2000; Various, 2000) Justin London (London, 2000) has pointed out that such complex meters conform to the rule of *maximal evenness*, so that there are only two different groupings of subdivisions in the meter, and the two groupings are as close as possible, for the given number of tactus beats.

One interpretation of meter, based on Lerdahl and Jackendoff’s influential Generative Theory of Tonal Music (Lerdahl & Jackendoff, 1983), includes an hierarchical system of accents relating to the several grouping levels. Meter of this kind is one of the meanings of the word as used here, but should not be thought of as privileged in any way.

### 1.1.3 Rhythm

*Rhythm* has been variously used to mean any of: pulse, meter, pattern of durations and accents, or as a collective term to refer to the interaction of all of these in a musical utterance. I will use the term *rhythmic pattern* to refer to patterns of interonset intervals only.

### 1.1.4 Groove

*Groove* is a term used to describe the rhythmic feel of music with a stable global tempo. Analysis of musical examples of groove, including jazz, Cuban drumming, and funk music, has found systematic *microtiming deviations* from quantised onset times. These microtiming deviations are the focus of the theory and model presented in this thesis. The theory and model ignore the contribution to groove of accent patterns, subtle loudness changes, offset timings, and timbre - although musicians would generally agree that these are all important components of the perception of groove.

*Groove music* here has the special meaning of music which is manifestly cyclical in nature, with clearly defined macroperiods, a focus on rhythm, and minimal variation of melodic and harmonic material (where it is present at all). Pressing (Pressing, 2002) defines groove as:

- '… emerging from one or more carefully aligned rhythmic patterns, characterised by:
- perception of recurring pulses, and subdivision structure to such pulses,
- perception of a cycle of time, of length 2 or more pulses, enabling identification of cycle locations, and
- effectiveness in engaging synchronizing body responses (eg. dance, foot-tapping)…'

- with what Pressing calls 'Black Atlantic groove' (based on the West African diaspora) showing 'relatively high levels of syncopation' while emphasising a 'metronomic' approach to timing, without rubato. (Pressing, 2002) The word 'metronomic' should not be taken to exclude microtiming variations, but does mean that rubato - variation in global tempo for expressive reasons – is not an attribute of groove music (see the discussion in Appendix A).

The definition of *groove* given above excludes jazz (which does not usually exhibit a defined, repetitive macroperiod) but includes funk and Latin music, and traditional African drumming; as will be argued, it also usefully fits some music to which the term groove is not usually applied, notably North Indian art music.

# Chapter 2 Review of theories of meter and entrainment

This chapter reviews experimental and theoretical literature on human entrainment to, and production of, an isochronous pulse, as a means of providing a temporal grid as a reference for microtiming deviations.

Two levels of timing control are identified. The first is the range of synchronisation or entrainment, with period in the range of ~200msec to ~1500msec – some kind of internal clock or interval timer is presumed to exist to produce these intervals. Motor programs are discussed as providing the second level of timing control, for events which occur between the beats of the (fastest such) internal clock or interval timer.

A theory of multiple internal clocks, with periods in hierarchical relation to each other is proposed. This new theory is constructed by drawing principles from existing theories and in later chapters is further developed into the *Covert Clock* theory of microtiming deviations which is implemented as computer software.

## 2.1 The concept of an internal clock

Most people can tap to a beat, dance in time and create music in time with others - this is evidence that people can entrain to an isochronous beat. In fact, even if the beat is not completely steady, entrainment is possible: even poor musicians keep time with each other, even though not as well as good musicians. (Fraisse, 1982)

Furthermore, people can produce the same behaviour without an external stimulus: they can tap an isochronous (or quasi-isochronous) beat without hearing anything but their own taps, dance without music, and sing or play the piano solo. In fact, even more basic activities display a quasi-isochronous temporal structure: most people walk, run, chew, engage in sexual intercourse, or ride a bicycle, in a rhythmic way. Babies suckle rhythmically and chefs chop vegetables rhythmically.

It follows that there is at least one kind of timing mechanism underlying the production of rhythmic behaviour in humans. Models of entrainment generally depend on the concept of an intrinsic frequency to give the period of rhythmic behaviour, with some degree of independence from the stimulus periodicity. (Large, Fink, & Kelso, 2002) The intrinsic frequency may relate to the period of either an emergent oscillation (for instance, as described in (Kelso, Holt, Rubin, & Kugler, 1981)), or some other kind of internal timekeeper mechanism (eg. (Vorberg & Wing, 1996)). The main difference between the two approaches is that phase information is presumed to be available from an oscillator, whereas a clock or interval timer provides only a trigger at the start of the timed interval. [In both cases, phase *correction,* relative to the stimulus, must be achieved either by adjusting the period of the oscillator or interval timer, or by a separate, direct phase correction process. A model described by Hary and Moore (Hary & Moore, 1985,1987), based on their experimental findings, postulates phase correction as provided through a mixed strategy of reference to either the external stimulus or the subject's own tap.] The theory proposed in this thesis does not directly address the oscillator versus clock debate: although an oscillator is used to give the sign (negative or positive) or asynchrony in the implementation of the computation model of entrainment, this is pragmatic choice rather than a theoretical preference.

Researchers have clarified a number of attributes of the phenomenon of human rhythmic behaviour, including the range of tempi and spontaneous tempo.

### 2.1.1 The tempo range of rhythmic behaviour

Fraisse stated that the range of periods to which synchronisation was possible was 200msec to 1800msec (Fraisse, 1982). He identifies a number of spontaneous activities of the human organism such as the suckling reflex of the newborn with (period between 600milliseconds and 1200ms); and walking (period about 550ms). (Fraisse, 1982)

Parncutt's (Parncutt, 1994) summary of the experimental literature on entrainment points to periods around approximately 600msec as having the greatest salience, with decreased salience at periods smaller than 400msec and longer than 900msec. Entrainment to periods below about 200msec and above about 1800msec is thought not to occur. (Parncutt, 1994)

### 2.1.2 Spontaneous tempo and the referent time level

When listening to music, most people will choose one level of pulsation as primary and use it as a frame of reference for faster and slower time levels in the music. The primary level has been termed the *referent time level* by Jones and Boltz (Jones & Boltz, 1989), and the tactus by Lerdahl and Jackendoff. (Lerdahl &

Jackendoff, 1981) If people are asked to tap to music, it is likely that they will make explicit their referent time level for the music. Parncutt (Parncutt, 1994) tested subjects tapping responses to six different cyclically repeating interonset interval patterns, each presented at six different tempos, from very slow to very fast. Parncutt found that responses gravitated toward a moderate pulse period of about 700msec.

Fraisse (Fraisse, 1982) states that "…sucking movements occur at a cadence that seems to be characteristic for each infant." He refers to studies of *spontaneous tempo*, found by free tapping, by Stern(Stern, 1900), also called personal tempo by Frischeisen-Kohler (Frischeisen-Kohler, 1933), and mental tempo by Mishima (Mishima, 1951-1952). Krumhansl (Krumhansl, 2000) follows Fraisse's review (Fraisse, 1982) which places the period of personal or mental tempo between 380ms and 880ms – with a period of 600ms as the most typical.

*Table 2-1 Summary of experimental findings for period of referent time level and spontaneous tempo.*

| Source | Method | Period |
|---|---|---|
| *Experiments to find most salient time level in response to musical examples* | | |
| Drake (Drake, 1998) | Test of sensitivity to tempo changes, infants to adults | 400 – 800msec |
| Parncutt (Parncutt, 1994) | Tapping to cyclic patterns to find referent time level | 700msec |
| Moelants (Moelants, 2002) | Tapping to music at a range of tempi | 370-740msec;<br>500msec most typical |
| *Experiments to find natural period of tapping, without a musical example being given* | | |
| Fraisse (Fraisse, 1982) | Review of studies of spontaneous tempo by Stern(Stern, 1900), Frischeisen-Kohler (Frischeisen-Kohler, 1933), and Mishima (Mishima, 1951-1952) | 380 – 880msec;<br>600msec most typical |
| Fraisse, Pichot & Clairouin (Fraisse, Pichot, & Clairouin, 1949) | Tests of spontaneous tempo | 200 – 1400msec |
| Drake et al (Drake, Jones, & Baruch, 2000) | Tests of spontaneous tapping | 600msec |
| Moelants (Moelants, 2002) | Tapping ‘at a comfortable rate’ | 370 – 740msec;<br>500msec most typical |

A summary of experimental findings for the referent time level and for spontaneous tempo is given in Table 2-1. The tempo ranges given for experiments to find the natural period of tapping without a musical example being given refer to the range over the sample of subjects. The variability within a subject has been found to be only 3% to 5% (Fraisse et al., 1949) and to be stable across different trials of the same subject. (Harrel, 1937; Rimoldi, 1951) On the other hand, an individual’s tempo range will be wider in experiments to find the most salient time level in musical examples, as it depends on the availability of a time level in the musical example at the preferred tempo. (Drake, 1998; Moelants, 2002; Parncutt, 1994)

Based on these findings, I will refer to the ‘moderate tempo’ as the tactus and assume its period varies around approximately 500msec, with the lower and upper limits of entrainment roughly half and triple that period respectively. The upper bound of entrainment has little impact on the current model. However the lower bound will be significant in differentiating between central clock-controlled timing, and motor implementation timing (see below).

## 2.2 A multiple clock model

A multiple-timer model has been developed by Ivry and Richardson  and Ivry, Richardson et al (Ivry & Richardson, 2002; Ivry, Richardson, & Helmuth, 2002) to account for decreased timing variance within both hands in bimanual *simultaneous* (in-phase) tapping, found in experimental studies of patients with unilateral cerebellar lesions, (Helmuth & Ivry, 1996) and subsequently in neurologically-healthy subjects. (Ivry et al., 2002) This multi-effector advantage was found even when people tapped with finger movements on one side of the body and forearm movements on the other side, or in finger and foot combinations.  Data analysis of these conditions showed increased *motor* variability when the two limbs were of unequal mass, but decreased central (clock) variability as calculated according to the Wing and Kirstofferson (Wing & Kristofferson, 1973a; Wing & Kristofferson, 1973b) two-process model . (Ivry et al., 2002) The multiple effector advantage persisted even for practised musicians, with very low baseline variances. The improved timing found with multiple effectors was even better when three effectors were moved in synchrony. The multiple timer model theorises that a new timer is engaged with each additional effector. The model incidentally accounts for the preference of musicians to move in time with the rhythm, to improve timing.

Ivry and Richardson characterise the multiple timers associated with the multiple effectors as gated activation functions which trigger both the effector action and the start of the next timer interval as the function exceeds a threshold. The synchrony of effectors and reduced clock variability arise from the averaging of interval start times between timers, through the summing of timer activations towards a higher threshold, the sum of the thresholds of the individual timers. The timer intervals thus remained temporally coupled. Temporal tuning in response to an external rhythmic stimulus, or period correction, is carried out by weighting the contribution of timers with different fixed intervals. (Ivry & Richardson, 2002) Once synchronisation with an external stimulus pulse is achieved, the timers operate in open loop mode, without feedback – thus there is no period correction, once synchronisation to an external pulse stimulus is achieved. Ivry and his colleagues explicitly reject either continuous or discrete feedback from the limbs as a source of improved clock stability in the multiple effector advantage. A discrete linear phase correction process added to a multiple timer simulation and based on feedback of asynchronies between the two limbs actually increased motor implementation timing variability. The authors argue that, although continuous proprioceptive-based feedback would probably reduce motor timing variability, it is unclear whether such a process would actually be viable during tasks such as repetitive tapping. (Ivry et al., 2002)

The Ivry and Richardson model may account for spontaneous tempo: if individuals have neural machinery or processes akin to a bank of timers, each with a fixed frequency, spontaneous tempo might arise from optimal period nesting as described above. Tempi other than spontaneous tempo would then be produced by adding weighted contributions from other clocks. The model allows for both spontaneous tempo stability and tempo flexibility. Period correction is achieved through weighting the contributions of timers with different fixed periods; and conversely the phases and slightly different periods of multiple timers required to control multiple effectors in synchrony is achieved in the same way. Note that period correction in response to an external pulse train will involve *discrete* time referencing – weights will be updated in response to perceived asynchronies between stimulus and response events. Internally, the timers utilise *continuous* referencing of each other's activation levels to achieve synchronisation.

The multiple timer theory is broadly compatible with the theory of mutual entrainment of clocks outlined above, with the added condition that separate motor effectors activate separate clocks even when engaged in simultaneous actions, when the clocks would be of the same period and phase. In fact, the activation functions can provide the foundation for a useful model of mutual entrainment, different from models of entrainment to auditory pulse trains which depend on discrete-time referencing of asynchronies between the stimulus and response. The averaging of two (or more) activation functions as described above, with the addition of a weighting parameter for the contribution of each function, is equivalent in output to the continuous mutual entrainment of the two (or more) interval timers involved.

However, while Ivry and Richardson’s gated activation function effectively accounts for the temporal coupling of multiple timers to control synchronous effector actions, a drawback appears when it is extended to accommodate nested clocks. Gated activation will keep timers in phase which are at or near the same period but does not readily adapt to entrainment at harmonic frequencies.

## 2.3 Hierarchical clocks

Although listeners may have a preference for tapping at a particular rate for a given piece of music, they can often choose to tap to a pulse level at a faster or slower tempo. In Western music this faster or slower pulse will often be at periods of half or double the tactus. Higher number subdivisions or groupings (eg. of three and four, or even of prime numbers such as five) are also found, but multiples of two and three are much more common.

It seems reasonable to assume that timing for tapping at a particular pulse level will be controlled by a clock with the same period as the pulse at that level. The fact that people who can tap at more than one pulse level can also generally move at will between pulse levels supports the further assumption that the clocks at different levels may be active whether they are the basis for the current level of tapping period or not. This leads to the concept of *nested clocks*.

A clock with a period at some multiple of the tactus (ie at a sub-harmonic frequency to the tactus) can be called a *super-tactus* clock; while a clock with a period at some factor of the tactus (at a harmonic frequency) can be termed a *sub-tactus* clock. A simple example is a tactus with period 600msec, a super-tactus with period 1200msec, and a sub-tactus with period 300msec. Of course, the tactus period may be too large to permit a super-tactus – a tactus period of 1000msec would give the smallest super-tactus a 2000msec period, outside the range of synchronisation as noted above. A fast tactus speed will put any subdivision below the level of the fastest clock in the range of synchronisation (200msec). On the other hand, a short tactus period of, say 250msec will allow the participation of clocks at 500msec and 1000msec, creating two clock levels above the tactus; and the reverse is obviously true for a long tactus period.

### 2.3.1 Evidence for an hierarchical clock system

Some of the earliest work on how people synchronise to patterns of interval durations other than isochronous series was undertaken by Povel, (Povel, 1981) who asked subjects to tap in time with sequences made up of intervals in various ratios to each other, and then to continue without the stimulus. In one experiment two intervals in were repeated and in another, the two intervals were presented in a S-L-L pattern. The shorter interval ranged over the two experiments from 200msec to 533msec. The ratios used were 1:4, 1:3, 2:5, 1:2, 3:5, 2:3, 3:4, and 4:5. The 1:1 ratio was not present. Subjects tended towards 1:2 timing in reproducing all intervals, from which Povel inferred that simple ratios are easier to reproduce. However, subjects had difficulty reproducing even the 1:2 ratio for the S-L-L patterns.

An internal clock account of Povel's (Povel, 1981) results for the two-interval experiment would attribute the dominance of the 1:2 pattern to its being the only one for which a subdivision of the pattern period is feasible. A 1:2 ratio allows for a subdivision into 3 equal parts, each part with a duration (in Povel's experiment) of 334msec – within the range of synchronisation and therefore capable of being clocked. Less simple ratios require the shorter interval to be subdivided putting the clock period below the range of synchronisation. For instance, both the 2:5 ratio and the 3:4 ratio require a subdivision into 7 equal parts. (The 1:4 and 1:3 patterns had a short interval of ≤ 250 msec and so were already at the edge or outside the range of synchronisation.)

The S-L-L pattern also used short intervals of about 250msec, except for three cases, two of which had complex ratios which would lead to subdivisions below that level. For the 1:2 ratio with the short interval at 400msec, standard deviations were *higher* than for the first experiment. The added difficulty, according to the hierarchical clock theory, here came from the addition of the second long interval, making the period of the pattern 2000msec – above the range of synchronisation.

Povel (Povel, 1981) presented intervals in longer patterns, which allowed for the emergence of a higher-level beat, covering repetitions of the short intervals. Patterns which were reproduced well were (with intervals in milliseconds):

- 250-250-500 (in the hierarchical clock theory, this would be a sub-tactus at 250msec, tactus at 500msec, and super-tactus a 1000msec);
- 250-250-250-750 (subtactus 250ms, tactus 750ms, supertactus 1500ms);
- 250-250-250-250-1000 (subtactus 250ms, tactus 500ms, supertactus 1000ms;

- 250-250-250-250-500-500 (subtactus 250ms, tactus 500ms, supertactus 1000ms)

Two patterns were not well produced:

- 200-200-600 (this produces a period of five 200ms beats – the short intervals are outside the range of synchronisation assumed here, and do not concatenate into a factor of the long interval);
- 400-400-400-600-600 (once again the subdivision of the 400ms tactus to produce the syncopation of the last onset requires a clock at 200ms, below the range of synchronisation.)

While Povel (Povel, 1981) introduced a theory of beats, and recognised that the range of synchronisation is approximately 250ms to 1500ms, he does not appear to have considered an hierarchical clock system to account for the his experimental findings. Povel (Povel, 1981; Povel, 1984; Povel & Essens, 1985) developed a rule-based coding system to account for the choice of period and location of temporal grid to match to cyclical patterns of inter-onset intervals but the preferences shown for a particular tactus in response to a pattern can also be accounted for by attributed to a system of nested clocks. The introduction of pitch can affect the selection and location of grid, as can the introduction of loudness accents. Pitch and accent appear to be used as cues to indicate the location of the temporal grid in the stream of onsets; however these factors are beyond the scope of the present model and do not contradict the principle of hierarchic meter.

Desain and Honing (Desain & Honing, 2003) note the ability of listeners to unconsciously abstract a rhythm matched to a temporal grid from timing information which is in fact quite noisy. In one experiment, subjects were asked to categorize rhythms generated from every combination of four onsets within a fixed duration on a fine temporal grid. Unsurprisingly from the point of view of the hierarchical meter theory, 97% of responses categorized the rhythms to a sequence of integers summing to a power of 2 or 3. The data are used to test their DECO (Desain & Honing, 2002) (de-composable theory of rhythm perception) model of the stability of rhythmic patterns, based on bonding between adjacent intervals preferably in a simple integer ratio to each other. However, once again hierarchical meter theory provides a simple explanation for many of the model’s rules: the interval pattern 2:5:1 is more stable than 5:2:1, for instance, because 2:5:1 has the reinforcement of a sub-tactus clock at a frequency four times that of the period of the cycle. The sub-tactus clock pulse falls on groupings of 2 subdivisions. So, in the 2:5:1 interval pattern it arrives at the start of the 2-interval, and the start of the 5-interval, reinforcing those onsets and stabilising the perception of the pattern. But for the 5:2:1 interval sequence, the subtactus clock (still falling on groupings of 2 subdivisions) coincides only with the start of the 5 interval (and not with the 2 interval and the 1 interval). Actual interval durations are not given and are apparently not a feature of the model, whereas in the hierarchical clock theory the upper and lower bounds of the range of synchronisation are a key factor in choice of tactus, and therefore in the stability of rhythmic patterns.

Vorberg and Hambuch (Vorberg & Hambuch, 1978) appear to have been the first to explicitly propose the possibility of two nested interval clocks to control different levels of the metrical hierarchy in music. Analysis of data from a synchronization tapping experiment where subjects were required to synchronize the first tap in a group of either one, two, three or four taps with a regular stimulus, found positive peaks in serial covariance groupings, at lags equal to the groupings (with the exception of some groupings of 500msec period taps.) Inter-tap intervals (ITIs) were either 300msec or 500msec. However, the authors finally reject the notion of hierarchically nested clocks, because they found that the variance of “natural” groupings (starting with the tap coinciding with the stimulus) had no tendency to be less than the variance for groups overlapping the stimulus boundary. However, the hierarchical clock models discussed by these authors involve intermittent activation of interval timers at lower levels (Vorberg & Hambuch, 1978). A full hierarchically-nested clock theory would include serial timing at lower levels, *together with* by serial timing at higher levels. In this case, the regulation of the lower-level period provided by the higher-level clock would be expected to stabilise the lower-level period, so that variance would be reduced overall, rather than only at the level of the higher clock. A further consideration is that groups of three or more 500msec period taps would produce a higher-level period outside the range of synchronization assumed here. The authors do note that testing over a larger tempo range might produce different effects. (Vorberg & Hambuch, 1978)

Hierarchically nested clocks were again postulated by Monahan (Monahan, 1993). Clarke (Clarke, 1985) had measured timing covariance at and below the beat level in performances by a concert pianist of Erik Satie’s Gnossienne No. 5, and proposed that positive covariances at the beat level combined with negative covariances  for nominally equal subdivisions divisions of the beat indicated direct timing of the beat interval but not of subdivisions (see discussion, below). However, small number divisions of the beat (two to four, at a beat period of over 1000ms) had positive covariance, which Monahan cited in support of nested interval clocks at the beat level and at twice the frequency (or half the period) of the beat.

Listeners in an experiment by Fraisse (Fraisse, 1967) were asked to adjust the interval between two isochronous pulse trains (at the same tempo) to produce isochrony (at twice the tempo of each separate sequence). Performance was best for periods of about 500-600msec (cited in (Parncutt, 1994)), consistent with a fastest clock  period of about 250-300msec.

Experimental data obtained by Drake (Drake, 1993) and cited by Parncutt (Parncutt, 1994) shows that binary/quaternary grouping dominates over ternary grouping for both children and adults. Binary groupings maximise the number of pulse levels which fall within the range of synchronisation  – ternary subdivisions mean that sub-tactus clocks must be faster, while ternary groupings of the tactus mean that super-tactus clocks must be slower. Data obtained by Fraisse (Fraisse, 1982) and others  supports this interpretation: when subjects were asked to tap in groups of three or fours, intertaps intervals were shorter than the usual 600ms, and averaged 420ms for groups of three and 370ms for groups of four. The period of a three grouping of 420ms intervals is 126ms, and the period of a four grouping of 370ms intervals 1480ms. Bisection of either of the component intervals would fall outside the synchronisation range. Drake (Drake et al., 2000) explicitly states that the ability to respond to other pulse levels than the referent (tactus) level depends on the activation of multiple oscillators .

The usual choice of ~600msec as a preferred ungrouped rate is consistent with covert bisection to take advantage of a 300ms sub-tactus clock, and covert grouping by twos, to take advantage of a super-tactus clock at 1200ms. Mutual entrainment of the three clocks would mean that the middle-period clock would be reinforced, and the stability of all three enhanced. Differences in spontaneous tempo between individuals would then relate directly to differences in the upper and lower bounds of clock periods they were able to produce. Indeed, according to Parncutt (Parncutt, 1994) the subjective accents which on the first and third elements when grouping by fours suggests the involvement of three consonant pulse sensations of periods 1, 2 and 4 beats.

One of Lerdahl and Jackendoff's (Lerdahl & Jackendoff, 1981) rules for metrical well-formedness is, 'Every beat at a given level must also be a beat at smaller levels.' Parncutt's (Parncutt, 1994) model of pulse salience and metrical accent states that:

> 'The salience of a perceived meter, or the probability of a given metrical interpretation, is proportional to the sum (or some other aggregate) of the salience of the pulse sensations that make it up. The most likely meter to be perceived is the one with the highest predicted salience.' (Parncutt, 1994)

And further, that there is mutual salience enhancement among consonant pulses, so that sub-tactus, tactus and super-tactus clocks, where they exist contribute to increased salience of pulses where they coincide. The frequently-noted predominance of binary/quaternary grouping over ternary grouping is implicit in Parncutt's model of pulse salience, as no consonant sub-tactus can occur in three-groupings (because, unless synchronisation can occur with the fastest pulse - giving three subtactus to each tactus – no even subgroup can be made within the three-grouping.) The fastest pulse is presumed here to be too fast for internal clock synchronisation.) Binary relations also maximise the number of consonant pulses within the range of synchronisation, since the period of a super-tactus clock will be 3 times that of the tactus and nine times that of the sub-tactus where ternary relations prevail, but only twice and four times the periods of the tactus and sub-tactus with binary relations.

Large, Fink & Kelso (Large et al., 2002) suggest the ability of subjects to synchronise tapping to different referent time levels may reflect the 'simultaneous representation of three periodicities within a single coordinative structure' (Large et al., 2002); and that the ability to synchronise with unpredictable patterns at prescribed metrical levels indicates the interaction of different levels in temporal tracking (Large et al., 2002) – in other words, a system of hierarchical clocks.

### 2.3.2 Nested clocks and better timing performance

So far, clocks (or oscillators) have been considered as entraining to external rhythmic stimuli. Tactus and sub-tactus clocks at adjacent pulse level and entraining to the same stimuli would synchronise the start of their timing intervals in response to the same stimuli. However, unless there were unambiguous cues as to the relative salience of sub-tactus intervals (which evidently is not always the case – see eg. Parncutt (Parncutt, 1994) on meter ambiguity) the super-tactus interval could conceivably begin on a different sub-tactus than the tactus. In other words, the tactus could coincide with the first of each pair of a binary sub-tactus, while the super-tactus coincided with the second of each pair. Mutual entrainment between internal clocks would ensure that the start of the timing interval of all nested clocks coincided with the start of the timing interval of the slowest clock.

Repp (Repp, 2002b) describes an experiment in which subjects tapped on-beat (ie. at the tactus level) and timing perturbations of the tactus onset or one or more subtactus-rate onsets were introduced. Definite phase resetting effects occurred for +/-60msec on-beat perturbations in both the 540msec period and the 720msec period conditions, without any subdivisions marked. By contrast, where on-beat perturbations were followed by unperturbed subdivision stimuli, correction responses were reduced to roughly 50% (540msec period) and, for the 750msec period, to the much lower 20%.

Why does the introduction of unperturbed subdivisions so dramatically alter the response in the 720msec period condition, relative to the 540msec period condition? A possible explanation might relate to nested internal clock intervals. Assuming that 270msec is at the lower edge of the range of synchronisation, a 270msec period (half of 540msec) clock would be unavailable, or of weak amplitude. Reinforcement of the 540msec tactus would be correspondingly weak (despite the presence of audible subdivisions in the stimulus), and phase resetting of the 540msec period tactus would be more likely to occur. By contrast, subdivisions at half the period of the 720msec tactus (360msec) would fall within the range of synchronisation. In this case, the 720msec tactus would be reinforced by the subtactus clock at 360msec. The 360msec subtactus clock would itself remain unperturbed because of the unperturbed subdivision stimuli, and despite the perturbation in the stimulus at tactus level.

If nested clocks are responsible for better correction in response to perturbed subdivisions, the response to a perturbation on the last subdivision of three in the 720msec period (ie. a perturbation 240msec before the next beat) should be only slightly less than the response in a bisected 540msec period to a perturbation on the subdivision (270msec before the next beat). In fact, the two responses are close, and well within the deviation as marked by error bars on the graph given in Repp's paper. (Repp, 2001a)

In contrast to theories which propose timers with oscillator properties including phase entrainment, the timing control hypothesis of Semjen and Ivry (Semjen & Ivry, 2001) posits a single timekeeper to provide a base duration and a counter mechanism which generates target timing intervals by counting multiples of the base duration. Subjects in an experiment continued either bimanual or unimanual tapping after synchronising to a metronome. The metronome provided two pulse trains at different pitches, both with 1000msec period and one in a phase relation between 0msec and 900msec to the other. The hypothesis predicts that timing patterns of alternate-hand tapping should be quite similar to those of single-hand tapping. As expected, variance for two-handed and one-handed tapping was nearly identical.

However, the prediction was also borne out that simple subdivision interval ratios, and particular the 1:1 ratio of anti-phase, would show less timing variance than complex interval ratios (Semjen & Ivry, 2001), supporting the idea that consonant clocks reinforce each other to produce better timing. On the assumption that internal clocks become less reliable (perhaps of weaker amplitude) as periods are further removed from an individual's spontaneous tempo, the better timing exhibited for simpler interval ratios could in part reflect poorer timing for smaller intervals towards the limit of the range of entrainment. In fact, a sharp rise in relative variation was found for periods below 300msec. (Semjen & Ivry, 2001) Three-level nesting is also likely to be a factor, given the tactus period of 1000msec: the tendency for 1:4 intervals (at phase delays of 200msec and 800msec) to assimilate to 1:3 intervals is likely due to the influence of a sub-tactus clock with period 500msec, so that the fastest active clock tends to entrain at 250msec, rather than the required 200msec. To achieve an accurate 1:4 ratio would require the 500msec clock (which in this case is not needed for direct timing of any interval) to be turned off. The 200msec clock would then receive reinforcement on one in every five timing intervals, rather than one in two, as is the case for the 250msec clock in this scenario.

Also supporting the timing control hypothesis was the reduction in variability in the full cycle (1000msec) duration when it was produced as the sum of two sub-intervals. Variance was most reduced in the anti-phase condition and some other subdivisions also produced a decrease, although extreme phase conditions actually increased variance of the full cycle. During the synchronisation phase, this increased variance applied only to the intervals between the taps which ended the longer of the two uneven intervals, rather than the taps which began the longer intervals, consistent with hierarchical timing. (Semjen & Ivry, 2001)

The decrease in variance for the full cycle in the anti-phase condition could be accounted for by mutual entrainment of the longer-period and shorter-period clocks (rather than hierarchical control of the faster clock by the slower). Mutual entrainment of the interval start, or clock trigger, would tend to cancel out variances in individual clocks; and the more clocks there were that were mutually-entrained, the stronger this effect would be expected to be. In support of this view, optimal tempo sensitivity for infants, children and adults, whether musically trained or not, is always found at periods between 400msec and 800 msec (Drake, 1998): just within the range at which both a sub-tactus clock (at period 200msec and up) and a super-tactus clock (at period 160msec or less) could be both entrained.

Evidence for mutual entrainment comes from studies of bimanual finger and hand movements. Kelso et al (Kelso et al., 1981) found mutual entrainment effects between the two hands when subjects were asked to move them at the same time. In right-handed subjects the left hand was attracted to the rhythm of the right hand. (cited in (Summers & Burns, 1990))

Yamanishi et al (Yamanishi, Kawato, & Suzuki, 1980) required subjects to maintain different phase differences between the two hands in a tapping task. Unstable performance was found at phase differences other than zero (in-phase) and $180^{o}$ (anti-phase), and a tendency to entrain to the nearest stable phase of $0^{o}$ or $180^{o}$. (Semjen & Ivry, 2001; Summers & Burns, 1990) This finding is consistent with the theory of mutual-entraining tactus and sub-tactus clocks in a binary relation.

Kelso et al (Kelso, 1984; Kelso et al., 1981) found that subjects made an abrupt transition to in-phase motion when asked to increase the cycling rate of anti-phase motion of the hand. The maximum anti-phase cycling frequency (averaged over the five cycles before transition) was 2.26Hz (period = 442.5msec), giving a minimum clock period according to the current theory of approximately 221msec. Frequency tended to increase to 2.50Hz in the transition to in-phase movement, possibly due to the relaxation away from the lower limit of available clock periods.

## 2.4 Timing control below the level of the fastest central clock

How is timing control achieved below the level of the fastest clock – whether tactus or sub-tactus – for periods of 200msec and below? I shall refer to this level as the *tatum* level (Bilmes's (Bilmes, 1993b) coining is a contraction of **t**emporal **atom**, and carries a reference to the massive technique of stride pianist Art Taum). Rhythmic subdivisions are often heard at periods below 200msec – the 16th hi hat pulse carried by drummer Clyde Stubblefield in James Brown's The Funky Drummer (Brown, 1986) for example, has a 149msec mean length; a recording of Ashanti drumming from Baule on the Ivory Coast has 12 micropulses per bell pattern, each 104msec long (Waterman, 1998); and Pressing (Monahan, 1993; Pressing, 1983) gives a range of smallest intervals down to below 100msec for dance drumming among the Ewe people of West Africa.

### 2.4.1 Motor programs

A motor program is a set of muscle commands which are assembled before the movement begins, and allow the whole sequence to be carried out without the necessity of peripheral feedback. (Keele, 1968) Motor programs are therefore especially useful for timeframes which are too short to allow sensory feedback.

A tactus clock is able to be corrected on the basis of the last perceived asynchrony with the stimulus pulse, but movements which are completed in 150msec or less are sequenced too fast for sensory feedback to play a role on the basis of the last movement. The time taken for subjects to make corrective hand movements in a (visual) tracking task was 190-220msec after delayed feedback was made available to them – strikingly close to the lower limit for the range of rhythmic synchronisation. Schmidt (Schmidt, 1975) puts the time to process sensory input into errors and initiate corrections at 120-200msec. Welford (Welford, 1974) gives simple auditory reaction time as about 150msec – 200msec.

Direct evidence for the existence of motor programs includes the success of subjects in an experiment in which they were required to point at the terminal position of a light after it was turned off in a dark room (Keele, 1968).The subjects' head position was kept fixed by a bite board, to minimise kinaesthetic feedback, and the light was moved to a new position either slowly enough to be tracked, or suddenly. If a motor program was invoked in the fast condition to move the eyes to look at the new position, it should be available to direct arm pointing; while bit-by-bit tracking in the slow condition would not be associated with a single motor program to move the whole distance. Subjects were able to point accurately when the light moved fast (but not when it moved slow), indicating the existing of *open loop* control - pre-programmed control in which the action-sensory feedback-modification of movement loop is not closed.

The tatum programs will need to issue commands to control position of limbs, amplitude, and so on. However, for present purposes I will deal with timing control in isolation from the rest of the program. The time for sensory feedback and preparation of a new motor program are not the only considerations for the duration of motor programs controlling tatum timing. Unless relative phase information is hypothesised to be available from the central clocks, timing control from the clocks consists in a trigger signal at the start of the timing interval of the fastest clock. Assuming that relative phase information is unavailable, the motor program which is triggered will have to provide all timing information up to when a new clock trigger is received. (Since we are dealing only with timing information, it is simpler to ignore kinaesthetic feedback for the moment. Schmidt (Schmidt, 1975)) indicates that proprioceptive feedback, which operates very quickly, is monitored during the action to ensure it takes the intended path.)

Where the fastest clock is available at the tactus level – the 600msec tactus subdivided into 4 tatums in the Funky Drummer drum pattern (Freeman & Lacey, 2001) - open loop control of tatum timing must span the 600msec period, that is, the timing of four 16th notes on the hi hat plus whatever kick or snare hits occur off the beat. As the tactus interval progresses, sensory feedback from the beginning of the tactus interval becomes available, and a motor program for the next interval can be prepared.

### 2.4.2 Motor program schema

How is a motor program modified to better match goals? In Schmidt's (Schmidt, 1975) account, as a motor program is run, information relating to instances of the action are stored:

- initial conditions (state of the body and its environment)
- parameters for the motor program, including force and timing information
- sensory and kinaesthetic feedback

- the error between the feedback and the planned goal

The motor program is assumed to be a greater or lesser variation of an action which has already been learnt. A *schema* is the abstraction of a pattern of rules to generate a prototype. As a number of instances of the action have been completed, a schema relating the four different elements given above begins to be abstracted from the stored sets of information. As enough sets of information are stored, and the motor response schema relating the different elements is refined, it becomes possible to successfully generalise the action to slightly changed initial conditions and goals, through the interpolation between parameter specifications. (Schmidt, 1975)

If we confine ourselves to specifying timing information – such as a pattern of temporal intervals which fills the tactus duration – the application of a schema becomes plausible. Assume that the goal is to fill the tactus duration with four evenly-spaced onsets, beginning on the beat, and ending with a fifth onset on the next beat. If entrainment to the current tempo has only recently been achieved and this is perhaps the first time we have played a note in this session (and perhaps we are not very expert), it may be that the motor program we call up is a little too fast. To repeat it correctly, we modify a parameter controlling overall speed, and play it again. Clearly, if we repeat that process many times (as people do when they practice a musical instrument) the speed of that particular program will become finely-tuned to the tempo of the tactus. A similar process would be applied if the intervals were perceived to be uneven when the motor program was executed. Changes are made offline and the program executed again.

Accurate feedback about the success of the executed program is essential for updating the schema. If the motor program is carried out without careful referencing of sensory feedback, the relations between the items of internally-stored information will begin to drift, and no longer specify achievement of the goal. This is a well-known phenomenon of music practice: careless practice is worse than no practice, and can led to deterioration of performance which it takes double the effort to correct. Conversely, precise knowledge of results can be used to refine the schema. The use of a MIDI sequencer to display the timing errors of a passage just played, for instance, leads to quick improvement.

### 2.4.3 Evidence for a different mechanism for timing below the tactus level

A comparison of timing at the beat level with timing at the tatum level has been made by Shaffer et al (Shaffer et al., 1985) in timing data from performance by a concert pianist of Erik Satie's Gnossienne No. 5. While the beat level is very slow (>1000msec), it is marked in the piece clearly and invariably by a bass note, alternating with a chord. The right hand plays a fluid melody, with groupings which (in notation) subdivide the beat evenly at different times into 4, 6, 7 and 8. At other times the RH has no onsets at all for two beats, or plays a variety of uneven rhythms. It was found that groups notated as comprising equal notes, were not in fact played evenly. If the duration of the beats depends on the sum of the durations of RH small notes, then the ratio of the variance of the sum of the intervals making up the beat to the sum of the variance of the intervals will be ≥1. It was found that the ratio was <1, indicating beat independently time from their sub-intervals, and sub-intervals constrained within the beats. The particular piano texture – a steady, even LH with fluidly-timed faster notes in the RH, is characteristic of Chopin – whose instruction to keep the LH steady and vary the RH is well known to piano students. Examples of similar freedom in the RH and clock-controlled rhythm in the LH, were also found in performances of Chopin and Bach. (Shaffer, 1981) Shaffer (Shaffer, 1985) distinguishes between biological clocks and a less-precise temporal pattern generator which is triggered by a clock external to it.

An experiment in judgement and production of beat fractions, relative to an isochronous pulse shows a marked deterioration in accuracy for intervals <200msec. (Sternberg & Knoll, 1994) The poor judgement and production of single intervals in this region seems to reflect a lack of an interval timer, and compared to the ability of musicians at least to produce reasonably even intervals at that level *if they are permitted to play all subdivision onsets* seems to indicate that motor expression assists relative timing at this level.

While runs of small subdivisions of intervals below 200Hz are achieved eg. by many pianists, dividing the intervals in the tatum region into uneven parts appears to be more difficult, with mutual assimilation of longer intervals in rhythms, and the - sometimes unintentional - production of non-harmonic subdivisions of the clock period. (Repp, Windsor, & Desain, 2002) Friberg's (Friberg & Sundström, 2002) analysis of jazz drummer's swing ratios show that complex ratios (involving in this case only two durations) can be reproduced consistently and accurately. Friberg (Friberg & Sundström, 2002) gives the range of ratios of the long note to the short note in jazz swing, at different tempi, as 1:1 to 3.5:1, while analysis of data for ten of the 16 different drummers given in Cholakis (Cholakis, 2003c) shows a variation between 1.5:1 to 2.7:1.

If concatenated, production of even intervals would be produced by an (unconscious) revision of the motor program following a cognitive review of its evenness. The perception of evenness at this level of timing would relate to a direct comparison of the pattern's component intervals to each other, rather than on asynchronies to a temporal grid produced by nested clocks.

Wright (Wright, 2002) notes that skilled drummers can control timing down to a single millisecond – when two drums are played together, such a small timing difference is heard as a timbre change, due to phase interference between the soundwaves produced by each drum. Such a level of control, combined with the non-simple ratios produced at sub-tactus timing levels, suggests a fast clock at ~1000Hz, with a counter mechanism.

### 2.4.4 Functional anticipation

Closely related to the problem of timing motor programs below the level of central clock control, is the phenomenon of synchronisation error (SE) or phase asynchrony in synchronisation tasks.

Experimental studies of synchronisation tapping (Dunlap, 1910; Pressing, 1998a; Repp, 2001a; Semjen, Schulze, & Vorberg, 2000; Thaut, Tian, & Azimi-Sadjadi, 1998; Vos, Mates, & van Kruysbergen, 1995) show that taps tend to anticipate the metronome, with the amount of anticipation increasing with period. In line with Paillard's (Paillard, 1948) view, Pressing (Pressing, 1999) ascribes the consistent small negative phase asynchrony found in human entrainment to compensation for consistent delays in neuroanatomical transmission and processing. Semjen et al (Semjen et al., 2000) raise a difficulty with this interpretation, finding tap-metronome asynchronies in a synchronisation task varied with IOI, whereas anticipation which depended purely on the different delays with which auditory stimuli and motor responses are registered centrally would presumably be nearly constant.

Semjen et al (Semjen et al., 2000) tested timing synchronisation tapping to a perfectly isochronous metronome at a range of periods for adult subjects with at least three years' experience playing a musical instrument. The mean phase difference across all subjects and trials in Semjen et al's experiment (Semjen et al., 2000) was about –10msec for the shortest period of 200msec (where a negative asynchrony means that the tap preceded the metronome), about –15msec for a period of 400msec, and about –30msec for the longest period of 640msec (data in the paper is provided graphically) (Semjen et al., 2000). Thaut et al (Thaut, Tian et al., 1998) found a mean asynchrony of –21msec for tapping to an isochronous pulse with 400msec period, and –30.3msec for a 600msec isochronous pulse.

As Semjen et al (Semjen et al., 2000) note, inter-response interval (IRI) variability is the result, at least in part, of error correction processes. In other words, the perceptual goal is stability of the phase error or asynchrony – also known as the synchronisation error (SE) - and correction processes will orient toward this goal. In an experiment by Repp (Repp, 2001a), auditory feedback from the subject's taps made it more likely that the SE would be reduced so that the sound of the tap was closer to the (antiphase, in Repp's experiment) synchronisation point.

Repp (Repp, 2001a) compared the effect on continuation tapping of a single changed IOI at the end of an isochronous stimulus train, perturbed by a subliminal 10msec in either the positive or negative direction, with the effect of three and five changed IOIs at the end of the sequence. The change was echoed in the immediately following inter-tap intervals (ITI), but for the singled changed IOI, subsequent ITIs returned to the previously established period. In the three-changed and five-changed conditions, ITI intervals gradually adjusted (following the initial echoing of the change), with more persistence of the new period exhibited in the five-changed condition. Repp (Repp, 2001a) attributes the gradual adjustment of ITI to the period correction process, and the initial echoing to phase resetting. Repp suggests that phase resetting may occur as the result of motor persistence competing with sensorimotor coupling, with phase resetting occurring where the tap – and therefore motor persistence - is absent and sensorimotor coupling takes over. Thisd hypothesis is in fact similar to the 'mixed reference' (ie. to either of the stimulus or the subject's own tap) strategy described by Hary and Moore (Hary & Moore, 1987).

A different account (hereafter referred to as the *functional anticipation* theory) might consider the consistent anticipation in terms of the function negative asynchrony of action to stimulus might serve in timing purposeful actions. Hary and Moore (Hary & Moore, 1987) use the term ‘reference interval’ for the delay between the clock trigger and the sensorimotor goal, occupied by neurophysiological transmission time, and possibly other latencies. Synchronisation tapping in experimental settings does not provide a clear perceptual goal for the subject, unless the tap produces audible feedback. If the evolutionary function of entrainment is to allow preparatory anticipation of and actual synchronisation with critical time points relative to processes external to the organism, then a consistent anticipation in the initiation of movements would enable synchronisation of the completion of the movement with external events. It would be necessary for the amount of anticipation to be variable to meet different the requirements of different motor-sensory goals; however, once the amount of anticipation was established for a particular goal in the current context, it would be necessary for it to be consistent within the tolerance required in order to achieve the goal. This ability would be required in actions such as catching, throwing to hit a target, striking a target, speaking in turn, or striking a drum in time with others. Musicians routinely compensate for the different amount of time to make a violin, piano, trumpet, voice, trombone or drum speak, so that the perceptual onset will be in time with the beat and with other instruments. A keyboard player, for instance, can adapt to different electronic instruments with different latencies – so long as the latency is consistent for each instrument. Wright (Wright, 2002) discusses the problem of latency for electronic instruments, and proposes 10msec as a good goal. Wright notes that, for timing differences down to a single millisecond to be achieved by skilled drummers, latency *jitter* - or variation - must be no more than 1msec. (Wright, 2002)

Musicians also routinely maintain consistent functional anticipation after having not played on one or more beats for musical reasons. In the experimental setting, where no prediction of a sensorimotor consequence of the action has been established - either via auditory feedback or the achievement of some physical goal - phase resetting following a missed tap (as found in (Repp, 2001a)) may simply indicate the absence of a clear perceptual goal for the subject. It is apparent that functional anticipation utilises absolute, not relative, timing, since the time required to make eg. a brass instrument speak, and the latency of an electronic keyboard, will not scale with tempo.

Synchronisation tapping experiments by Aschersleben and others where auditory feedback was introduced each time the subject touched the key showed a reduced asynchrony, (Aschersleben & Prinz, 1995; Mates, Radil, & Pöppel, 1992; O'Boyle & Clarke, 1996) supporting the theory that where a sensorimotor goal is set up, taps are timed to achieve the goal. In studies where a small delay (<100msec) was introduced between the touch on the key and the auditory feedback, (Aschersleben & Prinz, 1997; Mates & Aschersleben, 2000) an increase in asynchrony was observed with increasing delay, consistent with the functional anticipation theory.

The observable synchronisation error, in this view, is one component in the total *functional anticipation* amount, which also includes the motor delay which is modelled in the classic two-process model of repetitively-timed movements of Wing and Kristofferson. (Wing & Kristofferson, 1973a) To make explicit the whole chain from its source, the timed clock trigger travels through neural pathways to signal the start of the motor program, which is then executed so that a sensorimotor–temporal goal is achieved (such as the synchronisation of a drum sound with another onset.) The total functional anticipation is the time from clock trigger to the critical moment of synchrony with the external event.

Apart from the necessary and predictable delay associated with the transmission of the command actual physical execution of the movement, other factors may become part of the total delay between the clock trigger and the sensorimotor goal. For example, any latency inherent in the mechanism of a musical instrument will contribute to the delay – and therefore to the required degree of *functional anticipation,* so that the clock trigger would need to fire that much earlier at its central source.

If perhaps the muscles do not respond quite as predicted to the command, or some other variability is introduced such as a timing inaccuracy due to a change in the external goal, this will need to be accommodated by a correction to the amount of functional anticipation (equivalent to a phase correction in the current model). If a negative correction is required (if the process is to be completed a little sooner after the clock trigger time) it will be useful for the chain to have included a small amount of gratuitous delay, part of which can be taken up to compensate for the timing shift in the stimulus. The delays involved will be of the order of the synchronisation error observed in tapping tasks, ie. less than 100msec and probably closer to 50msec, but it will be essential for accurate motor timing (to a 1msec resolution, in an expert musician – see above) for functional anticipation to exhibit jitter (variability) of ≤1msec.

### 2.4.5 Possible functions of microtiming deviations in a musical context

A more complicated situation arises where a group of drummers is playing together. If the speed of sound in air is approximately 343metres/sec, the sound of a drum stroke will reach the player's ear, about .5m away, in approximately 1.5msec. If each drummer tried to adjust phase or period of their own drumming to their perception of the other's stroke timing, the result would be a gradual slackening of tempo. The sound from a drum played by another person 4m away which reach the first player's ear with a delay of ~12msec, and the second drummer will perceive the first drummer's strokes with the same delay – so there will be approximately an 11.5msec discrepancy in the perception each drummer has of a stroke by either. Were there are more than two drummers, necessarily at different distances from each other, the discrepancy between any given pair of drummers may be similar or quite different from that between any other given pair. In this musical context, a practice of varying synchronisation error relative to an external stimulus provided by one or more other musicians, so that both negative and positive microtiming deviations are produced, would help to maintain the stability of the global tempo over time for the whole group.

# Chapter 3 The macroperiod in groove music, and the psychological present

Chapter 2 described short motor programs to control timing via open loop control, with planning – construction of the motor program schema – occurring just in time for the execution of each motor program. This chapter reviews the evidence for *hierarchical* planning and execution of timed actions. It is proposed that the macroperiod of groove musics establishes a common psychological present, for which hierarchical planning and execution is carried out.

Further sections of the chapter describe various rhythmic and metrical devices in traditional African drumming, in the context of their function in defining the start and duration of a macroperiod, including cross meter, asymmetrical bisection of the macroperiod, and bell patterns or 'timelines'. The clavé pattern found in Afro-Cuban music is discussed as serving the same function as the African bell pattern. Finally, the tal of North Indian classical music is discussed in terms of defining the start and duration of the macroperiod.

## 3.1 Hierarchical planning of movement sequences

Evidence for motor planning segmentation in music comes from errors which suggest that musical sequences are partitioned during planning into phrase segments – errors replacing intended pitches in piano performance are more likely to originate from the same phrase as the intended event, than from different phrases. (Palmer, 1997; Palmer & Pfordresher) Larger serial distances for planning were found more often in the absence of musical phrase boundaries. (Palmer & Pfordresher) In speech, preparatory reaction time following a signal to respond depends on the number of syllables per word and the number of words that are to follow, as though utterances are preplanned. (Klapp, Anderson, & Berrian, 1973; Sternberg, Monsell, Knoll, & Wright, 1978)

Rosenbaum (Rosenbaum, 1985) asserted that if timing of actions is specified, their serial order is necessarily determined. However, as MacKay (MacKay, 1987) points out, proficient speakers never simply misorder components in time, as would occur with timing errors under this model. Instead, components exchanging places are nearly always members of the same sequential class, such as the initial consonants, in the error 'cake the ring of teas' in stead of 'take the ring of keys'.

In a review of motor program theory, Summers & Burns (Summers & Burns, 1990) note that it has developed to include the idea of hierarchic memory structure in which an abstract representation of action is elaborated into its more specific components as information descends the hierarchy. Speech errors where sounds are switched exchange consonants only with consonants, and vowels only with vowels; at a higher level of linguistic analysis, nouns usually exchange only with other nouns and verbs tend to exchange with other verbs. (Rosenbaum, Inhoff, & Gordon, 1984) Rosenbaum et al (Rosenbaum, Kenny, & Derr, 1983a) found interresponse time and errors in a rapid keypress task supported execution dependent on hierarchical organization. Evidence for hierarchical storage and retrieval of sequence content also comes from intertap pause structure in patterned sequences of finger taps (Povel & Collard, 1982); from analysis of reaction time (Rosenbaum, Kenny, & Derr, 1983b); and from analysis of errors when either subsequences of spoken nonsense syllables, or the syllable order within the subsequences, were reordered. (Gordon & Meyer, 1987)

### 3.1.1 Hierarchical meter and timing

Summers & Burns (Summers & Burns, 1990) relate hierarchic memory structure in motor program theory to hierarchical models of rhythm perception proposed by Povel (Povel, 1981) and Povel & Essens (Povel & Essens, 1985), as does Palmer. (Palmer, 1997) Recent experiments by Palmer & Pfordresher (Palmer & Pfordresher, 2003) support linking of an hierarchic memory structure to hierarchic metre: events intended for strong metrical positions were more likely to be confused in performance with an event on a strong metrical position, while events intended for a weak metrical position were more likely to be confused with an event on a weak metrical position in performance. Variance in timing is (at least in some contexts) controlled at different levels in a hierarchy. Shaffer et al's (Shaffer et al., 1985) study of expressive timing in repeated performances by one pianist of an Erik Satie piece found expressive deviation at the beat level, independent of expressive deviation of subdivisions of the beat.

## 3.2 The psychological present

Music – especially music which is markedly rhythmic – often falls into cycles of up to 3000ms. African bell timelines, and the *insiraf* rhythm of the music of the Magreb, have been noted as occupying periods up to 3000ms. (Elsner, 1990; Vos,

1976) (see also the discussion of macroperiod duration in North Indian classical music and traditional African music below.)

Seifert et al cite Pöppel (Pöppel, 1983) who defined the “subjective present” as the temporal integrating level where temporal gestalt perception appears. Two auditory stimulus events presented with a delay of more than 3000ms will not be perceived as grouped together. (Seifert, Olk, & Schneider, 1995) Turner and Pöppel state that a human speaker will pause for a few milliseconds at about three second intervals, in order to determine syntax and wording of the next three seconds. They identify this period as the “specious present”. They identify the average length of a syllable as about 1/3 of a second (~300ms) and the average number of syllables per line in poetry as about 10 – making the average line of poetry within the (approximately) three second “auditory present”. The line also usually constitutes a rhythmic, semantic and syntactical unit.

Turner and Pöppel (Turner & Pöppel, 1988) assert that meter in metered poetry – the arrangement of long and short syllables in accordance with a regular pulse – functions to synchronise the hearers with the speaker, “so that each person’s three-second present is in phase with the others”. Similarly, the start of the macroperiod in groove musics is always marked in a musically audible way, and sometimes visually – see discussion at **3.3.5,** below.

Observation and analysis by Richman (Richman, 2000) of gelada monkey vocalisations, and comparison with humans, found in both cases that friendly vocalizing was produced in units averaging a total length of about 9 or 10 syllables, produced at a rate of about 5 syllables a second. (Richman, 2000) A 10-syllable utterance at a 5-syllable-per-second rate would last about 2 seconds – within the duration of the psychological present.

Human speakers, independently of language or age, tend to construct closed verbal utterance up to about 2 to 3 seconds. (Kowal, O'Connell, & Sabin, 1975; Pöppel, 1988; Vollrath, Kazenwadel, & Krüger, 1992) The duration of spontaneous intentional acts falls in a temporal window of approximately 3 seconds. (Pöppel, 1996; Schleidt, Eibl-Eibesfeldt, & Pöppel, 1987) Pöppel (Pöppel, 1996) concludes that there is a general temporal segmentation mechanism, and he terms the single states of 3-second segmentation “states of being conscious” (STOBCON). Each of Pöppel’s STOBCONs represents ”a mental island of activity distinctly separated from the temporally neighbouring ones” - continuity of experience is an illusion formed by linking the contents of each temporal window to the one that follows. (Pöppel, 1996)

As argued at **3.3.6** below, the *clavé*, the West African bell patterns, and the *tal* all seem designed to synchronise the phase of the psychological present for musicians and listeners alike, in a similar way to the universal 10-syllable line of metered verse noted by Turner and Pöppel above, and to Pöppel's STOBCONs.

### 3.2.1 Rhythm, the body, and the psychological present

Todd (Todd, 1999), with others (Todd, 1992a,c; Todd, O'Boyle, & Lee, 1999), has elaborated a theory of bodily motion underlying musical rhythm. A fundamental notion motivating Todd's explorations is the musician's concept of the phrase as a 'single, continuous gestural form', which organises discrete events into a structural gestalt. (Todd, 1999) Todd's 'single, continuous gestural form' creating a structural gestalt at the musical phrase level corresponds, in the current theory, to the broad planning level of the psychological present. In the theory, the temporal trajectory of the phrase (and non-temporal aspects, such as loudness dynamics) are planned before the phrase begins. Physical implementation is then carried out through motor programs, which are planned individually immediately prior to execution. Todd (Todd, 1999) is not explicit about how long is required to plan each motor program, but it may be assumed at least that it is a shorter time than is required to execute the preceding program since otherwise the planning for two programs would overlap. In the theory proposed in this thesis, motor programs are triggered at the tactus level of timing, ie within approximately 250msec – 1200msec. The question of how far in advance motor programs can be parameterised and stored for execution seems to be so far unanswered.

It seems probable that the phrase or macroperiod level would also be when decisions are made as to which rhythm patterns will be played (whether these are improvised, or 'read into' working memory from a score or long term memory). Presumably, the start and duration of motor programs would be decided at the level of Todd's gesture, or the level of the macroperiod in groove music.

## 3.3 Period segmentation in various groove-like musics

Anku (Anku, 2000) describes African dance drumming as depending on a circular concept of time which defines a structural set from with the entire performance is derived – both the cyclical nature of the rhythm and the constrained variation are characteristic of groove.

### 3.3.1 Cross meters

The pulsation (tactus) in Central African music is organised into periods, usually an even number, from 2 up to 12 and sometimes more. (Anku, 2000; Arom, 1991) Arom (Arom, 1991,1994) reports finding instances of polyrhythm in traditional African drumming music, with two simultaneous meters made up of different groupings of subdivisions in ratios such as 2:1, 3:1, 2:3. 3:4, and multiples of these. In a 3:4 ratio of subdivision groupings, for instance, a total of 12 subdivisions within the macroperiod would be divided (in one of the two simultaneous meters) into 3 groups of 4; and in the other, into 4 groups of 3. The groupings are marked by repeating patterns of inter-onset intervals (IOIs), so that each grouping is marked by an audible drum stroke at its start, but not all the subdivisions with the grouping are marked. In addition, either or both of accents or tone colour changes may be added. (Arom, 1991)

Since the main beat or tactus is not always clearly evident to a listener not used to the cultural and musical traditions of African music, the issue of whether a different tactus pulse may be held by different performers has been discussed at different times in the ethnomusicological literature. Arom notes that the cross meters are in fact hemiolas, which he defines as ‘the repetition, within a single period, of an identical configuration *at different positions with respect to the pulsation*.’ (Arom, 1994) [my emphasis] In other words, as with other authors (Hodges, 1992; Iyer, 1998) and practitioners of African drumming (Ladzekpo, 1995), Arom‘s account allows for only one main tactus pulse to be active. Although metrical ambiguity may be a feature and the main pulse may in fact not be strongly marked, drummers and dancers are still expected to maintain a sense of the main pulse at all times. (Ladzekpo, 1995)

Two kinds of cross-rhythms are found in traditional African dance drumming: those which employ a different grouping of tactus subdivisions from the main beat, as in a 6-beat cross meter to 4 main beats with triplet subdivisions (as in Figure 3-1); and those which subdivide the tactus in a different way, as in an 8-beat cross rhythm to 4 main beats with triplet subdivisions (as in Figure 3-1). (Anku, 2000; Ladzekpo, 1995)

Subdividing the tactus in a different way simply reinforces the tactus.

*Figure 3-1 An 8-beat cross rhythm to 4 main tactus beats is a different subdivision of the tactus*

| | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Main beat (4) | **\|** | **-** | | **-** | **\|** | **-** | | **-** | **\|** | **-** | | **-** | **\|** | **-** | | **-** |
| Crossrhythm (8) | **\|** | | **-** | | **\|** | | **-** | | **\|** | | **-** | | **\|** | | **-** | |
| Subdivisions (12) | **x** | **x** | | **x** | **x** | **x** | | **x** | **x** | **x** | | **x** | **x** | **x** | | **x** |

**|** = location of main and cross rhythm beats

**-** = location of subdivisions for main and cross rhythms

**x** = location of subdivisions for main rhythm

*Figure 3-2 A 6-beat cross meter to 4 main tactus beats is a different grouping of subdivisions*

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Main beat (4) | **\|** | **-** | **-** | **\|** | **-** | **-** | **\|** | **-** | **-** | **\|** | **-** | **-** |
| Crossrhythm (6) | \| | - | \| | - | \| | - | \| | - | \| | - | \| | - |
| Subdivisions (12) | x | x | x | x | x | x | x | x | x | x | x | x |

**|** = location of main and cross rhythm beats

**-** = location of subdivisions for main and cross rhythms

**x** = location of subdivisions for main rhythm

One of the effects of different tactus subdivision groupings is to define a macroperiod, which begins where both the main beat and the cross rhythm fall on the same subdivision (Figure 3-2), fulfilling one of the points of the definition of groove given above – identification of cycle locations.

### 3.3.2 Asymmetrical bisection of the macroperiod.

A notable device for producing asymmetry consists in dividing an even number of subdivisions into a section containing one subdivision less than half the total; and a section containing one subdivision more than half the total. For example, a figure with four pulsations divided into 16 subdivisions would be ordered by accent, the pattern of tone colour, or IOI pattern into a section (8 – 1 =) 7 subdivisions long plus a section (8 + 1 =) 9 subdivisions long. (Arom, 1991,1994) The asymmetrical periods produced in this way will not coincide at the bisection with any other sub-period of the total. As in the case of cross meters, the start of the macroperiod is defined by the coincidence of the downbeat of (otherwise overlapping) periods.

Arom (Arom, 1994) cites an example of this device in a piece from the repertoire of the Aka people of Central Africa. The piece involves 4-part vocal polyphony, plus three layers of drums, and a percussion part played on iron blades – 8 layers in all. One of the drum parts repeats a 3-pulse ostinato throughout, the smallest period in the piece. (Pulses here are evidently not tactus beats, but a subdivision). A second drum part defines 12-pulse periods through joining three 4-pulse periods together. The third drum also defines a 12-pulse period, using asymmetrical bisection into 7 + 5. The voices define what Arom calls the 'song period' of 36 pulses, divided into for each voice into two periods of 12 + 24; 24 + 12; 8 +28; and 16 +20. All parts coincide at the beginning of the 36-pulse song period, except the iron blades part. This part asymmetrically divides a 24-pulse period into 11 + 13; three of these periods span two of the 36-pulse song period to make the 72-pulse macroperiod. (Figure 3-3)

*Figure 3-3 Period layers in a song from the Aka people, from a figure provided in (Arom, 1994)*

<table>
<tr><td colspan="16">16</td><td colspan="20">20</td><td colspan="16">16</td><td colspan="20">20</td></tr>
<tr><td colspan="8">8</td><td colspan="28">28</td><td colspan="8">8</td><td colspan="28">28</td></tr>
<tr><td colspan="24">24</td><td colspan="12">12</td><td colspan="24">24</td><td colspan="12">12</td></tr>
<tr><td colspan="12">12</td><td colspan="24">24</td><td colspan="12">12</td><td colspan="24">24</td></tr>
<tr><td colspan="12">7 + 5</td><td colspan="12">7 + 5</td><td colspan="12">7 + 5</td><td colspan="12">7 + 5</td><td colspan="12">7 + 5</td><td colspan="12">7 + 5</td></tr>
<tr><td colspan="12">4 + 4 + 4</td><td colspan="12">4 + 4 + 4</td><td colspan="12">4 + 4 + 4</td><td colspan="12">4 + 4 + 4</td><td colspan="12">4 + 4 + 4</td><td colspan="12">4 + 4 + 4</td></tr>
<tr><td colspan="3">3</td><td colspan="3">3</td><td colspan="3">3</td><td colspan="3">3</td><td colspan="3">3</td><td colspan="3">3</td><td colspan="3">3</td><td colspan="3">3</td><td colspan="3">3</td><td colspan="3">3</td><td colspan="3">3</td><td colspan="3">3</td><td colspan="3">3</td><td colspan="3">3</td><td colspan="3">3</td><td colspan="3">3</td><td colspan="3">3</td><td colspan="3">3</td><td colspan="3">3</td><td colspan="3">3</td><td colspan="3">3</td><td colspan="3">3</td><td colspan="3">3</td><td colspan="3">3</td></tr>
<tr><td colspan="24">11 +13</td><td colspan="24">11 + 13</td><td colspan="24">11 + 13</td></tr>
</table>

Inspection of Figure 3-3 reveals what is evidently a carefully integrated system of periods to allow the different parts to support each other rhythmically, while systematically developing longer period asymmetries. The basic 3-pulse is presumably the tactus; the 4-pulse cross-rhythm coincides with the tactus at every 12 pulses and supports the 12-pulse (7 + 5) drum part shown immediately above it, the fourth voice part (12 + 24) and the third voice part (24 + 12); and the first period of the first and second voice parts, which are 8 and 16 respectively, both multiples of 4. The third voice part (24 + 12) and the fourth voice part (12 + 24) each at different times support the 24-pulse (11 + 13) period of the iron blades (shown at the bottom of the figure), while conflicting with its asymmetry elsewhere in the 72-pulse macroperiod. The asymmetrically bisected parts (7 + 5 and 11 +

13) produce prime number periods which therefore do not accommodate nesting of the 3-pulse and 4-pulse periods. The rigorous interlocking of the whole leaves little doubt that defining the macroperiod is a principal focus of the song.

### 3.3.3 Bell patterns in Central and West African drumming, and other reference patterns

The macroperiod in Central and West African drumming is often defined by a recurrent bell pattern. (Anku, 2000; Hodges, 1992; Seifert et al., 1995) Three patterns from different cultures are show in Figure 3-4.

Each of the patterns shown in Figure 3-4 has the same number of fastest pulses (12 in each case), and the same number of realised beats (7), which are organised into the same pattern of distances between onsets – so there is rotational symmetry between the three examples. For instance, rotating the pattern used by the Anlo Ewe people of Ghana two pulses to the left makes it the same as the pattern shown for the Bemba people of Central Africa. In the terminology of the Derler system (Seifert et al., 1995) for classifying rhythmic patterns, they are members of the same equivalence class. The only difference between them is their relation to the tactus (shown beneath the bell rhythm), and the point in the pattern at which they start.

The bell patterns shown in Figure 3-4 exhibit no rotational symmetry within themselves individually, except at the level of the whole pattern. If a particular pattern is rotated through any number of pulses not a multiple of 12, some onsets and unrealised pulses within the rotated pattern will inevitably not match those in the original.

Only one element of the equivalence set from the appropriate rhythm set is used by a particular culture, while other cultures will utilise a different element from the same equivalence set. Once a starting point for the pattern is fixed (by cultural convention), the expert listener (one familiar with the musical conventions of the particular culture) can always unambiguously identify the starting point of the whole pattern – the macroperiod – from hearing the realised pattern.

*Figure 3-4 Drawn from figures given in (Anku, 2000).*

<table>
<tr><th>Region</th><th>Culture</th><th>Bell pattern<br>(equivalence class element)</th></tr>
<tr><td>Ghana</td><td>Anlo Ewe people</td><td>x . x . x x . x . x . x<br>|     |     |     |</td></tr>
<tr><td>Central Africa</td><td>Bemba people</td><td>x . x x . x . x . x x .<br>|     |     |     |</td></tr>
<tr><td>Nigeria</td><td>Yoruba people</td><td>x . x . x . x x . x . x<br>|     |     |     |</td></tr>
</table>

### 3.3.4 The clavé pattern in Afro-Cuban music

The standard clavé pattern of Afro-Cuban music functions in a similar way to the African bell pattern. Once again, the timbre of the instrument chosen (in this case, the clavé) has a sharp attack, a quick decay, and a penetrating frequency. The clavé pattern (shown in Figure 3-5), like the West African bell patterns, exhibits no rotational symmetry except at the level of the whole pattern. It effectively defines the start, progress and duration of a macroperiod, similarly to the bell patterns. (Figure 3-5)

*Figure 3-5 Cuban clavé pattern*

<table>
<tr><td></td><td>x . . x . . x . . . x . x . . .</td></tr>
<tr><td>tactus</td><td>|       |       |       |</td></tr>
</table>

### 3.3.5 The *tal* in North Indian music

According to Clayton (Clayton, 1997) the *tal* in North Indian art music is an hierarchical meter with *matra* (beat), *vibhag* (section) and *avart* (cycle) levels. The 20 in common use each have an avart or cycle composed of between 6 and 16 *matras* or beats. The beats are collected into at least two *vibhags*, or sections, which may not be of equal length. (Clayton, 1997) For instance, the example of *jhaptal* given by Clayton is shown in Figure 3-6.

*Figure 3-6 North Indian* jhaptal *metric pattern, from a figure in (Clayton, 1997) (**+** = clap; **o** = wave)*

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | *avart* |
| **+** | | **+** | | | **o** | | **+** | | | *vibhag,* marked by clap pattern |
| **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | *matra* |

In *jhaptal,* the four vibhags (sections) are either two beats or three beats long, and the 2 + 3 + 2 + 3 pattern divides the *avart* (cycle) into two equal parts. However, the start of each *vibhag* is marked with either a clap (shown in Figure 3-6 as **+**) or a wave (**o**), with the first given a clap and the third (otherwise indistinguishable) section given a wave. The *avart* as a whole is thus asymmetrical, and musicians and musically-educated listeners know where the first beat lies. In addition to the clap pattern, the *tal* has a *theka* or drum pattern, expressed through different strokes of the drum. (Clayton, 1997)

The *tal* is always explicitly performed in North Indian music; in addition, singers and members of the audience count out the *tal*, and musicians and audience may also perform the hand pattern. There are thus always audible and visible clues ensure neither the performer nor the listener can misinterpret the *tal*, and it is 'an idea alien to Indian musical thought' (Clayton, 1997) that listeners might find two or more metric interpretations of a single piece equally valid. Once, again cultural familiarity is required to properly interpret the auditory aspect of the music: for instance, there is nothing to prevent an uneducated listener hearing the start of the second *vibagh* in *jhaptal* as the beginning of the cycle. Given this familiarity, the start and the the duration of the macroperiod is clearly defined for both performer and listeners.

### 3.3.6 Duration of the macroperiod interval

How long are the macroperiods defined by these different devices? According to Clayton (Clayton, 1997), the North Indian *tal* is actually played at a wide range of tempi which produce cycle length ranging from about 1.5 seconds to more than a minute. However, Clayton suggests that the influence of explicit theory in the teaching of North Indian music may be responsible for the very long tal, and points out that at very slow tempi, musicians make use of smaller divisions of the tal cycle. (Clayton, 1997) Arom (Arom, 1991,1994) gives two- and three-part examples drawn from Aka pygmy music of macroperiods four pulses long, and three-part examples which are eight pulses long. Arom (Arom, 1991) does not indicate tempi for the examples given: at a very fast tempo of 180bpm, a 4-beat macroperiod would last ~1.3sec and the 8-beat macroperiods, ~2.6sec, with slower tempi giving a proportionately longer macroperiod. The typical African bell pattern described in (Anku, 2000; Hodges, 1992; Seifert et al., 1995) is 12 pulses long – at a very fast tempo, with each pulse of period 100ms, the macroperiod would last 1200ms. The 72-pulse macroperiod described by Arom (Arom, 1994) and shown in Figure 3-3 would last for 7200ms, with a very fast smallest pulse of period 100ms. However, 7 of the 8 cross-rhythm layers in the pattern coincide at the half-period, which (at the same tempo) would have a duration of 3600ms.

The range of macroperiod duration in North Indian classical music and traditional African music can therefore be assumed to fall between a lower bound of ~1200ms and an indeterminate upper bound. However, many examples have a macroperiod of duration up to ~3000ms, and may be considered to be within the possible duration of the psychological present. In musical contexts where macroperiods are longer, performers probably focus on some regular subdivision of the whole macroperiod.

These findings are strikingly similar to average bar durations for JS Bach’s Well Tempered Clavier, given in (Vos, 1976). Vos gives the average bar durations (drawn from recorded performances) for different meters as: 1.75sec for 2/4 meter; 3.0sec for 3/4 meter; and 4.8sec for 6/4 meter (Vos, 1976)

.

## 3.4 Summary

Hierarchical planning for rhythm performance takes place on at least two temporal levels. At the lowest level, schemata for individual motor programs are generated just in time for the execution of the motor program. At a higher level, corresponding to the macroperiod in groove music, and to the phrase in other musics, the overall choice of timing trajectory is made. Decisions regarding the actual rhythm patterns to be played, and the start and duration of motor programs to express those patterns, are probably also made at the higher level.

# Chapter 4 The Covert Clock Theory

Various authors have undertaken studies of expressive timing variation in Western notated art music. Typically, this involves both local tempo changes – where the tactus rate is varied relative to the average or base tempo of the music - as well as systematic variation of tactus subdivision durations that do not affect tempo. (Beran & Mazzola, 2000; Clynes, 1983,1987,1995; Clynes & Walker, 1982; Gabrielsson, 1982; Rasch, 1988; Repp, 1990,1998; Shaffer et al., 1985; Stange-Elbe & Mazzola, 1998; Timmers et al., 2000) If motor programs are triggered by the tactus clock, expressive variation of tactus subdivision onset times can be freely undertaken without affecting the tempo (which is equivalent to the tactus rate), by means of parameterisation of motor programs.

Less work appears to have been undertaken on systematic timing deviation in groove music, where tempo, if not precisely invariant, does not vary locally, and I address this lacuna.

## 4.1 Review of studies of microtiming deviation in groove music

Several studies of microtiming deviation in groove musics have argued that the deviations are an essential component of the groove, rather than a product of poor timing by performers. The main argument for the musical, rather than random, nature of microtiming deviations rests on the consistency with which deviation patterns are repeated in the same or similar musical contexts. (Several analyses of data provided in these studies are undertaken in Chapter 5 of this thesis.)

### 4.1.1 Consistency of microtiming deviation patterns

Alén (Alén, 1995) found deviation patterns in *Tumba Francesca* music of Cuba, which were consistent for each drum part to each pattern in a dance; but which varied between drums in the pattern, and between patterns.

Freeman & Lacey’s work (Freeman & Lacey, 2002) on jazz swing in recordings by individual performers has identified consistent delays on specific beats of the bar that fall within 30ms windows. Their analysis of The Funky Drummer break (Freeman & Lacey, 2001) by Clyde Stubblefield (Brown, 1986) also shows consistent microtiming deviations

### 4.1.2 Recreating and applying microtiming deviation patterns

Bilmes (Bilmes, 1993a,1993b) applied computer beat extraction techniques to a multitrack recording of a Cuban drum ensemble, and found the variance for attacks of three of the drums, in relation to a reference timing pattern. Analysis of the deviations showed repeating patterns of deviation, mainly in a range of approximately 20ms-80ms ahead or behind the tactus subdivision. Experiments in which resynthesised recordings were played to subjects, found listeners preferred the recording which had been quantised and then had recorded deviations added to it, over both the quantised track (which they found mechanical), and the quantised track with random deviations added to each attack time (which they characterised as “sloppy” or “random”). (Bilmes, 1993a,1993b)

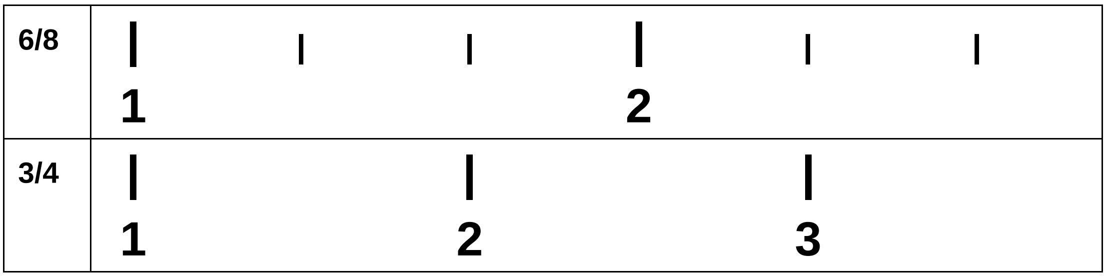
At least as convincing of the musical import of microtiming deviations as the studies above, are Cholakis’s DNA Beat Templates (Cholakis, 1999a,b) - because of their commercial viability. The Beat Templates are files which can be loaded into a MIDI sequencer and which enable quantisation of rhythm patterns in the sequencer, according to typical deviation patterns as played by a famous drummer (eg. Bernard Purdie, or Clyde Stubblefield). Each commercially available CD of Beat Templates files provides a variety of feels, each with its own characteristic microtiming deviation, which Cholakis extracts from audio files via his own beat detection software. (Cholakis, 2003a)

Although these studies are limited in number, they provide definite evidence that microtiming deviations are a musical phenomenon, and at some level under the control of the performer (although the control may not be conscious). The generation of microtiming deviations by Bilmes (Bilmes, 1992,1993a,1993b) and Cholakis (Cholakis, 1999b) have focussed on limited variation of deviation patterns obtained through analysis of audio recordings. The *Covert Clock* theory of microtiming deviations represents an attempt to construct a more holistic method of generating deviation patterns, based on the musical and psychological processes described in previous chapters.

## 4.2 The Covert Clock theory of microtiming deviations

I propose a more general theory of interlocking clocks, of which the fully-consonant metrical hierarchy of harmonically-nested clocks described in Chapter 2 above constitutes a special case. Consider a 3/4 hemiola over a 6/8 meter (see Figure 4-1). The two meters are consonant at the bar level and share subdivisions (which may be below the range of synchronisation). The hierarchical *consonance* of harmonically nested clocks is absent – the two clocks of this much more ambiguous meter are interlocking, rather than nested. The hemiola device can be extended to any meter which allows a different grouping of lowest-level subdivisions to produce a cross-rhythm at a higher level. For example, the 3:4 hemiola over 12 subdivisions is common in traditional African music. (Anku, 2000; Arom, 1991,1994; Ladzekpo, 1995; Magill & Pressing, 1997)

*Figure 4-1 3/4 hemiola against a 6/8 meter*



Note that a hemiola meter is ambiguous only in that there is a choice to be made by each individual as to which clock constitutes the tactus – it is generally agreed that only one clock can be perceived as the tactus, the beat to which physical movement is synchronised, at one time. Parameterisation of models to fit to data from two-handed hemiola rhythms indicates that, while there may be two intermittently-consonant clocks separately driving the two hands, one clock is always dominant. (Pressing, Summers, & Magill, 1996) I will term the non-dominant clock the *Covert Clock.*

Summers et al (Summers, Rosebaum, Burns, & Ford, 1993) concluded from error rates and timing variability of hemiolas and other polyrhythms (eg. 5:2, 5:3 and 5:4, where the two isoschronous streams do not share subdivisions) that a strategy of timing the fast hand to the internal clock trigger and the slow hand to appropriate delays (not matched to any central clock trigger) following the triggers was used by both musicians and non-musicans. (Note that the tempi specified in the text are too fast to allow the non-dominant or covert clock to synchronise to a sub-tactus clock belonging to the main tactus.)This finding was confirmed by Pressing et al (Pressing et al., 1996) for experienced performers with training in jazz and West African percussion ensemble music.

However, these studies and the one by Magill and Pressing (Magill & Pressing, 1997) with a master Asante drummer from Ghana cannot be taken as definitive of timing processes for polyrhythms in an actual musical setting. Ensemble playing, with musical interaction, and polyrhythms and other parts distributed among the players, is certainly a different context from laboratory trials; repetitions of the cycle will probably exceed the 40 trials per block used by Magill and Pressing; and Magill and Pressing were obliged to repeatedly request their subject not to play in his normal fashion, with added flams and musical tempo variation.

### 4.2.1 Contrasting properties of cross-rhythms and hierachical meter

Whatever the timing control process used to achieve it, hemiola and cross-rhythms have these properties in contrast with hierarchical meter depending on harmonically-nested clocks:

- clock dissonance between the tactus clock and the covert clock within the macroperiod boundary (although clock consonance occurs at the boundary, as with harmonically-nested clocks);
- no central-clock-timed temporal grid is marked below the tactus level. In fact, a triplet subdivision of a moderate tactus can be viewed as (in part) a strategy for avoiding clocks below the tactus level – a binary subdivision of a 600ms tactus period produces a clockable interval (300ms) but a ternary subdivision (200ms) does not;
- maintenance of a covert clock at a cross-rhythm to the main or apparent tactus will presumably conflict with and so prevent the activation of a super-tactus clock at the next hierarchical level;
- therefore the *only* consonance between the tactus clock and the covert clock will occur at macroperiod boundaries; and
- meter of this kind is therefore not hierarchical, at least in the sense of levels of salience graded according to the importance of the beat to the bar by the aggregate consonance of clocks in harmonic period ratios.

Thus, hemiola and cross-rhythms can be seen as serving to break down a strong sense of metrical grid to create a more fluid space within the boundaries of the macroperiod. Other rhythmic devices, discussed in Chapter 3, produce the same effect.

The Covert Clock Theory posits that the relative freedom from entrainment clocks thus obtained within the duration of the macroperiod allows for the development of motor programs of varying length over successive iterations of the macroperiod. Thus, motor programs in successive iterations of the macroperiod will not necessarily begin or end on the same tatum. Motor programs provide timing control below the range of synchronisation. The covert clock – whether in actual musical practice it actually functions as a central clock in opposition to the tactus clock, or is always timed in terms of delays relative to the tactus clock – provides a timing reference by which motor program triggers can be shifted from exact concurrence with the tactus clock. Shifting some of the trigger points in time means that some motor programs will need to be spread over a slightly longer period – in other words, the local tempo will vary, producing microtiming deviations, while the global tempo over the macroperiod remains stable (since the tactus clock itself was not affected).

The Covert Clock Theory does not require that the trigger timing point be shifted all the way to the point where it would coincide with precisely-timed covert clock timing points: instead, the tactus timing point and the covert clock timing points constitute the boundaries of a temporal space within which the musician can flexibly choose to place the next trigger point. A choice of trigger point very close to the tactus timing point amounts to no deviation, and it is presumed that this option is available to the performer and may be applied intermittently. This would account for both the range of variation in microtiming deviations found in some iterations of The Funky Drummer break (see analysis below), and through stability of motor program schemata, the consistency found between other iterations.

### 4.2.2 Some predictions of the Covert Clock Theory

The Covert Clock Theory predicts that microtiming deviations for different parts played by an individual musician during any given macroperiod will show substantial agreement amongst themselves. Since the individual is assumed to operate with only one active tactus, and one covert clock, the deviation patterns for the different parts (for instance, the kick, snare and hi hat parts in a drum kit pattern) are expected to be identical.

In contrast, the Covert Clock Theory predicts that different drummers playing at the same time would produce microtiming deviations which do not coincide with other drummers. Since the tactus can be assumed to be at the same phase and period, and the Covert Clock itself culturally determined, drummer might be expected to show similar contours in their deviation patterns. The variation between individuals would then be a matter of amplitude, which would depend on how closely the actual Covert Clock triggers approximated the ‘ideal’ timing of the Covert Clock triggers, according to a precise subdivision of the macroperiod.

Due to consonance of the overt tactus and the covert clock at the start of the macroperiod, deviations at the macroperiod boundary are predicted to be close to zero.

A final prediction of the Covert Clock Theory is that first-order phase correction will affect the timing of the first tatum onset which follows a deviated tactus onset.

These predictions are discussed in this thesis in Chapter 5.

# Chapter 5 An examination of the Covert Clock Theory using microtiming data analysis

## 5.1 Extraction of microtiming data, and general approach to analysis

Several sets of microtiming deviation data are analysed and tested against the Covert Clock theory, in this chapter. The sets of data are drawn from the following studies:

- Freeman & Lacy's (Freeman & Lacey, 2001; Freeman & Lacey, 2002) data from the solo drum break from **The Funky Drummer** track, as played by Clyde Stubblefield on James Brown's album *In the Jungle Groove; (Brown, 1986)*
- Alén's (Alén, 1995) data from recordings of ***tumba francesca*** dance drumming groups in Cuba; and
- data provided by Ernest Cholakis of hi-hat onset times in the Stan Getz and Joao Gliberto performance of the Antonio Carlos Hobim tune ***O Grande Amor***, from the album Getz/Gilberto.

A general introduction to the approaches used to extract deviation data from an audio file is given in the following paragraphs; more detailed explanations are given as required under individual headings for the analyses.

Microtiming deviation data are typically given in the form of deviations for each onset in a repeating pattern (or bar, or loop), from an idealised temporal grid. The *Tumba Francesca* analysis undertaken by Alen (Alén, 1995) gives the deviations as the difference between the idealised internonset values (corresponding to notated values) and the actual interonset values, in the form of cumulative percentages, where 100% constitutes the whole of the pattern. It is useful to convert these tables to millisecond values, on the basis of millisecond values for notated values given in Alen's text (Alén, 1995). The DNA Beat Templates (Cholakis, 1999a) give deviations in the form of MIDI sequencer ticks, also easily converted to milliseconds, for the given tempo. The Funky Drummer deviations (Freeman & Lacey, 2001) were given as absolute onset times in milliseconds (from the beginning of the excerpt), which were converted to millisecond deviations from an idealised grid derived by dividing the length of the excerpt by the number of tatums in it.

Given the deviation data in a uniform format, there are other pertinent differences between the examples studied. Kit drum examples where the various parts (kick, snare, hi-hat) are played by the different limbs of a single drummer, such as The Funky Drummer break and the Clyde Stubblefield DNA Beat Template examples, exhibit conformity in the deviations for the different parts. Patterns involving individual drum parts which are played by several drummers together (as in the *Tumba Francesca* examples) show disparate deviation characteristics for each part. This is hardly a surprising conclusion, since the experimental literature on pulse and entrainment consistently indicates that one tactus applies for an individual at any one time. However, the conclusion does not appear to have previously been explicitly stated or validated against deviation data.

The deviations for each iteration of a groove pattern are often understood to conform broadly to a characteristic trend for the given pattern – that is, they will at least tend to be ahead of the idealised grid, or behind, in roughly the same position in the pattern, for each iteration. For the *Tumba Francesca* drum ensemble data and for the DNA Beat templates analysed below, the given deviations are in fact averages of a number of iterations of the pattern.

The Funky Drummer data was the only set of data I was able to obtain with the raw data for successive iterations of the same pattern intact. Comparison of successive iterations did show marked similarities (discussed below) but also significant differences. Importantly, comparison of successive deviations shows the range of variation, even where the contour of the deviation graph is similar.

## 5.2 The Funky Drummer

The famous Funky Drummer break can be found on the James Brown CD *In the Jungle Groove. (Brown, 1986)*[1] The rest of the band drops out while drummer Clyde Stubblefield continues the same 1-bar pattern solo for 8 bars, after which the band enters again. In a band where funk groove is the goal and reason for existence, the Funky Drummer break is a highlight – Brown's call to the rest of the

---

[1] Microtiming deviation data for successive iterations of the same rhythmic pattern proved difficult to obtain, and I am indebted to Peter Freeman for providing his and Lachlan Lacey's data for Clyde Stubblefield's drum break from James Brown's The Funky Drummer. (Freeman & Lacey, 2001)

band to drop out so that the drummer can be heard solo is evidently a spontaneous decision, made in response to the sense of groove Brown was hearing in Stubblefield's playing.

### 5.2.1 Data acquisition and treatment

The drum pattern itself takes up the four beats of a 4/4 bar and is consistent over the eight bars of the break, with some extra snare or kick drum hits, for decoration. For ease of reference, tatum onsets will be referred to by their beat number and tatum number within the beat: tatum 2-1 is the first tatum of the second beat, which is the fifth tatum in the pattern.

A constant 16th note pulse is maintained on (mostly closed) hi hats, the kick plays two eighth notes at 1-1 and 1-3 and there are snare backbeats on 2-1 and 4-1. There are additional kick and snare hits following beats 2,3 and 4 but these vary between iterations of the pattern.

Freeman and Lacey's data for the Funky Drummer break were obtained by visually placing markers and reading out times in a digital stereo wave editor, providing absolute onset times from the start of the sound file to the nearest millisecond. The complete data set provides absolute onset times to the nearest millisecond for all onsets (including kick drum, snare drum and hi-hat) for the 8-bar break and the down beat of the following bar. The data set represents the tatum onset times of all events (kick drum, snare drum and hi-hat) for the 8 bars – 129 onsets, including the down beat of the ninth bar. Durations between tatum onsets were calculated by finding the difference between successive onset times. Where a kick or snare hit was present, hi-hat onsets were masked. (Freeman & Lacey, 2001). Unless explicitly stated otherwise, all other manipulation of the data was undertaken by myself.

For the present analysis, the difference between tatum 1-1 and tatum 9-1 (the downbeat of the bar following the break) was calculated and divided by the number of tatums (128) to provide a tatum mean duration of 148msec and a tactus period of 594msec. (All timings are given to the nearest millisecond, however calculations were carried out in a spreadsheet program using floating-point calculation.) Deviations were calculated by finding the difference between actual onset timings and undeviated timings based on 148msec tatum durations. A positive deviation value means the tatum onset was late, relative to the undeviated onset time calculated from the tatum mean duration. Tatum 1-1 was arbitrarily assumed to have zero deviation. A chart of deviations for the eight iterations of the pattern is given in Figure 5-1.

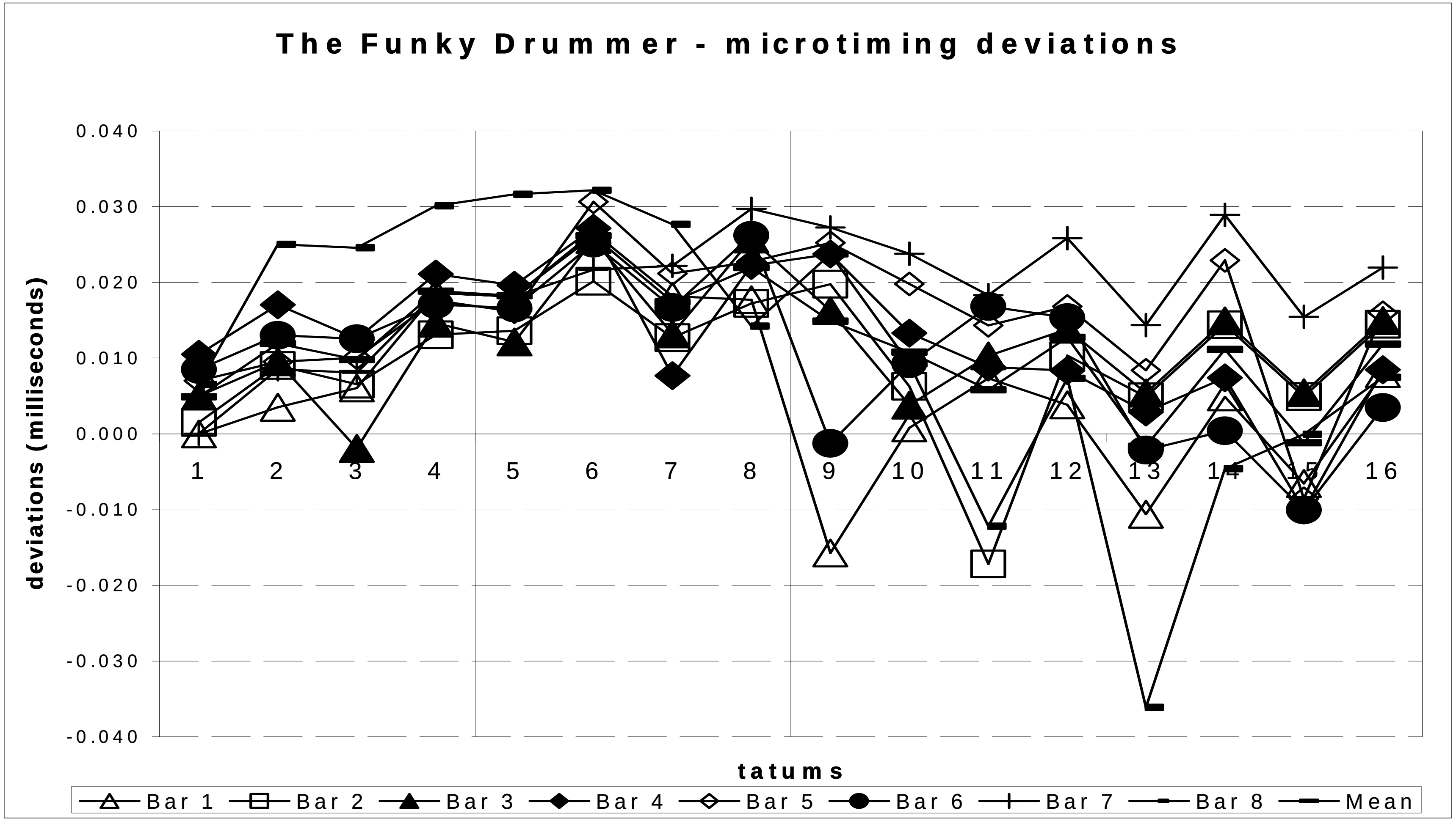


*Figure 5-1 Microtiming deviations for the drum break in The Funky Drummer, by bar.*

### 5.2.2 Analysis

The Covert Clock Theory makes a number of predictions, including:

- that while an individual's performance on drum kit will show substantial agreement between microtiming deviations for the different parts of the kit played by different limbs, different drummers playing in an ensemble will likely show completely different deviation patterns;
- that due to reinforcement of internal clocks, deviations at the start of the macroperiod, where the tactus and covert timers coincide, will be close to zero; and
- that first-order phase correction will affect the timing of the first smallest subdivision onset following an onset which corresponds to a point on the temporal grid provided by the tactus interval timer.

*Prediction 1: An individual's performance on drum kit will show substantial agreement between microtiming deviations for the different parts of the kit played by different limbs.* Inspection of Figure 5-1, in which the kick drum, snare drum and hi hat parts of the Funky Drummer pattern are included together, shows that the deviations within each part follow a reasonably smooth curve.

The largest successive difference in deviations in the whole data set is 43msec in Bar 8, from 3-4 to 4-1. However, this is largely accounted for by the size of the deviation itself (-36msec). The next-largest successive difference 33msec is in Bar 1, from 2-4 to 3-1. It might be thought that this represents a kick drum onset, with timing which is independent from the surrounding hi hat onsets, but in fact there is no kick drum onset on Bar 1 3-1, whereas there is one on Bar 1 3-3, with differences ≤6msec from its neighbours. In most cases where there is a sharp change in deviation amount, it is followed by a more gradual progress in the opposite direction, over successive tatums – for example, at Bar 6 3-1 and Bar 8 4-1, a larger step is in one direction is in each case followed by two smaller steps in the opposite direction. Bar 2 3-3 appears to be the only exception to this general trend. *Prediction 1)* above, that timing will be consistent between all parts for a pattern played by an individual kit drummer, is supported. Contrast with the independent deviations for individual drummers in the *Tumba Francesca* drum ensemble (see below).

*Prediction 2: The start of the pattern or macroperiod is always close to zero deviation.* Apart from the arbitrary zero deviation of bar 1, Bar 7 has zero deviation at tatum 1-1 and the deviations of all other bars at 1-1 are contained between zero and the +11msec deviation of Bar 4. All 1-1 deviations are less than the immediately preceding 4-4 deviation. The late onsets (relative to the calculated mean for the whole 8 bar passage) at the beginnings of bars 2 to 6, and bar 8 could be the result of actual tempo deviation at the macroperiod level.

If the position 1-1 deviations covary with the position 4-4 deviations, it would indicate contribution of tempo change at the macroperiod level to the non-zero deviations at the 1-1 positions. Covariance of 1-1 and 4-4 deviations is given by the ratios of differences between successive 1-1 deviations, relative to difference between the 4-4 deviations in the preceding two bars.[2] The value of 11.000 for Bar 4 of the co-variances (shown in Table 5-1) is an outlier and is excluded from the calculation of the mean. The mean of 0.385 indicates some tempo variation. Tempo adjustment for each macroperiod was calculated on this figure, but the difference was trivial. All analysis has been carried out on the data without correction for tempo changes.

*Table 5-1 Covariance of 1-1 deviations with preceding 4-4 deviations*

| Bar 3 | Bar 4 | Bar 5 | Bar 6 | Bar 7 | Bar 8 |
|---|---|---|---|---|---|
| 0.538 | 11.000 | 0.538 | 0.200 | 0.680 | 0.351 |
| *mean:* | *0.385* | | | | |

Despite the relative global tempo stability of groove music, it may be that there are small fluctuations in tempo between iterations of the macroperiod pattern, which could contribute to the microtiming deviations, which are small in terms of tempo change. Positive deviations at the 1-1 position, if not due to underlying tempo changes, are accounted for by the theory as the failure of muscle effectors to respond consistently to the clock trigger. In the passage under analysis, this could be attributed to a physical carry-over of the positive deviations in the preceding beat: due to inertia or

---

[2] The results shown in Table 5-1 were found through application of a covariance process similar to that used by (Shaffer et al, 1985) to determine whether beats in a piano piece were performed as a concatenation of smaller durations played by the RH, or whether they had a rhythm to some extent independent of the RH durations (the timing of which would then be constrained by the timing of the beats). If the duration of the beats depends on the sum of the durations of RH small notes, then the ratio of the variance of the sum of the intervals making up the beat to the sum of the variance of the intervals will be ≥1. It was found that the ratio was <1, indicating beat independently time from their sub-intervals, and sub-intervals constrained within the beats.

some other reason, the faster movement required to make up the difference and play on the downbeat with zero deviation is not completed successfully.

The deviations from Bar 7 3-4to Bar 8 1-4 are noteworthy. Bar 7 tatum onsets are later from 3-4 to the end of the pattern than in any other bar, and oscillate (due to shuffle feel) between +14msec and +29msec. Tatum 1-1 of Bar 8 arrives much closer to zero deviation, at +7msec – and then immediately moves back to +25msec and gradually rises from there to the peak at 2-2. The sequence of deviations as described has the appearance of a nearly correctly-marked clock pulse at Bar 8 1-1, dropped in between a series of consistently late onsets.

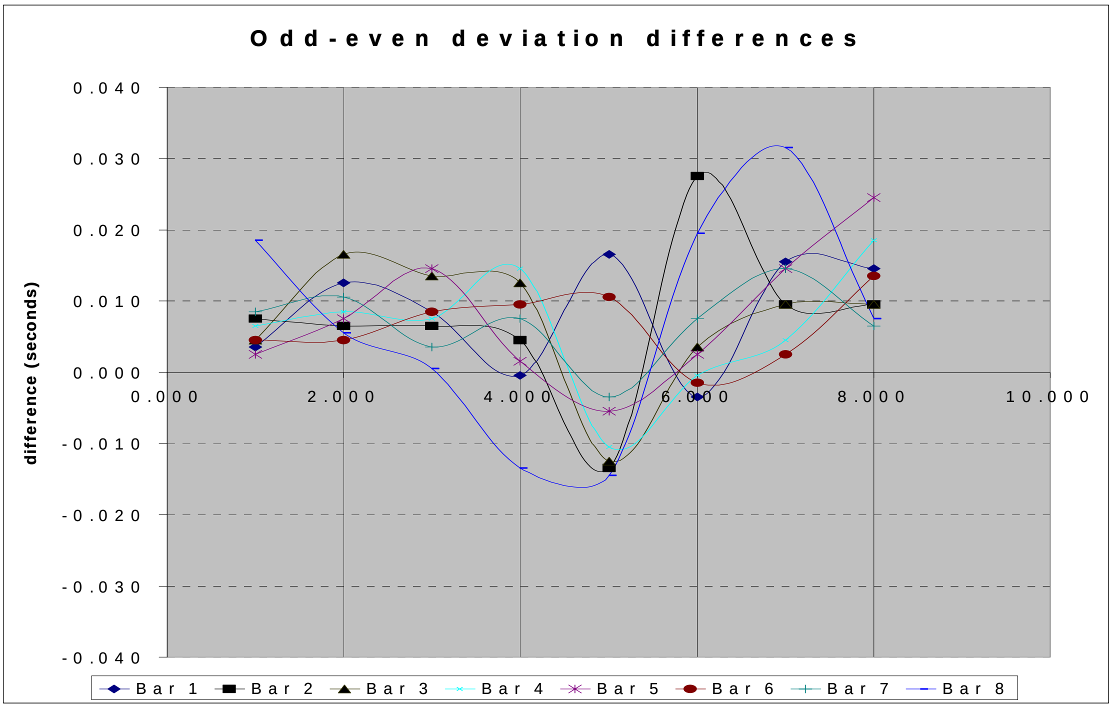
In general, deviations at the 1-1 position, relative to deviations at other positions, are consistent with the *Prediction 2)* above that minimal deviations will occur at the start of the macroperiod, where tactus and covert clock triggers coincide. Further support for the coincident clock aspect of the model are the James Brown band dictum of "Everything on the one", and the implicit assumption of zero deviation at the downbeat of patterns shown in the Groove Templates of Cholakis (Cholakis, 1999a) and in the microtiming data of Alén (Alén, 1995) (data from these sources is discussed below) and Bilmes (Bilmes, 1993b).

*Prediction 3: First-order phase correction will affect the timing of the first smallest subdivision onset following an onset which corresponds to a point on the temporal grid provided by the tactus interval timer.* According to *Prediction 3)* above, phase corrections should be found to tatum 1-2 following the start of the macroperiod at 1-1 where a clock trigger is presumed to have occurred. For instance, where the downbeat of the bar is not at zero deviation (with any tempo change taken into account), the theory predicts that phase correction will occur at position 1-2 in response to the non-tempo component of the positive deviation at position 1-1. It was not possible to conclusively test for first-order phase correction however, because of the difficulty of estimating both swing amount and deviation trend at once, together with uncertainty about the contribution of tempo differences to 1-1 deviations.

### 5.2.3 Swing

The characteristic zig-zag pattern of early-late and odd-even tatum numbers represents a swing or shuffle feel (mainly carried by the hi hat part). There seems to be a tendency for deviations 4-1 to 4-4 to oscillate between two deviation amounts which are consistent within each bar. With Bar 8 excluded because of the extreme negative deviation on 4-1, a mean absolute swing value of 12 msec was calculated from 4 –1 to 4-2 and 4-3 to 4-4 differences. The odd-even mean difference for the whole data set was 7msec. This value is probably too small to give rise to a swing feel, and the low value may be partly due to the rising slope of the deviations in all bars from 1-1 to 2-2 masking (or in fact obliterating the performance of) swing differences. While swing was not a consistent amount from bar to bar, Figure 5-2 shows that it did vary smoothly between successive pairs of tatum onsets, although less so from tatum 3-1 to about tatum 3-4 in each bar – also the region where the deviations are less consistent between iterations. It may be that the greater variation in swing amount and the greater variation in deviations between iterations both result from a lack of timing certainty in the region 3-1 to 3-4.

*Figure 5-2 Funky Drummer odd-even deviation differences by bar*



### 5.2.4 Application of the Covert Clock Theory to The Funky Drummer deviations

The Covert Clock Theory posits a covert clock at a cross-rhythm to the tactus as an alternative reference which distorts tatum timing positions. If the covert clock theory is applied to a macroperiod comprising four beats of four tatums each, the microtiming deviations produced should fall within the range of microtiming deviations found in successive iterations of the Funky Drummer break.

A cross-rhythm of three against a four-tactus macroperiod is commonly found in African traditional music (Arom, 1991), and funk is one of the musics of the African diaspora (Pressing, 2002)*.”*For the Funky Drummer break, it was assumed a covert clock dividing the macroperiod into three equal durations was active. The 3/4 clock would fall on tatum 1-1, 345msec after the second tactus pulse at 2-1 (49msec after the nominal time of tatum 2-2), and 395msec after the third tactus pulse at 3-1 (99msec before tatum 3-4). One way of reconciling the tactus clock with the covert clock would be to generate a motor plan for the first five tatums in the bar which would spread them evenly between the macroperiod start and the *covert clock* pulse at 2-2

The five tatums would be payed a little more slowly than usual, and a positive deviation would progressively result. Tatum 2-2 itself would then be triggered by the covert clock pulse and the next motor program would fit the six tatum onsets from 2-2 to 3-3 evenly into the duration until the third covert clock pulse, with tatum 3-1 coinciding with the third (main) tactus pulse. Tatum 3-4 would be triggered early by the third covert clock pulse and the remaining the five remaining tatums beginning with 3-4 would be spread evenly over the remaining duration of the macroperiod. The result is shown graphically in Figure 5-3, with an absolute swing value of 12msec added to every even-numbered tatum deviation.

The ratio of mean squared error between the prediction of this application of the model – with 100% shift towards covert clock trigger times - and the actual deviations of the 8-bar Funky Drummer break, was calculated at 2.54. Despite the high error, visual inspection of the deviations produced by locking the motor program triggers directly to the covert clock (see Figure 5-3 shows that deviations produced do share some features with the actual deviations, most notably the peak at tatum 2-2, which is found in all iterations of the break.

However, the Covert Clock theory does not predict that motor program trigger times will be shifted to coincide exactly with covert clock onsets. Rather, the space between the tactus clock onset and the covert clock onset becomes a free temporal space in which the musician can flexibly choose to place the motor program trigger. If the model deviations are scaled by a factor of 0.5 – which corresponds to placing the trigger at the midway point between the tactus clock onset and the covert clock onset - the deviations show close agreement with the first 7 tatum deviations (up to and including the peak at 2-2). With the Covert Clock model scaled to 0.5, 13 of the 16 tatum deviations of the model fall within the range of actual deviations in successive iterations of the Funky Drummer break.

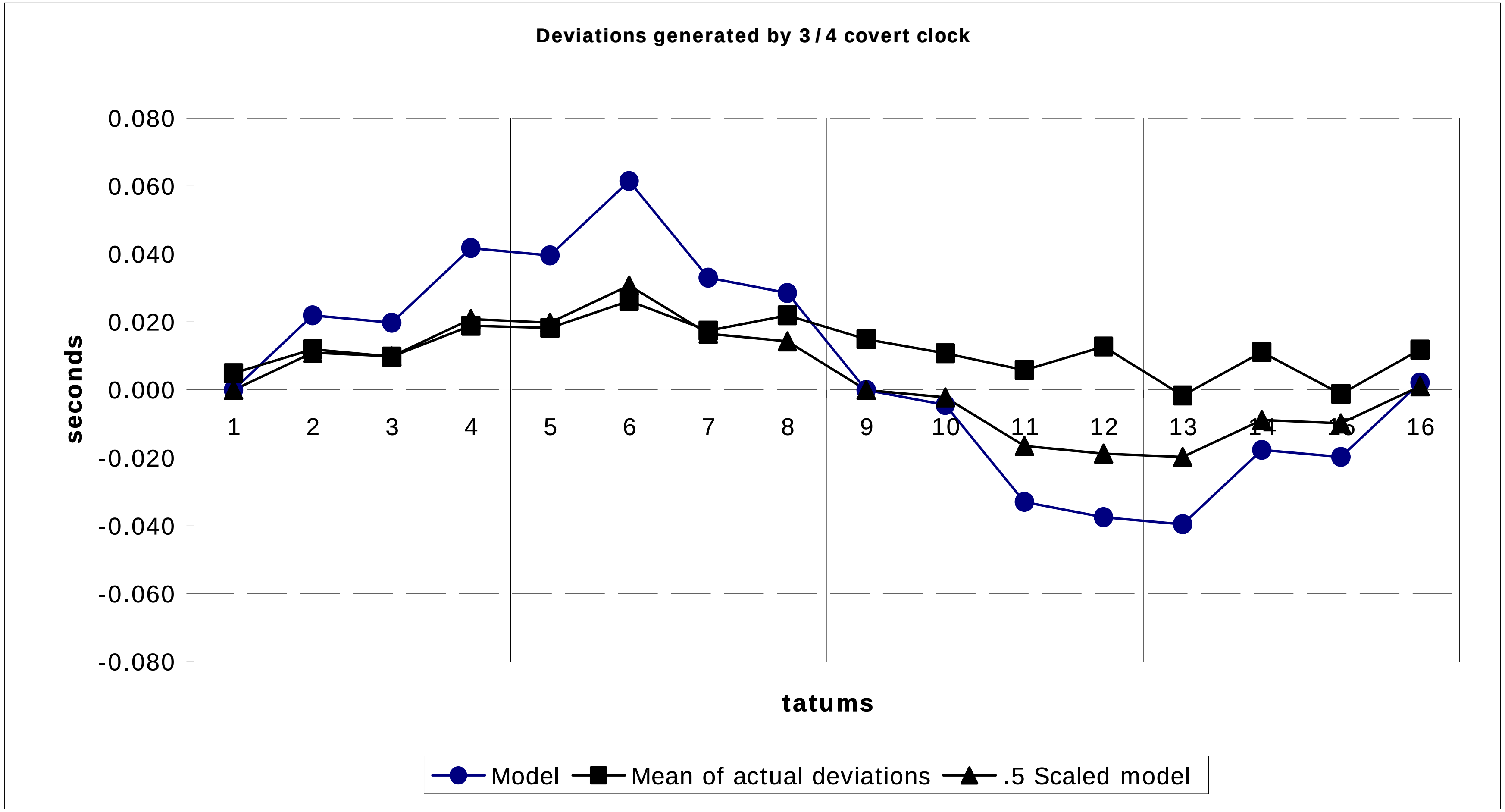


*Figure 5-3 Deviations generated by a covert 3/4 clock against a 4/4 tactus with 16$^{th}$ tatums (see text). An absolute swing value of 12msec has been added to every even-numbered tatum onset time.*

After the peak at 2-2, the two sets diverge. The actual deviations are grouped closely around this peak, within 12msec of each other. After 2-2, the data for the different bars show much less agreement with each other, as well as with the model, although they do have a general trend to the positive (as shown by the mean), in contrast to the model. The furthest the 0.5 model is from the range of actual deviations is at tatum 3-4, where the 0.5 model give a deviation of –19msec, while the lowest actual deviation is +4msec.

The model also shows points of agreement within individual iterations of the performed break. Bars 1 and 6 both show a steep dive from 2-4 (+18msec and +36msec respectively) to 3-1 (-16msec and –1msec, respectively) at 3-1, the same position that the 0.5 scaled model drops from 14msec to 0msec to coincide with the third pulse from the main tactus clock. Bars 2, 5, 7 and 8 of the actual deviations all have local minima at 3-3, for bar 2 it is a global minimum, and for bars 2 and 8 it is below zero. At this position the .5 scaled model falls to –16msec, although its global minimum comes two tatums later at –20msec. Overall, there seems to be a tendency for deviations in successive bars of the Funky Drummer to stabilise the mean deviation, usually at a positive level, towards the end of beat 3, sometimes after rebounding from a local – or global – minimum. From 3-3 in the actual deviations a pattern of oscillation gradually sets in and amplifies to some extent by 4-2.

The minimum for the whole data set at 4-1 in Bar 8 appears to be an anomaly, perhaps caused by Brown rushing the count as he cues the rest of the band to re-enter. The almost flat deviation curve for Bar 8 1-2 to 2-3 – but remarkably, topping the next-highest deviation at 2-2 by only 1msec – follows the oscillation in Bar 7 around a similar positive value, which is the positive boundary of the data set from Bar 7 2-4 to 4-4. Bar 8 is then the positive boundary for the whole set from 1-2 to 2-3, and contains the most extreme deviation of the set at 4-1, the last tactus beat before the band re-enters. The microtiming strategy from Bar 7 2-4 to the end of Bar 8 appears to be a mix of the previously-applied deviation pattern (Bar 7 follows the slope of the mean closely from 2-4 to 4-4) and a push to extremes, possibly as a structural marker for the close of the 8-bar antecedent-consequent group. Structural markers are also provided in the form of the snare drum stick-bounce effects in Bar 4 beat 3 and Bar 8 beat 3.

Overall, the 3/4 covert clock model, scaled at 0.5, provides a good fit for the data from the 1-1 to the 2-3 positions, after which the data diverges widely within itself although still displaying similarity in some features, these necessarily located at the same tatum position in each iteration. The pattern of close fit to the model, up to the same point in each bar, followed by variation, suggests a repeated strategy of utilising the covert 3/4 clock for each bar up tatum 2-3 and then allowing some freedom for the rest of the bar, with a tendency to be behind the beat rather than ahead. This characteristic is consistent with a strategy of stretching the first beat by reference to the covert clock or cross rhythm and following a trend back towards zero deviation afterwards – perhaps loosely referencing the underlying tactus but without constraining the beginning of motor programs until the start of the next macroperiod.

If the function of microtiming is at least in part to break down the clear sense of tactus in favour of both motor-level timed tatum groupings and the containing macroperiod, the strategy described in the preceding paragraph would be effective. The deviation timing pattern is combined with a corresponding kit rhythm pattern which could be designed to reinforce clear timing in the first half of the macroperiod and undermine it in the second half. Apart from constant 16ths on the hi hat, tatums 1-1 to 2-1 in every bar consist of two kick drum hits (on 1-1 and 1-3) and one snare drum hit (on 2-1); this very regular and clearly-marked pattern is immediately followed for the rest of the macroperiod in each iteration by a much busier, less regular and less distinct pattern of off-beat snare hits (played softer than the backbeat), plus variable kick drum placement.

## 5.3 *Tumba Francesca*

### 5.3.1 Data acquisition and treatment

The deviation data provided by Alén (Alén, 1995) was drawn from recordings made of a drum ensemble at a *tumba francesca* society in 1976 in La Pompadou, Guantanamo in Cuba. Recordings of various dances were made of the ensemble as a whole, and then each drum was recorded individually, using the *catá* as a guide. The drummers for each part could not play without reference to the *catá* - a kind of log drum – which carries the most stable rhythm (Alén, 1995) and appears to serve a similar function to the bell pattern in traditional African music. Other drums in the ensemble were the *premier,* the *tambora,* and two *bulás.* Within a given dance piece, such as the *Yubá,* different standard patterns, or *toques,* apply for different sections of the piece, such as *toque macota* and *toque cobrero.*

Transcriptions from the sound recordings (on analogue linear tape at 76 cm/s) were made using a Winckel repeater, which converts tape lengths to time, with an accuracy of 0.1 msec. On each instrument, a total of 20 recordings of each *toque* was made. Since the total length of the *toque* might vary between 1200 msec and 1150 msec, the interonset intervals (IOIs) were converted to a percentage of the total length of the *toque* in which they occurred, to make possible comparison between different iterations of the *toque.* [Note that this process seems to make the implicit assumption that each pattern begins with zero deviation.] Alén gives these values in a table, together with ideal or 'schematic' values corresponding to the musical notation of the transcription. For instance, in a pattern comprising 12 eighth notes, an idealised quarter note constitutes 16.7% of the total pattern length, while the mean over the 20 iterations of the IOI in that position of the *toque* might be 16.4%. (Alén, 1995) For the present study, the percentages were taken to refer to a standardised duration for the *toque* of 1200 msecs. The attack time, within the 1200 msec duration, was calculated for both the idealised IOI percentages and for the mean of the actual performed durations. The difference was then taken as the millisecond deviation.

Alén (Alén, 1995) states that great difficulty was encountered in transcribing the recordings into standardised musical rhythm notation, due to the complex ratios between IOIs intervals in individual parts, and that it was this experience which led to a detailed study of the exact temporal relationships in the rhythmic motifs of the drums which support the improvisation of the *premier.* In terms of the theory elaborated in this thesis, the preponderance of complex ratios would indicate periods of control at the motor program level, rather than being referenced to each tactus clock.

### 5.3.2 Application of the Covert Clock Theory to Tumba Francesca deviations

Visual inspection of the graph of the deviations of the *catá*, *bulá I,* and *bulá II* of the *toque macota* of the *Yubá* dance (see Figure 5-4 shows that *bulá II* deviations are in all cases the most extreme, with the *catá* in all cases but one being the least extreme, and *bulá I,* again in all cases but one, lying between the two. The one exception occurs at eight note number 7, where *bulá I* has zero deviation while the *catá* has –3.6 msec deviation. at this point , however, the deviations are minimal – less than 5 msecs from zero, so this discrepancy in the overall relative amplitude of the respective deviations of the different drums may not be significant. What is striking, is that the deviations for at least the two *bulás* appear to conform to a similar contour, although with a different degree of amplitude for each drum.

The catá deviations, for this and other toques given in (Alén, 1995), rarely exceed ±10 msec, and are more often in the ±5 msec range. The catá deviations reach 14.4 msec and 15.6 msec in two different toques, but otherwise remain within the ±10 msec range. Attempts to apply the Covert Clock theory to the tumba francesca deviations meet with more difficulty with regard to the catá than with other parts. It may be, in fact, that deviations of < ±5 msec are not musically significant. Where deviations in the catá part are > ±5 msec, they appear to follow a deviation pattern which is independent from other parts – see especially the Frenté dance shown in Figure 5-5.

### The Yubá

The *Yubá* dance has a meter of 12 eighth notes. If a covert clock of 5 against 12 is hypothesized, the clock times for the covert clock will fall between the 3rd and 4th the 5th and 6th; the 8th and 9th; and the 9th and 10th eighth notes of the *toque.* By selection of which of the two eighth notes to alter the timing of, by reference to the covert clock, the contour of the deviations in the *toque* can be approximated. Thus, the 4th, 6th, 9th and 10th onsets are deviated towards the covert clock onset timings, although of these, only the 6th eighth-note onset of the *toque* is actually the closer one of the pair framing the covert clock onset.

The *catá* is ignored, for the reasons given above. As for The Funky Drummer analysis, deviations of the onset time towards the covert clock onset were scaled, where a scaling of 1.0 would represent coincidence of the deviated onset with the covert clock onset. The results for the Bula I and Bula parts*,* scaled by 0.17, are given at Figure 5-6 and 5-7, respectively. Where application of the covert clock theory provides deviation data at the same eighth note as an onset occurs, differences between the covert clock application and actual deviations are < ±5 msec; and where no onsets and covert clock data do not occur on the same eighth note, the covert clock data interpolates between actual onset deviations with a similar degree of accuracy. Similar results are obtained for the *bulá II* deviations in the same *toque,* this time with a scaling factor of 0.29 – again, covert clock-generated deviations are either within ±5 msecs of the means of actual deviations, or interpolate between them with equivalent accuracy.

Thus, this example may be seen as strongly supporting the covert clock theory of microtiming deviations.

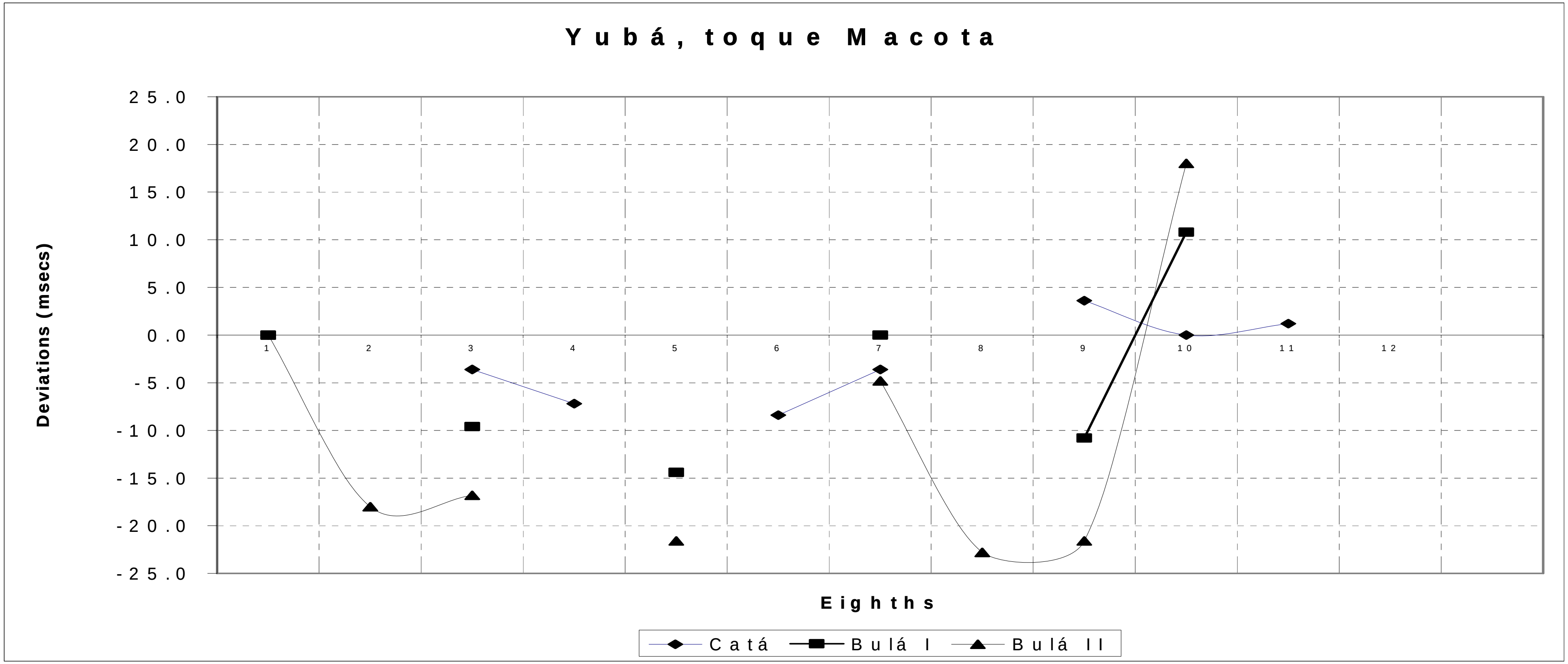


*Figure 5-4 Deviations for t3he* catá, bulá I *and* bulá II *of the* toque macota *of the* Yubá *dance. Individual drum strokes on successive eight notes (the smallest rhythmic value) are joined by lines.*

### The Masón

The *Masón* dance also has a meter with 12 eighth notes, and once again a covert clock of 5 against the 12 is chosen. For the *Masón,* the 4th, 7th, 9th, and 11th eighth-note timing positions are altered with reference to the covert clock onset times. This differs from the *Yubá,* where the 6th, rather than the 7th, and the 10th, rather than the 11th eighth-note timing positions were altered. The covert clock onset time actually lies between the 5th and 6th eight-note onset times, so deviating the 7th eight-note position is a more extreme choice, necessitated by the large (-139 msec) deviation at that position in the *bulá* part. Similarly, the covert clock onset timing lies between the (undeviated) 9th and 10th eighth-note timings, and altering the 11th eighth-note onset timing by reference to the covert clock timing is a more extreme measure than altering the 10th eighth-note onset, which lies closer to the covert clock onset and between it and the 11th eighth-note onset.

The tambora part for the Masón has only two onsets, at the 1st and 9th eighth-notes. Zero deviation is assumed at the first eighth-note position of the toque. Scaling the covert clock influence by 0.22 produces a deviation within 1 msec of the actual deviation. (Figure 5-8)

A scaling of 1.0 is required to come close to the –139 msec deviation in the bulá part. (Figure 5-9) Although this scaling works well for the deviation at the 4th eighth-note, coming within ±5 msec, it only produces a –120msec deviation at the 7th eighth-note, where the actual deviation of -139 msec occurs. Also, the interpolations at the end of the bar would appear to be approximately 35 – 45 msecs out from a linear interpolation between the actual deviations. Application of the covert clock theory to the Masón is therefore less successful than its application to the Yubá, toque macota. However, as with application of the theory to The Funky Drummer, the covert clock approach may be seen as a useful model, if some flexibility is allowed in its use by musicians. Rather than a consistent and rigid application of a single scaling factor between the tactus and covert clock timings, in a performance situation the drummer might apply a covert clock strategy to generate a contour of microtiming deviations, which over many successive iterations of the pattern is itself used only as a reference, with variations to a greater or lesser degree around it.

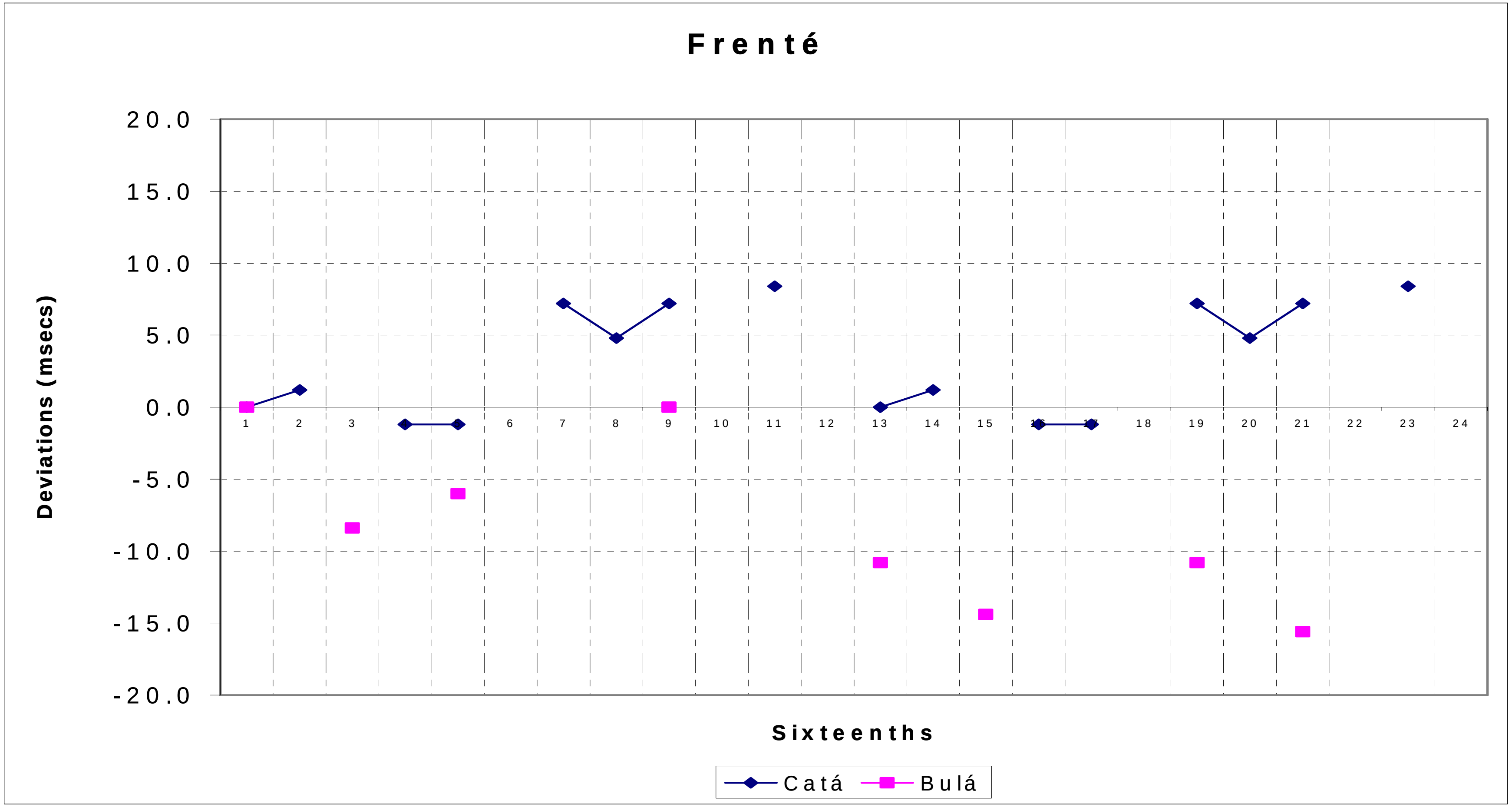


*Figure 5-5* Catá *and* bulá *deviations for the* Frenté *dance. The* catá *appears to follow a deviation strategy which is independent to that of the* bulá.

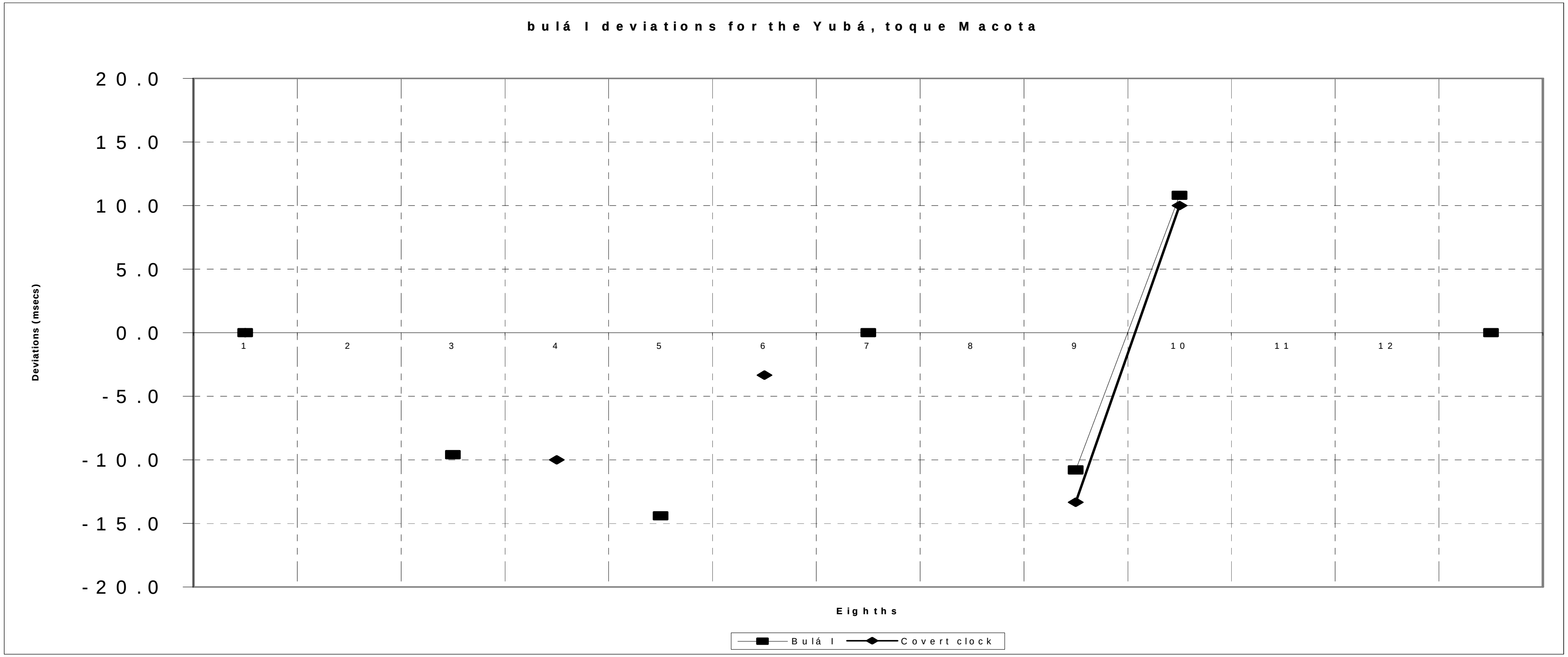


*Figure 5-6* Bulá I *deviations using a covert clock of 5 against 12, scaled by 0.17*

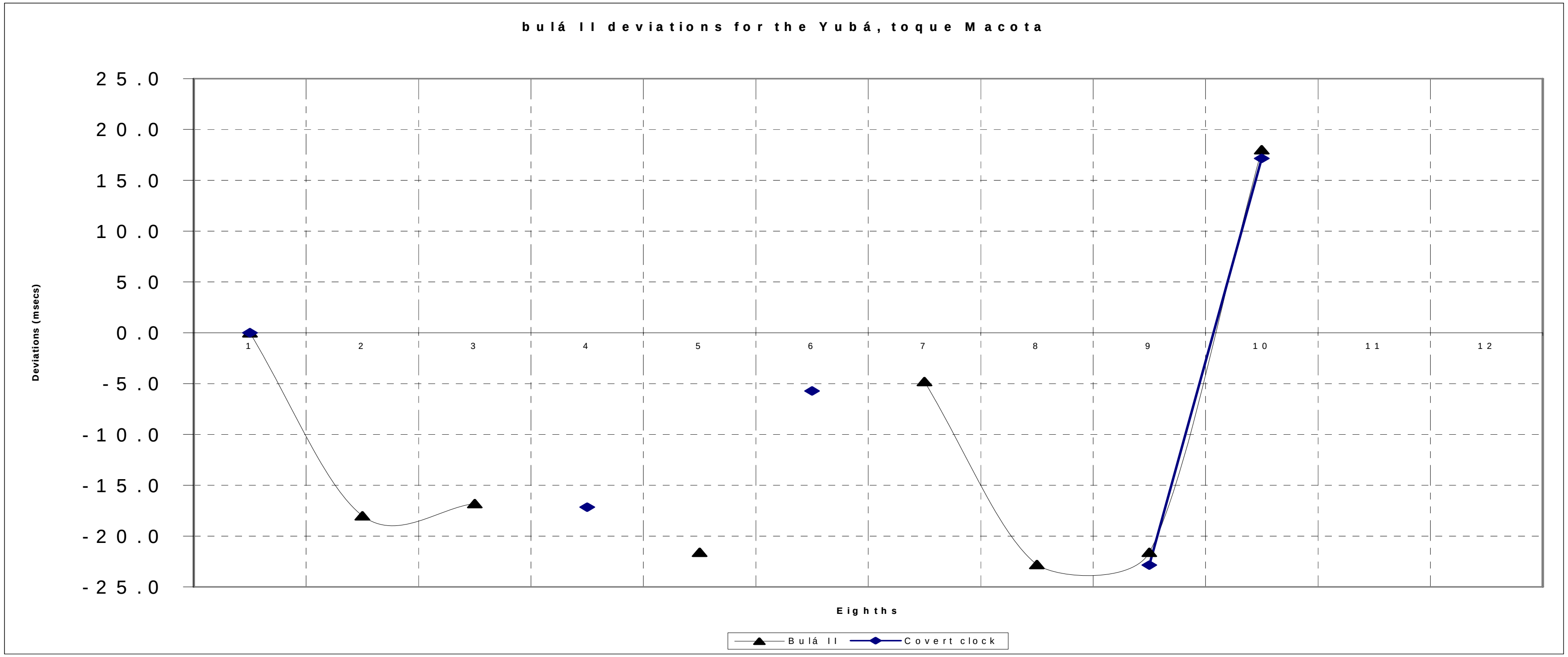


*Figure 5-7* Bulá II *deviations using a covert clock of 5 against 12, scaled by 0.29*

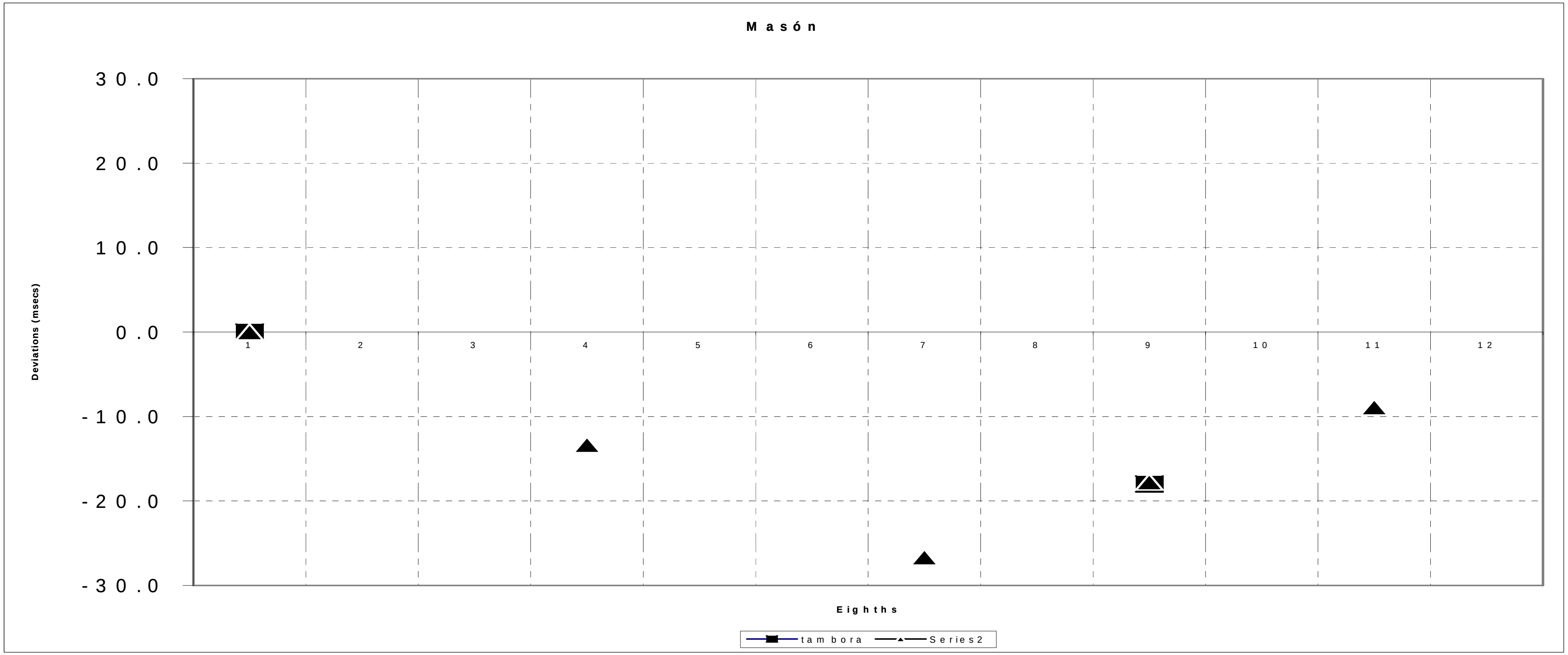


*Figure 5-8* Tambora *deviations for the* Masón *using a covert clock of 5 against 12, scaled by 0.22*

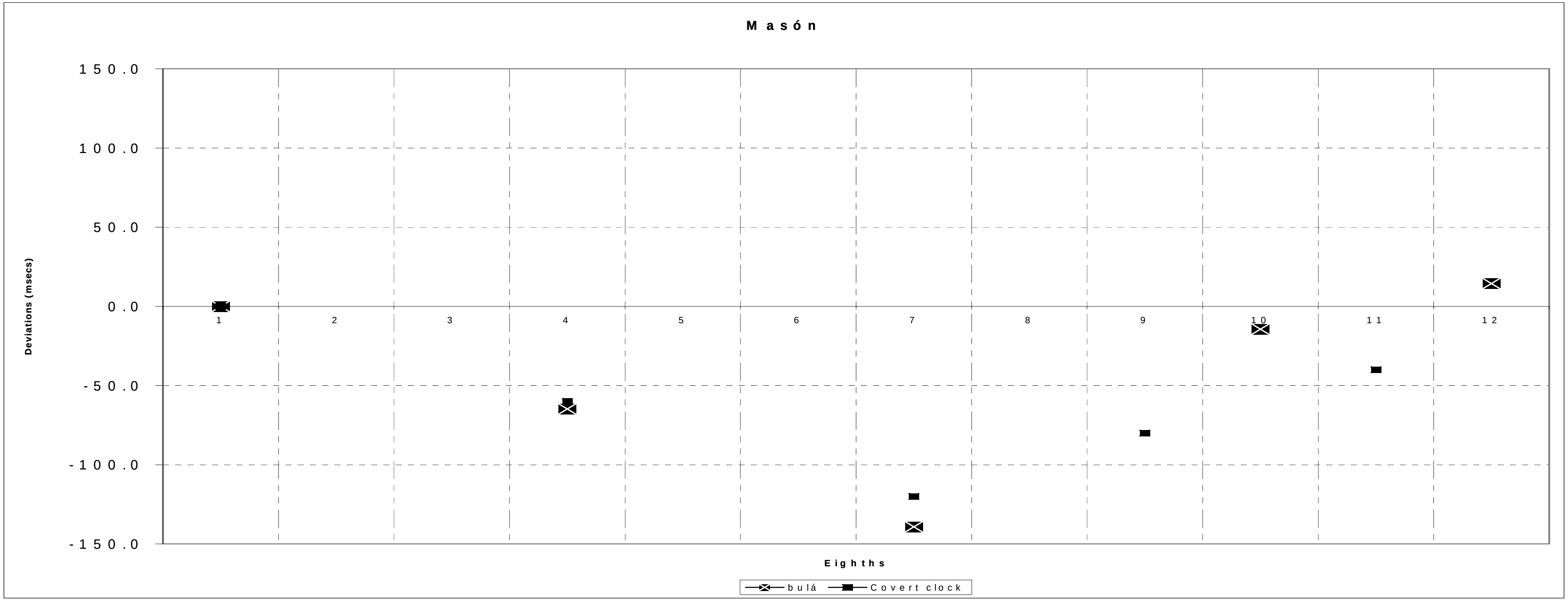


*Figure 5-9* Bulá *deviations for the* Masón *using a covert clock of 5 against 12, scaled by 1.0*

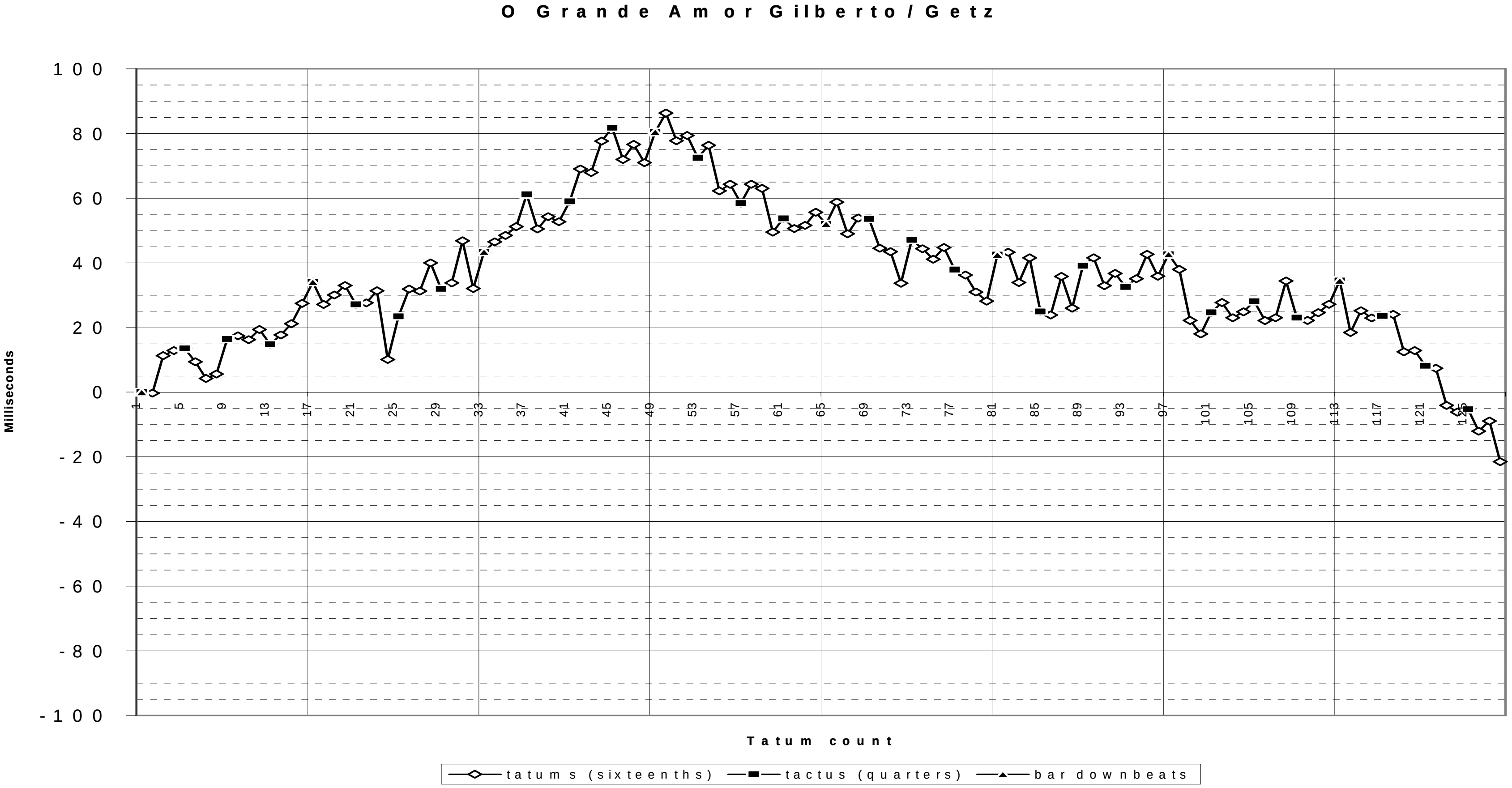


*Figure 5-10 Hi-hat deviations for Getz/Gilberto's* O Grande Amor

If it is accepted that the Covert Clock theory, with the caveat just described, can account for microtiming deviations in the two *tumba francesca* examples above, a notable feature is the agreement between individual drummers in how to apply the clock for the particular *toque* being played. Both for the *toque macota* of the *Yubá* dance and for the *Masón* dance, the timing deviations for each part are generated by an application of the covert clock which is the same for each part in the *toque* (not counting the *catá)* but different for different *toques.* Although a 5 against 12 covert clock is used in both instances, different eighth-note positions are referenced to the covert clock onset times for each instance, thus producing different deviation patterns. The result is particularly convincing for the *toque macota* of the *Yubá* dance, where application of the covert clock theory in the same way for the two different parts, with different scaling, produces deviations within ±5 msecs of all onsets, or within ±5 msecs of linear interpolation between onsets, where the covert clock-generated deviation does not fall on an onset of the *toque.*

## 5.4 *O Grande Amor*

The analysis of the eight bars of the Stan Getz and Joao Giberto's performance of the Antonio Carlos Hobim tune *O Grande Amor*, from the album Getz/Gilberto, was provided to me by Ernest Cholakis, producer of DNA Beat Templates. (Cholakis, 1999b) Cholakis has developed software to finds onsets in soundfiles. He describes the way it works:

> It is proprietary software that I developed over the years. The transient separation is achieved primarily through filtered based (FIR Impulses). The filters have to be adjusted for each project (different instruments/overall sound) and adjusted in order to eliminate strong instrument resonance's in the spectrum so that the transient is more apparent. The waveforms are often converted into decibels representation -so that low level details can be seen on a screen (the dynamic range is often greater than 60db which is one part in a thousand - something that cannot be viewed when looking at data in a linear way).(Cholakis, 2003a)

Hi-hat onset times are given in samples at 44.1 kHz. The tempo – 64.91 bpm – is calculated by dividing the number of beats into the total duration of the excerpt. Nominal attack times are found in relation to the tempo and deviations calculated by subtracting the nominal attack time from the actual attack time. The result is converted from samples to milliseconds. (Cholakis, 2003b)

The excerpt is not a repeating groove pattern macroperiod but rather one section of a bossa nova tune. The microtiming deviations are interesting as a contrast to the Funky Drummer and *tumba Francesca* groove patterns analysed above: here, rather than beginning from close to zero deviation at each start of a pattern which repeats every one or two bars, there is an extended excursion and return in the positive (late) direction, over the eight bars of the song section. (Figure 5-10)

The Covert Clock theory clearly will not apply in this case: microtiming deviations for *O Grande Amor* are referenced to the 8-bar section of the song, not to the meter. The deviations – apart from providing the laid-back, behind-the-beat feel which is typical of bossa nova – function to articulate structure, by remaining well behind the beat for the duration of the song section but returning to the neighbourhood of zero as the section ends. The deviations reach their peak (86 msec) at the second tatum of the 4th bar – nearly a whole bar before the middle point of the section, a choice possibly made simply to avoid a too symmetrical effect. Following tatum 4-2, deviations begin to reduce in amplitude. If local fluctuations in amplitude are discounted, the reduction is more or less linear and if continued would reach zero deviation about 1.5 bars before the end of the section. However, a kind of plateau around ~40 msec is reached (again discounting local fluctations) at bar 6, which drops to another quasi-plateau of roughly 25 msec in bar 7. The return towards zero deviation in the second half of the excerpt (bars 4 to 8) can be seen as an exercise in musical expectation: the return to zero is first indicated with the linear reduction in deviation amplitude; next, the fulfilment of the expectation is delayed, by plateauing in bars 6 and 7; and finally, there is a steep return to zero deviation in bar 8, with in fact the last 6 tatums in negative territory – they are early. As noted, the whole arch of the deviations over the course of the song section serves as a strong structural marker; the slight anticipation of the last six tatums will work to smooth over the join between sections, to lead the listener's ear smoothly into the next section.

## 5.5 Summary

Analysis of microtiming deviations in the Getz/Gilberto's performance of *O Grande Amor*, a bossa nova tune, shows that deviations are referenced in this case not to a groove macroperiod, but to a song section. The function of the deviations in this case relates to articulation of song structure and is not susceptible to analysis by application of the Covert Clock theory.

The Covert Clock theory, as applied to James Brown / Clyde Stubblefield's Funky Drummer break and to *tumba francesca* performances appears to match the data well, if allowance is made for a less rigid application of the theory in some instances. The closer fit of the *toque macota* of the *Yubá* dance may reflect the need for the drummers of the two parts analysed to fit with each other and with the *catá* reference, whereas Clyde Stubblefield in The Funky Drummer break is playing completely solo. Similarly, the small number of onsets played by the *tambora* and *bulá* in the *Masón* dance of the *tumba francesca* may mean that the drummers play with more freedom, referencing more loosely to the covert clock.

# Chapter 6 Design and implementation of the VROOGE model of entrainment and microtiming deviation

This chapter describes the computational implementation of the Covert Clock Theory in the VROOGE model of microtiming deviations. [The name *VROOGE* is simply an anagram of ‘groove’.] The VROOGE model integrates entrainment and microtiming deviation functions in one model. The model is designed for robustness of function for use in a music production context and some aspects of the implementation depart from the theory discussed in earlier sections of thesis. For instance, a second-order period correction process (similar to that decribed in (Mates, 1994a)) provides faster and more stable entrainment than the first-order process indicated in the experimental literature discussed below.

## 6.1 Outline of the VROOGE model

In order to generate new, musically convincing microtiming deviation patterns, a model which takes account of human perception and production of rhythm is needed. The VROOGE model locates separate phase and period correction processes respectively in the motor (implementation) and clock (internal) areas of the basic Wing and Kristofferson (Wing & Kristofferson, 1973a; Wing & Kristofferson, 1973b) two-process model for repetitive actions.

Microtiming deviations – which according to the Covert Clock theory, are the result of expanding or compressing the duration of motor programs, relative to the duration of the macroperiod – are implemented in the VROOGE software by means of a percentage variation of the playback tempo of the recorded sequence of onset times, where the base time is the current global tempo.

The tempo variation is calculated by the model in accordance with parameters set by the user. The number of tatums of the sequence is set, and the user chooses the Covert Clock subdivision of the macroperiod – for example, there might be 12 tatum subdivisions in the macroperiod, and 5 covert clock onsets. The user then chooses which tatum to associate with which covert clock onset; and a percentage amount by which to influence the timing of the chosen tatum.

A 100% timing shift for the tatum will result in that tatum occurring at the same instant as the (idealised) covert clock onset: the step-shift in tempo is the derivative of the accumulated microtiming deviation. Associating a tatum with an earlier covert clock onset will result in a shorter duration for the given motor program, and vice versa, while association a particular tatum with a later covert clock onset will result in a longer duration for the motor program. A shorter-duration motor program is equivalent to a faster tempo for that section of the macroperiod. Typically, a combination of faster and slower local tempi within the macroperiod will be chosen (by choosing earlier or later covert clock onsets to associate with a given tatum), resulting in both negative (early) and positive (late) deviations, or at least a return to near zero deviation towards the end of the macroperiod.

The VROOGE model can be shown to accommodate stable global clock timing, while allowing the microtiming deviations typical of groove music. A comparison of the features of the VROOGE model with other models of microtiming deviations is given in Table 6-1. Note that, for simplicity, the comparison is limited to features related to onset timing variation only – Bilmes's Xcite software in particular is a full-featured drum sequencer including control over MIDI velocity (loudness) and providing flexible, on-the-fly capability to combine different rhythmic patterns. Since the focus of the current theory and model is exclusively on entrainment and micro deviation of onset timings, and for reasons of space, these features of Bilmes's model are omitted from the table.

## 6.2 Existing models of microtiming deviations

The *groove template* approach developed by Cholakis (Cholakis, 1999b) found in most commercial MIDI sequencing software today and Bilmes's (Bilmes, 1993b; Iyer, Bilmes, Wright, & Wessel, 1997) re-application of timing deviations derived from analysis of audio recordings are essentially the same process. Individual deviation values are applied to pulse onsets within a repeating rhythm pattern, and the only way to create a new pattern is through more analysis of audio recordings. Bilmes also describes an interface which allows manual setting of individual deviations. (Bilmes, 1993a) Waadeland's Rhythm Frequency Modulation approach (Waadeland, 2001) is radically different in its holistic, systematic, parameterised control of timing. A table summarising and comparing the main features of the three models and the VROOGE software is given at Table 6-1.

### 6.2.1 Groove templates

Composers and arrangers working with a MIDI sequencer can make use of 'groove templates' to emulate the microtiming variations of a performing musician. A groove template defines how far ahead or behind its undeviated timing each fastest pulse will sound, with the same pattern of microtiming deviations usually repeated every two bars.

The deviations are defined in terms of “ticks”. Ticks do not have an absolute duration but are relative to the quarter note (or crotchet) in the sequence. A resolution of 96 parts per quarter note (ppq) is the usual minimum, with resolutions up to 960ppq available in some sequencers. Groove templates are applied to a quantised performance, where the timing of each onset is mapped exactly to the fastest pulse on which it would be shown in music notation.

Numerical Sound is a company which sells groove templates in their own proprietary format. Numerical Sound owner Ernest Cholakis says:

> Designing a groove template from scratch can be a hit and miss affair. If you know from past experience exactly where each grid point should be, then you can easily create a template. You can use the create template feature (available on most midi sequencers) to design very obvious rhythmic signatures, like shuffles, lags, etc. However, beyond this it is difficult to know where to place each grid position.
> With DNA Beat Blocks, we created templates that are based on actual acoustic performances, that already have "proven" feels. The timing of each pulse is extracted and provided in the form of a quantized template. All the DNA groove template are two bars long. (Cholakis, 1999b)

Cholakis notes that variations over subsequent sections of a song would be similar but would contain different timing features. He recommends that when a groove template is applied to a MIDI arrangement it should be changed subtly over the course of the song, not simply cut and pasted to different sections. (Cholakis, 1999b) However, modifying a groove template requires specifying a change to the amount of variation applied to each pulse over the two-bar template. There is no systematic way to vary the template overall. The 16- and 32-bar groove templates included with numerical Sound’s Groovin’ Rhythm Streams released in 2003 (Cholakis, 1999b) appear to be an attempt to overcome this limitation.

### 6.2.2 Bilmes’s approach

Bilmes develops an approach allowing selection of rhythmic patterns and control over aspects of microtiming variation in live performance. (Iyer et al., 1997) Bilmes et al’s Xcite software was implemented in MAX, a graphical programming environment for musical applications(Puckette, 1991). The software is used to trigger MIDI modules. Rhythmic cells can be created in advance for selection during performance, or created as required using a graphical editor. Each note event occurring within a cell has its onset mapped to a pulse (tatum) within the grid and has values attached for note type, duration, and deviation. Deviations take on continuous values from –0.5 to +0.5 tatums, allowing all possible rhythmic placements. Deviation and accent values can be exaggerated or reduced (separately) with a non-linear (power law) tool for compression and expansion. Bilmes et al's model seems to be a powerful and full-featured application suitable for live performance, with control of dynamics and note offset, which experienced musicians say are important components of groove. Control of tempo variation over large timescales is also provided. However, deviations must still be chosen manually – the authors do not elaborate on the "probabilistic processes" which are a feature of the software – despite the continuous control of overall deviation amount.(Iyer et al., 1997) The necessity of making manual adjustments for each deviation in a pattern would militate against creating new timing deviation patterns during performance – it seems much more likely that deviation patterns would be applied to the cells prior to performance.

Elsewhere(Bilmes, 1993a,1993b), Bilmes describes in detail his technique for extracting deviation timing from audio recordings. Presumably using a similar approach to Cholakis(Cholakis, 1999b), Bilmes is able to extract note onset times from digitised multitrack recordings (with each drum on a different track), calculate the tempo, the tatum grid times (ie. where undeviated onsets would occur) and actual microtiming deviations.(Bilmes, 1993a,1993b) Bilmes recreated the performance with quantised attacks, with random Gaussian deviations added to the quantised attack times (using the same mean and variance as the actual deviations), and with the extracted deviations added back to the quantised attack times. Most listeners found the quantised, undeviated version "mechanical"; the random variation version "sloppy" or "random"; and said the version with the correct deviations added sounded most like the original.

Bilmes (Bilmes, 1993a) describes a deviation experimentation program called xited (for experimental Interactive Tatum-Editor of Deviations) which runs on SGI IRIS Indigo workstations. Similar to standard software drum machine emulators, xited provides a pattern window with rows of toggle buttons corresponding to tatums which are used to program whether or not a drum stroke occurs on any individual tatum. Different rows correspond to different types of drum. A slider for each tatum affects deviation times for all drum onsets on that tatum. Once again, an individual choice must be made regarding microtiming variation for each onset in the pattern, although the sliders provide an intuitive way to change the amount of variation as well as a visual overview of the variations chosen for the whole pattern.

Bilmes has, however, undertaken some interesting investigation into systematic specification of microtiming deviation. (Bilmes, 1993b) Having found that a random Gaussian process with the same mean and variance of the whole instrument track (drawn from a multitrack recording of a performance by Cuban drum ensemble Los Muñequitos de Matanzas) did not produce a musically coherent effect, he applied the same process separately to onsets grouped according to their position in the bar. Thus, mean and variance for the $i^{th}$ onset in the 16 bar pattern as it was repeated over the duration of the recording was used as the basis for a random Gaussian process applied to all $i^{th}$ onsets in the track. Again, this was not musically successful.

Bilmes then used a function approximation technique to map microtiming variation patterns to particular rhythm patterns which recurred in the track. To do this, he first applied a Clustering Criterion Function to the rhythm patterns, to identify identical sequences of strokes. These were required to begin at the same point in the measure, to have the same number of tatums between their start and finishing strokes, to have the same sequence of stroke types (each drum can be hit in two different ways), and the inter-onset times between strokes had to be identical. Patterns meeting these criteria were subject to an optimisation strategy which minimised the number of different patterns. (Bilmes, 1993b)

The mapping function was implemented by means of neural networks. A set of input units was allowed for each tatum in the longest pattern identified. A particular unit in the set was activated if a drum stroke of the type corresponding to that unit occurred on that tatum. No unit was activated where a rest occurred. One neural network output the probability (presumably taken over all the patterns in the performance) relating to the deviation value of each tatum in the rhythm pattern being analysed; the other neural network learned the sign of the deviation. The probability value and sign were then plugged into an equation (the inverse of the probability of deviations in a Gaussian distribution – so mean and variance were pre-computed) to give the signed deviation value for each tatum in the pattern.(Bilmes, 1993b)

The Cuban ensemble analysed by Bilmes (Bilmes, 1993b) can be presumed to include a greater variety of rhythm patterns in a given track, than for instance, the Phil Collins track (and similar western pop music songs) analysed by Cholakis. (Cholakis, 1999b) Thus, Cholakis expects to find subtle variations of the same deviation pattern in different sections of the song (albeit different for each of hi-hat, snare, kick and bass guitar), while Bilmes finds distinct deviation patterns attached to specific rhythm patterns in the parts played by different drummers in the Cuban ensemble. Bilmes's analytic process is useful in identifying and grouping rhythm patterns and the deviation patterns associated with them. As sophisticated as his analysis is, the outcome is finally only a set of deviation values attached to particular pulses – that is, a set of static values which can be increased or decreased overall, but with otherwise no holistic method for creating or varying deviation patterns.

Both Cholakis (Cholakis, 1999b) (and similar approaches to extracting groove templates from audio) and Bilmes (Bilmes, 1993a) describe what might be termed a ‘black box’ approach to microtiming deviation patterns: a particular pattern of deviations is associated with a particular rhythm pattern, with no attempt to construct a system which can generate deviation patterns for new rhythm patterns. Apart from increasing or decreasing the amount of deviation overall (multiplying all deviations by a constant, or raising all deviations to the same exponent) there is no model of how particular patterns might come into being. A new pattern of deviations can only be created (rather than extracted from a recorded performance) by making individual deviation settings for each tatum or fastest pulse with the rhythm pattern. That is, there is no means of relating a deviation setting at one point in the rhythm pattern to a deviation setting at any other point.

### 6.2.3 Simulating Expressive Timing By Modulated Movements – Waadeland’s model

Waadeland’s (Waadeland, 2001) Rhythm Frequency Modulator – on which the present model is based - utilises Frequency Modulation of the tatum rate to produce timing deviations. The Rhythm Frequency Modulator timing deviation model is unique among models of microtiming deviation in groove – first, in its attempt to model the process for creating deviation patterns on human movement dynamics; second, in that the process creates patterns of deviations with interrelationships between the individual deviation and the history of deviations over the course of the rhythm pattern. [For a survey of approaches to expressive timing in non-groove musics, where the global tempo is free to vary for expressive purposes, see the Appendix A to this thesis.]

Waadeland points out that research on rhythm performance which focuses on onset times as a finite number of discrete points along the time axis ignores the continuity of movement in time and space through which the performing musician produces the rhythm. (Waadeland, 2001) He points to empirical studies which show that listeners can recover motion information encoded in music and convert it into body movements. (Becking, 1928; Clynes, 1977,1986a,1992; Shove & Repp, 1995; Sievers, 1924; Todd, 1992a,b,c; Truslit, 1938) Waadeland takes a cosine wave (vertically shifted so that the lowest point of the function is equal to zero) as a possible curve described by the movement of a hand beating to an isochronous pulse:

*Equation 6-1*

$$\mathrm{p}(t) = A[1 - \cos(\, ft \,)]$$

where *i* is time, *f* is frequency, and 2A is the maximum distance of the hand from the minimal value, 0. The pulse coincides with the lowest point of the curve, which is zero in this formulation. Note that the vertically-shifted cosine function gives a value of 0 at $t = 0$. Introduction of a phase-shift variable, $\phi$ allows the use of a sine function:

*Equation 6-2*

$$\mathrm{p}(t) = A[B + \sin(\, ft + \phi \,)]$$

Values of $B = 1$, $\phi = 3\pi/2$ in Eq. (2) will give the same curve as Eq. (1).

Waadeland goes on to elaborate the equation in order to introduce ties and subdivisions, which enable the equation to define rhythmic patterns rather than simply isochronous pulses.

So far, Equation 6-1 and Equation 6-2 define undeviated rhythms. However, Waadeland introduces into the model the Frequency Modulation techniques pioneered as a form of sound synthesis by Chowning (Chowning, 1973). Of course, where FM synthesis uses frequencies in the audio spectrum (above 20Hz) for both the *carrier* (corresponding in Waadeland's model to the pulse equation given in (1) and (2)) and the *modulator* (the wave which controls deviations to the frequency of the carrier) – for Rhythm Frequency Modulation, both carrier and modulator will be at sub-audio rates, typically within the range ~1Hz to ~10Hz. The instantaneous amplitude of the output wave as a simple FM instrument, with one carrier and one modulator, is given by:

*Equation 6-3*

$$A\sin[f_c t + d\sin(\, f_m t \,)]$$

Where $f_c$ is the *carrier frequency*, $f_m$ is the *modulator frequency*, and *d* is the *peak frequency deviation.* (Chowning, 1973) With $d = 0$, the carrier will be unaffected by the modulator and will be a pure sine wave – or in the Rhythm Frequency Modulator, the basic pulse will be undeviated. Waadeland gives the basic rhythm frequency modulation equation as:

*Equation 6-4*

$$r(t) = A[1 - \cos[ ft + d\sin( f't + \phi )]]$$

which allows a phase offset $\phi$ of the modulator, and where $f'$ is a sub-harmonic of $f$. Thus this equation:

*Equation 6-5*

$$r(t) = 1 - \cos[ 3t + \sin( t + \pi/2 )]$$

produces uneven beats with periods in the ratios 0.42 : 0.32 : 0.26 (and with the first beat anticipating $t = 0$ by a small amount). Since the period of the modulator is exactly three times the period of the carrier, or basic pulse, the pattern of beat lengths - which are timing deviations – is exactly repeated every three beats, giving the aural impression of a 3/4 time signature. It is important to note that the length of the 3/4 "bar" thus defined is the same as if the three beats had been undeviated – globally, the tempo is unchanged. (Waadeland, 2001) This is different from the expressive rubato discussed in eg. (Beran & Mazzola, 2000; Bresin, 1998,2000; Bresin et al., 1992; Canazza et al., 1997; Canazza, Roda et al., 1999; Cannazza et al., 1998; Desain & Honing, 1996; Friberg, 1995; Friberg & Bresin, 1997; Friberg et al., 1998; Gabrielsson, 1982; Honing, 1992; Mazzola & Zahorka, 1994a,b; Repp, 1992,1998; Shaffer, 1980; Stange-Elbe & Mazzola, 1998; Todd, 1985,1989; Todd, 1992b; Widmer, 1994,1996,2000,2001), where the global tempo fluctuates in the service of musical expression (mainly in performances of Western notated art music). Waadeland introduces the condition $0 \leq d \leq f/f'$ to ensure that the deviated rhythm has the same number of pulses as the undeviated rhythm over any interval of length $2\pi$. Waadeland also gives examples of a rhythm deviated through frequency modulation where the timing deviations result in one of the quarter notes being actually realised with a *shorter* duration than the following eighth note. (Waadeland, 2001)

Waadeland and Saue's computer program implementation of the Rhythm Frequency Modulator (in C++) calculates the onset times of a given MIDI file prior to playback of the (now-deviated) file. (Saue, 2000; Waadeland & Saue, 1999) The parameters of the RFM-instrument include a choice of modulating waveforms (sine, triangle, sawtooth and square) and permit a natural number exponent of the modulating waveform. (Waadeland, 2001) This last will effectively alter the shape of the modulating waveform (since $\sin^2 (x) = \sin(\sin(x))$ ). Both the choice of modulating waveforms and the ability to introduce an exponent of the modulating waveform will compensate in some degree for the limitations of simple FM (one carrier and one modulator). Settings for carrier and modulator frequencies, modulator phase, and peak frequency deviation can also be made.

Waadeland gives parameter settings which produce typical deviation patterns found in eg. waltz rhythms (short-long-medium); the Norwegian “springar” folk dance (different deviation patterns in different regions of Norway: long-medium-short in Telemark, short-long-medium in Valdres); and varieties of jazz swing (ranging from a 2:1 ratio to close to a 1:1 ratio for the two notes). (Waadeland, 2001) Different parameters may be applied to different instrument tracks in the same composition, resulting in asynchronies between instruments as found in jazz performances. (Waadeland, 2001)

Waadeland’s model provides a way of generating deviation patterns with systematic interdependency between the deviations within any individual pattern, the only model for microtiming deviations which I have found with this ability. The deviation patterns which Waadeland’s model is described as generating (Waadeland, 2001) are musically realistic, if somewhat simple, being confined to variations in onset times of at the most three beats (in 3/4 time dance music). It seems likely that (as Waadeland notes (Waadeland, 2001)) more complex FM implementations involving various configurations with more than a single modulator could produce more complex patterns of timing deviations. If it could be shown that this approach was capable of producing timing deviation patterns similar to those found in music with longer and more complex repeated rhythm patterns – such as over the course of the bell reference pattern timeline found in some traditional African drumming; or the length of the repeated riff in funk music – the rhythm frequency modulation approach would have a clear application to groove.

*Table 6-1 Feature comparision of microtiming deviation models.*

| | How deviations are generated | User-set parameters to generate deviations | Other comments / features |
|---|---|---|---|
| **Cholakis's (Cholakis, 1999b) Groove Templates** | Pre-set, with variations provided in separate files (for use in MIDI sequencers) | Choice of files to load | Provides ready access to a variety of feels, in the style of a particular well-known drummer, or genre. Based on data extraction from recordings, and not designed to generate deviation patterns. |
| **Bilmes's (Bilmes, 1992,1993a,1993b) Xcite drum machine software** | Set individually and manually per tatum | Individual tatum deviations manually set for each patterns. Patterns can be preset and loaded on the fly | Strong flexibility, especially for live performance using the full-featured drum machine software. Would appear to require extensive preparation prior to performance (particularly setting deviation patterns). |
| **Waadeland's (Waadeland, 2001; Waadeland & Saue, 1999) Rhythm Frequency Modulator** | Deviations generated by frequency modulation of the tempo, with the period of the modulator at some multiple of the macroperiod | User chooses modulator waveform and modulation strength, for simple FM. The number of parameters increases rapidly with more complex FM algorithms | The only model (except for the VROOGE model) which actually generates new deviation patterns. Difficulties arise in trying to generate a desired deviation pattern. Since deviation amount is the integral of tempo modulation, a sine modulator wave produces only negative deviations. Concatenation of different modulator waves during the course of the macroperiod would be required to produce negative and positive deviations (eg. a sine wave at half the macroperiod, followed by a negative sine wave). Results are difficult to predict as algorithm complexity increases. |
| **The VROOGE Model** | Deviations generated step-shifting the tempo of sections of the macroperiod, to simulate temporal expansion and compression of motor programs | User chooses number of covert clock subdivisions per macroperiod and relates a tatum to each subdivision. Percentage strength of influence of covert clock is chosen once for a given pattern (but can be varied). | Deviation generation allows for intuitive experimentation by user. The model also entrains to an external beat (utilizing period and phase correction), and so is suitable for use in live performance, with a human drummer.<br>The VROOGE model is unique in its attempt to model psychological and musicological processes underlying the generation of microtiming deviations. |

## 6.3 Development of the VROOGE model

### 6.3.1 The basic model for repetitive timing and correction processes

As noted in (Large et al., 2002), models of entrainment generally depend on the concept of an intrinsic frequency to give the period of rhythmic behaviour, with some degree of independence from the stimulus periodicity. The basic two-process model for the timing of repetitive discrete motor responses introduced by Wing and Kristofferson (Wing & Kristofferson, 1973a; Wing & Kristofferson, 1973b) has been influential in the literature. A central timekeeping process triggers motor responses. Neuromuscular transmission lags and the time taken for physical movement introduce a delay between the clock trigger and the motor response; and independent noise in both the central clock and the motor processes introduces variance to the timing of the inter-response interval (IRI).

The Wing and Kristofferson (Wing & Kristofferson, 1973b) basic equation is

*Equation 6-6*

$$I_n = C_n + (M_{n+1} - M_{n-1})$$

where $I_n$ are the interonset intervals, $C_n$ are the internal clock intervals, and $M_n$ are motor output delay intervals. The equation allows for both a noisy clock and a noisy motor effector – each $C_n$ and each $M_n$ may be varied according to the clock noise and motor noise respectively.

The fundamental model described by Wing and Krisofferson includes neither period correction nor phase correction and is intended only as a model of repetitively-timed actions – it provides the basic, independent rhythmic behaviour which is the basis for models of entrainment to an external stimulus. The VROOGE model, in common with models by various researchers (e.g. (Aschersleben, 2002; Madison & Merker, 2002; Mates, 1994a,b; Mates & Aschersleben, 2000; Mates, Müller, Radil, & Pöppel, 1994; Mates et al., 1992; O'Boyle & Clarke, 1996; Pressing, 1998a; Pressing, 1998b,1999; Pressing & Jolley-Rogers, 1997; Repp, 2000,2001b; Semjen et al., 2000; Thaut & Miller, 1994; Thaut, Tian et al., 1998; Thaut & Kenyon, 2003; Thaut, Miller, & Schauer, 1998)) retains the concepts of the clock and motor response.

### 6.3.2 The introduction of correction processes to the basic two-process model

For entrainment to occur, error correction processes must be introduced to the basic two-process equation (Equation 6-6). Error correction for entrainment has only one perceptual reference on which to base its process: the asynchrony between the stimulus and the produced response. Error correction processes for the internal clock can alter either its period, or its phase.

As Repp (Repp, 2001a) demonstrates, the behavioural consequences of phase correction of the internal clock and period correction correction of the internal clock, in terms of the following observable asynchrony, are indistinguishable. In fact, the stimulus which produces corrective behaviour in the model - a perception of a change in asynchrony - is itself ambiguous, in the sense that there is no way of knowing whether the first change in asynchrony represents a phase change or period change in the stimulus train. A phase shift (say, of 100ms in the positive direction) which affects all subsequent stimulus onsets could theoretically elicit period change in the model (a longer period, then a return to the original period), although a one-off phase shift would seem more economical. Conversely, a step tempo change (say, adding 100ms to the initial period) could in theory be accommodated by an ongoing series of phase shifts, although in this case a corresponding period change would probably be expected. Any model incorporating separate phase and period correction will have to deal gracefully with these two stimulus conditions and a variety of others.

Vorberg and Wing (Vorberg & Wing, 1996), extend the fundamental two-process model of Wing and Kristofferson (Wing & Kristofferson, 1973a; Wing & Kristofferson, 1973b) by the addition of a linear feedback mechanism to adjust for phase differences. A fixed proportion of the last synchronisation error is subtracted from the timekeeper interval.

### 6.3.3 Pressing’s Referential Behaviour Theory

By contrast, Pressing (Pressing, 1998a; Pressing, 1998b,1999) extends and elaborates the Wing and Kristofferson (Wing & Kristofferson, 1973a; Wing & Kristofferson, 1973b) model to produce an entrainment model with *period* correction. Correction processes in Pressing’s model are based on sine basis functions, with Fourier harmonics to account for subdivisions of the stimulus period. Pressing accepts the possibility of some control by expert musicians of phase shifts (Pressing,

1999) but does not introduce error correction for phase independent of period: rather, adjustments to the clock interval are made to minimise phase asynchronies.

The low parameter dimensionality of Pressing's model makes it transparent. The transparency of the model (by contrast, for instance, with the "black box" approach of neural networks) is an important feature for a music production model of microtiming production where user generation of new microtiming patterns is a desired goal, as it enables inspection of computational processes underlying microtiming deviation.

The error correction (or control) function for Pressing's (Pressing, 1998b) behavioural model is a scaled sine function:

*Equation 6-7*

$$A^*_{n+1} = A^*_n - (\alpha P/2\pi)\sin[2\pi A^*_n/P] - (\beta P/2\pi)\sin[2\pi A^*_{n-1}/P] + Q_n$$

where P is the fundamental period of tapping, $A^*_n$.is the asynchrony between stimulus and response at beat *n,* and $Q_n$ is clock and motor noise combined*.* The error-correcting sine functions are scaled by the period, and by the first- and second-order error correction parameters, $\alpha$ and $\beta$ respectively.

The use of the sine function for error correction might be taken to suggest the presence of an oscillator which can provide phase information during its period; rather than a clock or interval timer which provides only a trigger at the onset of its period. However, Pressing intended RBT to model human entrainment behaviour, and did not intend to create a plausible model of underlying processes. (Pressing, 1998b) The sine function is convenient for error correction in entrainment, first, because it provides a sign (positive or negative) for the asynchrony. A response tap may occur slightly early or slightly late, relative to the stimulus time. Without some means of determining which, error correction will always assume either that the response is early (so that slightly late responses will result in a large and inappropriate positive correction), or late (when the opposite will occur). For this pragmatic reason, the VROOGE model uses a sine-based error correction process.

The use of a sine-based process also means that correction will reach a maximum at one quarter the value of the period and reduce to zero at halfway through the period. However, the zero value of error correction halfway through the period is not stable – it is a repeller, as error values around the halfway point will correct back to the stable attractor where the phase error is zero.

Accepting the use of phase information, for practical reasons, we divide the whole of Equation 6-7 by P and define $\theta_n = A^*_n / P$, so that:

*Equation 6-8*

$$\theta_{n+1} = \theta_n - (\alpha/2\pi)\sin[2\pi\theta_n] - (\beta/2\pi)\sin[2\pi\theta_{n-1}] + \xi_n$$

where $\xi_n = Q_n/P$. (Pressing, 1999)

The introduction of 2:1 tapping – a binary subdivision of the tactus – would require error correction terms at the second harmonic of the fundamental frequency, and with phase lags at half the fundamental period (Pressing, 1999); however, in the VROOGE model, only entrainment to the tactus is required.

Despite in-depth analysis of the nature of the separate contributions of clock and motor noise (Pressing, 1999), Pressing's model does not rely, in its error correction processes, on the two-process paradigm introduced by Wing and Kristofferson. As with the Vorberg and Wing model described above (Vorberg & Wing, 1996), adjusting the clock period in response to each asynchrony erases the distinction between phase and period correction, since it effectively undermines the concept of a stable internal clock. With the introduction of microtiming deviations in the VROOGE model, a stable internal clock will be an essential requirement, so that the performance or motor implementation of the rhythm can both reference the stable period, and depart from it where required.

To allow for microtiming deviations – variations in the phase asynchrony for expressive purposes - within a stable tempo, a model which includes both period and phase correction as separate processes will be required. In such a model, the separate treatment of clock (or central) and motor (or implementation) processes will be crucial. An extension of the original Wing and Kristofferson (Wing & Kristofferson, 1973a; Wing & Kristofferson, 1973b) two-process model allows the location of the period correction process in the clock (internal) area of the model; and the phase correction process in motor (implementation) area of the model.

Once the clock and the motor implementation are separated, the central clock requires an error correction process which results in correct (or nearly correct) frequency for the internal clock. The stability of the internal clock provides the basis of microtiming variability in implementation

### 6.3.4 Simultaneous phase and period correction

Entrainment models incorporating separate phase correction and period correction terms have been described by Large and Kolen (Large & Kolen, 1994), Mates (Mates, 1994a), Thaut et al (Thaut, Miller et al., 1998) and Large and Jones (Large & Jones, 1999).

Thaut et al (Thaut, Miller et al., 1998) point out that the interdependence of the phase and period timing mean that the original phase relationship can only be restored following an unanticipated step change in ISI duration by a temporary ‘overcorrection’ – because changing the IRI duration to simply match the ISI will not correct for the phase change introduced by the first and any other ISI periods before the IRI duration was adjusted. Accordingly, their synchronisation model is a real two-process model, with separate terms for clock period and synchronisation error (taking into account the normal negative asynchrony found in synchronisation tapping tasks and discussed above, synchronisation error becomes naturally associated with the implementation, or motor, process in the two-process paradigm). Thaut et al’s model also includes separate (linear) correction terms for both period and phase. (Thaut, Miller et al., 1998) Note that the model does not attempt to separate conditions under which either of the phase and period correction processes might be activated without the other. In fact, their model could predict initial overcorrection - not just for the IRI to restore the original phase asynchrony, but overcorrection of the phase asynchrony as a result of the two correction processes working together - although this will depend on the gain parameters for the two correction processes. In fact, phase asynchrony overcorrection is found in the +50ms condition but not in the –50ms condition.

### 6.3.5 Period correction triggered by different conditions from phase correction

Before attempting to specify separate error correction processes for period correction (of the internal clock) and phase correction (located in the motor implementation area), it is important to establish how the two processes respond same asynchrony conditions.

The existence of phase corrections in response to subliminal timing perturbations (below detection threshold for ISI changes) has been established in a number of studies. (Repp, 2000; Repp, 2001a; Repp, 2001b,2002a,c; Thaut, Tian et al., 1998; Thaut & Kenyon, 2003) In all, there is strong evidence that phase correction occurs in response to timing perturbations below the detection threshold of about 4% of ISI, and definite evidence that period correction does not occur below this threshold. Apart from studies by Repp (Repp, 2002a,c) in which subjects were required to attempt to suppress corrections in response to event timing shifts or permanent phase shifts, there is no concrete evidence that phase correction disappears above the detection threshold.

The strongest evidence for engagement of phase correction below the perturbation detection threshold, and for period correction above that threshold, comes from a series of experiments performed by Repp (Repp, 2001a). Repp's first experiment introduced a small timing perturbation at the end of the stimulus sequence in a synchronisation-continuation task. Either the last ISI was changed only (a phase shift in the stimulus train, since no further tones were heard), or the last three or the last five were changed (a step tempo change in the stimulus train). The introduced changes – at were +10ms or –10ms or 0ms - were below the threshold of detection and participants reported not noticing any ISI changes. Where only the last ISI was changed, the change was echoed immediately but IRIs during continuation tapping returned to the baseline period of the stimulus train - indicating a probable phase correction response. (Repp, 2001a) Where the last three ISIs were changed, the IRI adapts to the new period after a lag of one position and the adaptation partially persists during continuation tapping – indicating a probable period correction response. A stronger adaptation effect appears in the condition where the last five ISIs are changed, although the full 10ms correction to the period is still not evident in continuation tapping. Repp concludes that rapid phase correction is followed by more gradual period correction.

The threshold for the crossover to where the larger effects (presumed to be the result of period correction) are felt is variously given as 5% to 7% of ISI (Thaut, Tian et al., 1998) and 2% of ISI (for musically experienced subjects) (Repp, 2001b).

### 6.3.6 Separating phase and period correction processes

To write an equation for motor implementation phase correction separate from internal clock period correction, it is necessary to clarify exactly the timing of the motor response to the (auditory) stimulus. In the situation where synchronisation is considered to have been achieved, the sensorimotor goal can be taken to be synchronisation of auditory response feedback (the sound of the drum) to the auditory stimulus, with variance within the mean given by experimental studies.

Let:

*Equation 6-9*

$$M^{*} = M - S$$

where $M_j$ is the time of the auditory response feedback onset and $S_j$ is the time of the auditory stimulus onset. For a human being, there will be a delay between the clock trigger which initiates the tapping action and the arrival of auditory feedback from the tap in the central representation area, resulting from the cumulative delay of the nerve path from brain to muscles, the time taken to complete the physical action, the time taken for the sound to travel through the air to the ear drum, and finally the delay from ear drum to central representation area. In cases where there is no external driving stimulus, such as when a drummer is playing alone, S can be taken to be the expected time of the response feedback. In that case, $M^{*}$ will vary as a result of motor noise, or any introduced timing perturbation which does not affect the central clock.

Where all aspects of the model (including sound generation) are contained within software, there will be a shorter delay produced by the time taken for the clock to trigger a MIDI event, and for the MIDI event to return to the synchronisation process – no more than 2 or 3ms. However, if events were triggered at a MIDI module external to the computer and a MIDI note-on message was mirrored back, the model would adjust to the additional latency without difficulty, since it works on the actual asynchrony between the stimulus and the note-on as their respective signal arrive at the synchronisation process. A longer latency would in practice result in a correction to increase the asynchrony between the clock trigger and the stimulus time, thus accommodating the latency. The stimulus itself will originate as an external trigger generated either as a MIDI message from eg a keyboard, electronic drum, or drum trigger. The model will be regarded as successful if it succeeds in reducing $M^{*}$ to within the mean for variance, after an appropriate number of beats.

Note that synchronisation of the auditory response to the auditory stimulus allows an unspecified latency between the trigger from the internal clock (the start of the clock interval) and the arrival of the produced auditory response as sensory input in the central representation area.

Since error correction for the motor implementation phase is now divorced from clock error correction, it is feasible to consider motor or implementation phase in relation to tatum subdivisions, while clock period remains referenced to the tactus. Assuming linear error correction for phase (without deviations for the moment), and first-order correction, the *implementation phase correction* equation is:

*Equation 6-10*

$$M^*_{j+1} = M^*_{j-1} - \alpha_1 M^*_j + Q_{Mj+1}$$

Where $M^*_j$ is auditory stimulus/response feedback asynchrony for the j$^{th}$ tatum, $\alpha_1$ is the phase correction gain parameter, and $Q_{Mj}$ is the motor noise for the j$^{th}$ tatum production. First-order correction only is assumed, on the basis that phase correction has been shown (Repp, 2001a; Thaut, Tian et al., 1998) to be a subconscious process, and therefore less likely to utilise remembered information.

In fact, Equation 6-10 is already capable of producing microtiming deviations of a sort, through the interaction of the noise term $Q_{Mj+1}$ and the error correction term. However, applying error correction to the noise effects would not produce the systematic patterns found in analysis of musical examples, such Freeman's (Freeman & Lacey, 2001) analysis of the Funky Drummer break by Clyde Stubblefield on James Brown's In The Jungle Groove album (Brown, 1986).

Assuming the sine function still provides the clock period, the *clock period correction* process becomes:

*Equation 6-11*

$$\theta_{n+1} = \theta_n - (\alpha_2/2\pi)\sin[2\pi\theta_n] - (\beta_2/2\pi)\sin[2\pi\theta_{n-1}] + \xi_n$$

where $\theta_n = M^*_n / P$, and $M^*_n$ is the asynchrony between the n$^{th}$ tactus audible response onset and the n$^{th}$ tactus auditory stimulus onset. $\alpha_2$ and $\beta_2$ are respectively first- and second-order period correction parameters, and $Q_{cn}$ is the clock part of the total noise.

In accordance with the evidence that the total of the first- and second-order parameters is constant, $\beta_2$ can be replaced by:

*Equation 6-12*

$$\beta_2 = \lambda - \alpha_2$$

where $\lambda$ is the constant total of $\alpha_2$ and $\beta_2$ , giving:

*Equation 6-13*

$$\theta_{n+1} = \theta_n - (\alpha_2/2\pi)\sin[2\pi\theta_n] - ((\lambda - \alpha_2)/2\pi)\sin[2\pi\theta_{n-1}] + \xi_n$$

Note that tactus asynchronies $M^*_n$ are a subset of tatum asynchronies $M^*_j$. There is as yet no definite evidence whether tatum timing perturbations trigger clock period correction, independently of any nested clocks. Restricting period correction to tactus asynchronies simplifies the model.

It is a feature of the VROOGE model that period and phase correction are separate processes, located in separate areas of Wing and Kristofferson's two-process model and involving a central clock and separate motor implementation (Wing & Kristofferson, 1973a; Wing & Kristofferson, 1973b). Hence, the two processes are independent, except for the trigger from the clock to the motor implementation process, at the start of each clock interval:

Note that focus of the VROOGE model is the separation of period and phase correction, to allow microtiming deviations while maintaining a stable internal reference clock. No attempt is made to model tatum subdivisions of the tactus in a psychologically plausible way.

### 6.3.7 First- and second-order correction processes

Pressing (Pressing, 1999) observes that higher-order control is more difficult for humans than lower-order control. In continuous (as opposed to discrete) models, zero-order control in a spatial task is based on position, first-order control on velocity, second-order control on acceleration, and so on. In the Referential Behaviour Theory (RBT), (Pressing, 1999) control based on position is called 'first-order', because the effect of correction is delayed until the next response, which is also the next sampling point. In an isochronous tapping synchronisation task, both the frequency (equivalent to *velocity)* of the central clock, and its phase (corresponding to *position)* are relevant, yet error correction for both must be based on one item of information per clock cycle, the relative asynchrony between the reference audio tone and the tap – which amounts to phase, the lower order variable. [Note that I am assuming the train of audio tones has a period within the range of the tactus or spontaneous tapping. Pressing (Pressing, 1999) explicitly assumes a period between about 500ms and 1000ms, which is within the range of spontaneous tapping. Tatum subdivisions will be dealt with below.] In a continuous model, frequency is the derivative of phase over one clock cycle. In RBT, second-order error correction references asynchrony values from the current and the preceding clock cycle. It follows that frequency correction would be most accurately achieved by referencing the *difference* between successive phase values, as in Mates's (Mates, 1994a,b) model – in other words, by making the second-order parameter negative.

A summary of estimates of first- and second-order gain parameters (by parameter estimation of computational models to fit experimental data) is given in Table 6-2. It should be noted for all results regarding first- and second-order correction, that the contribution of phase correction and period correction processes are unclear in each case. Two assumptions will impact on the interpretation of the data: first, that period correction is absent below the Weber fraction for tempo change detection (approximately 5%); and second, that period correction is triggered only for changes to the tactus onset stimulus, while phase correction will occur in response to perturbations in timing for subdivisions to the tactus as well.

To complicate the picture, the possibility of nested clocks for longer tactus periods must be considered – presumably a change to a subdivision which falls on a nested clock pulse will trigger period correction (see earlier discussion). For instance, the very small second-order values found by Semjen et al 1998 and by Semjen et al 2000 for 735ms period and 640ms period respectively might be explained as the result of the joint correction between the longer tactus and a nested clock. (In this interpretation, the.18 value for the 682ms period in Semjen et al 1998 is an anomaly). Presumably this effect falls away for periods below 640ms in the table because half-period clocks are not available at faster tempi.

Data from Repp's (Repp, 2002b) subdivision study support this account – the difference in correction in response to perturbations on the subdivision or on the beat of a subdivided stimulus interval is less for 540ms period than a 720ms period, possibly due to the influence of the hypothesised half-period clock in the 720ms condition.

Similarly, the very low second-order values for longer periods might reflect the absence of a longer-period clock. Assuming a narrower range of possible clock periods for individuals than found across the whole population, the period of an individual's spontaneous tempo could reflect a value at which either a half-period or double-period clock could be engaged, as a stabiliser to the main tactus. Despite these considerations, actual tactus periods between 486ms and 534ms in the table do show (small) second-order values.

The data from the various studies is in broad agreement in two basic areas. All studies found first-order correction tends to zero at lower intervals. There is some disagreement about the actual interval. Pressing and Jolley-Rogers (Pressing & Jolley-Rogers, 1997) give 150ms as the interval at which first-order parameter values begin to decline and Thaut and Kenyon (Thaut & Kenyon, 2003) claim

*Table 6-2 Experimentally determined gain parameter estimations for entrainment models*

| | **Short-interval / first order error correction gain** | **Long interval first order error correction gain** | **Short-interval / second order error correction gain** | **Long-interval / second order error correction gain** |
|---|---|---|---|---|
| Pressing and Jolley-Rogers 1997 (Pressing & Jolley-Rogers, 1997) | ~150ms / progressively less at smaller intervals | | ~150ms / negligible above this interval | |
| Semjen et al 1998 (Semjen, Vorberg, & Schulze, 1998) | 682ms / .18<br>534ms / .26<br>535ms / .25<br>429ms / .20<br>486ms / .29<br>*Mean: .24* | 735ms / .58 | 682ms / .24<br>534ms / .03<br>535ms / .15<br>429ms / .15<br>486ms / .16<br>*Mean .14.6* | 735ms / .04 |
| Semjen et al 2000 (Semjen et al., 2000) | 200ms /.06 | 640ms /.51 | | 640ms / tends to zero |
| Repp 2001 (Repp, 2001a) | 250ms / as strong as for 500ms | | | |
| Repp 2001 (Repp, 2001b) | 500ms / .06 (phase correction); period correction 0.25 (where timing perturbation not detected) | 0.55 (where timing perturbation detected); | | |
| Repp 2002 [values are current author's estimates from actual corrections shown in (Repp, 2002b)] | 135ms / ~0<br>180ms / .1<br>270ms / .4<br>360ms / .6 | | 270ms / .21<br>360ms / .25<br>[see note in text, above] | |
| Thaut and Kenyon 2003 (Thaut & Kenyon, 2003) | 200ms/ adaptation to a +/-2% step change in ISI | | | |

Where no value is shown for second-order parameters, first- and second-order values were not distinguished adaptation at 200ms, but Semjen et al estimate the first-order parameter reduced to .06 at 200ms, and Repp's data (Repp, 2002b) appear to indicate low values of 0.1 at 180msec and 0 at 135msec. Despite these discrepancies, and allowing for individual differences, it appears clear that first-order correction disappears as the beat period becomes smaller, somewhere between ~200ms and ~135ms.

Second, if small-interval-low-gain data and long-interval-high-gain data are excluded, there is general agreement that the total gain (first-order and second-order parameters values added) is in the region of 0.4.

Overall, available data supports a total gain of ~0.4, distributed between first- and second-order parameters, with first-order values tending to zero at some value below 250ms and achieving their maximum at about 640ms periods.

### 6.3.8 Entrainment latency

A final consideration in evaluating the simple success of an entrainment model (as opposed to whether its behaviour is consistent with human experimental data) is *entrainment latency* or *entrainment lag.* Thaut and Miller (Thaut & Miller, 1994) found that a stable synchronisation response to a stimulus train is attained within 2 to 3 stimulus repetitions. Repp (Repp, 2001b) quotes Michon (Michon, 1967) as finding synchronisation to the new tempo for an 8% or more tempo step change was generally achieved within 4-5 taps. This is consistent with the findings of Repp's (Repp, 2001b) experiment, for step changes of 3% to %5 of an initial 500ms inter-stimulus interval (ISI). Asynchronies for these conditions actually overshot the baseline in subsequent taps. Step changes less than 2% required returned to the baseline asynchrony much more slowly.

### 6.3.9 Adding microtiming deviations to the model

Given a central clock (which may, depending on the musical situation may be referenced to an auditory input such as a reference bell pattern) we now need a model to allow constrained variability of phase during motor implementation, while referencing the central clock.

Microtiming deviations made for expressive purposes should not engage the period correction process, as it is essential the clock period remains stable. Repp (Repp, 2000; Repp, 2001a) has established that period correction is very small or nonexistent for timing perturbations which are less than an amount approximately equal to the Weber fraction for tempo discrimination – approximately 5% of IOI.

Accordingly, a parameter $\eta$ is introduced into the period correction process, to reduce response to zero, for perturbations below this level:

*Equation 6-14*

$$\theta_{n+1} = \theta_n - \eta_n(\alpha_2/2\pi)\sin[2\pi\theta_n] - \eta_{n-1}((\lambda - \alpha_2)/2\pi)\sin[2\pi\theta_{n-1}] + \xi_n$$

where $\eta_n$=1 when $M^*_n$ is greater than an amount approximately equal to the Weber fraction for tempo discrimination, and $\eta_n$=0 for smaller values of $M^*_n$.

In practice, microtiming deviations found in groove music normally range up to about 5% of period, with only a few individual deviations in some examples exceeding the 5% level by a significant amount (see discussion in Chapter 6), so a value of ~.05 for $\eta$ will probably produce the desired result. However, Pressing (Pressing, 1999) notes that adaptation must include the possibility that over time some initial terms in the equation may fall to zero, and suggests that such changes can be handled formally through a master equation of control for the system. (Pressing, 1999) Increasing the value of $\eta$ - perhaps over time as the value of C stabilises – would allow greater phase difference without affecting the underlying clock period.

### 6.3.10 Two different approaches to generating microtiming deviations

Systematic deviations may be introduced into the phase relationship through direct control of phase, or through control of period, as with Waadeland's (Waadeland, 2001; Waadeland & Saue, 1999) 'Rhythm Frequency Modulation'. An implementation of the microtiming deviation model incorporating frequency modulation has in fact been completed, but several drawbacks were found. Estimating parameters to reproduce deviation patterns from real-world musical data proved difficult in most cases (see discussion, below). The musical results of modulator phase and period parameter settings could not be intuitively predicted. A simple frequency modulation (one modulator) arrangement appeared insufficient to model the deviation patterns studied and parameter estimation for more complex models would have been correspondingly more complicated. In all, rhythm frequency modulation did not satisfy the goal of producing a musically intuitive tool, capable of both emulating deviation patterns gleaned from analysis of musical examples, and generating new patterns in accordance with an underlying principle transparent to the user. Nevertheless, the rhythm frequency modulation approach does produce systematic microtiming deviations, which may be musically useful.

A second approach relies on metrical subdivision of the macroperiod with a meter conflicting with the main, or overt, meter –the Covert Clock approach. For example, a macroperiod of four tactus beats, each subdivided into four tatums, is overlaid with a covert clock of three tactus beats without tatum subdivisions. Both period correction and error correction processes are active in the covert meter , and mutual entrainment between the meters occurs.  Both meters (the overt tactus and the covert clock) have a downbeat at the start of each macroperiod, and the total effect is to stabilise the duration of the macroperiod. Tactus beats (and their tatum subdivisions) of the main meter are shifted in the direction of the closest tactus beat of the covert clock. The shifted tactus triggers motor implementation.

The choice of covert clock macroperiod subdivision (eg the choice of a cross rhythm of three tactus pulses against the main meter's four tactus pulses per macroperiod) will be determined externally , and will in fact be  a user parameter.

# Chapter 7 The VROOGE software

The name *VROOGE* is an anagram of 'groove'. The VROOGE software (constructed in Cycling 74's Max/MSP visual programming environment) consists of an ENTRAINMENT patch, a SEQUENCE patch, and Covert Clock deviator mechanism located in the top-level patcher.

The version of the software provided loads automatically with the *Macota, Bula II* sequence discussed in Chapter 5 of the thesis, at 5.3.2. The parameter settings for this sequence load automatically, so that the user unfamiliar with the software can run it easily.

The VROOGE software runs only under OS 9, on a Mactinosh computer.

## 7.1 Entrainment

The ENTRAINENT patch synchronises to a *driving tactus pulse*, and is an implementation of the entrainment model described in Chapter 6 of this thesis. It contains a central clock period correction process, separate from the motor implementation phase correction process. The motor implementation phase correction process allows for some *functional anticipation* (see Chapter 2 of the thesis), the amount of which is not predetermined. The amount of *functional anticipation* will vary to accommodate longer clock-trigger-to-motor-implementation delays (for instance, if an external MIDI module was involved).

The driving tactus pulse for the entrainment pulse is provided in this version of the VROOGE software by a metronome incorporated into the patch. The period of the driving metronome is calculated from user-set parameters of the *number of tactus beats* in the macroperiod, and the *duration of the macroperiod.* The correction processes operate to bring the period of the *internal clock* usually within 1ms of the period of the driving pulse: however, if there is a dramatic change from one period to another, the clock may lock to a sub- or super-harmonic of the tactus frequency – in this case, a 'cognitive' intervention is required (the user should double-click the ENTRAINMENT patch to open it, and set the current clock period close to the period of the driving pulse.)

Since playback of onset timings is achieved in the software by varying the speed at which the buffer of onset times is played back, the ENTRAINMENT patch sends a 'global tempo' parameter to the SEQUENCE patch, calculated from the current period of the internal clock (which is found using error correction processes, in response to the driving pulse). This is a pragmatic solution in software terms, and does not invalidate the theory of timed clock triggers and motor programs, described in the thesis.

The ENTRAINMENT patch can also entrain to an external driving pulse from a MIDI or audio source (such as a drum).

## 7.2 Sequence

VROOGE stores onset timings in an audio buffer, which is read back at a variable rate, according to a *global tempo* parameter calculated from the current period of the entraining clock; and a *local tempo* parameter which provides the microtiming deviations.

## 7.3 Microtiming deviations

Microtiming deviations are calculated according to the scaled differences between paired sequence tatum timings and covert clock timings. Local tempo variations and the loop index - the point in the macroperiod - where these should be applied, are calculated.

## 7.4 User-set parameters

Parameters for the Macota, Bula II sequence (discussed in the thesis at 5.3.2) are loaded automatically when the software is opened. The user should first turn audio on. The *number of tatums in the loop* and the *number of tactus in the loop* (or macroperiod) should be set, and the *duration of the macroperiod*. The number of *covert clock onsets in the loop* must be set.

Sequence tatum onsets must be paired with covert clock onsets: this sets the microtiming deviations and the local tempo changes required to achieve them. A graph of the deviations appears at the right of the VROOGE screen.

Click the green START button to begin playback (with microtiming deviations); click the button again to stop.

## 7.5 Evaluation of the VROOGE software

A graph of the microtiming deviations output by the VROOGE software for the Bulá 2 part for the Macota toque of the Yubá dance is shown at Figure 7-1. The results of the application of the Covert Clock Theory to the same part are shown at Figure 7-2. Both the Covert Clock theoretical application and the VROOGE model output are based on a Covert Clock of five (against the 12 tatums of the macroperiod), and are scaled by a factor of 0.29. It can be seen that the software produces deviations with broadly the same contour as the data from the recording. However, the deviations at tatums 7 to 10 vary more from the original than do those at tatums 1 to 3.

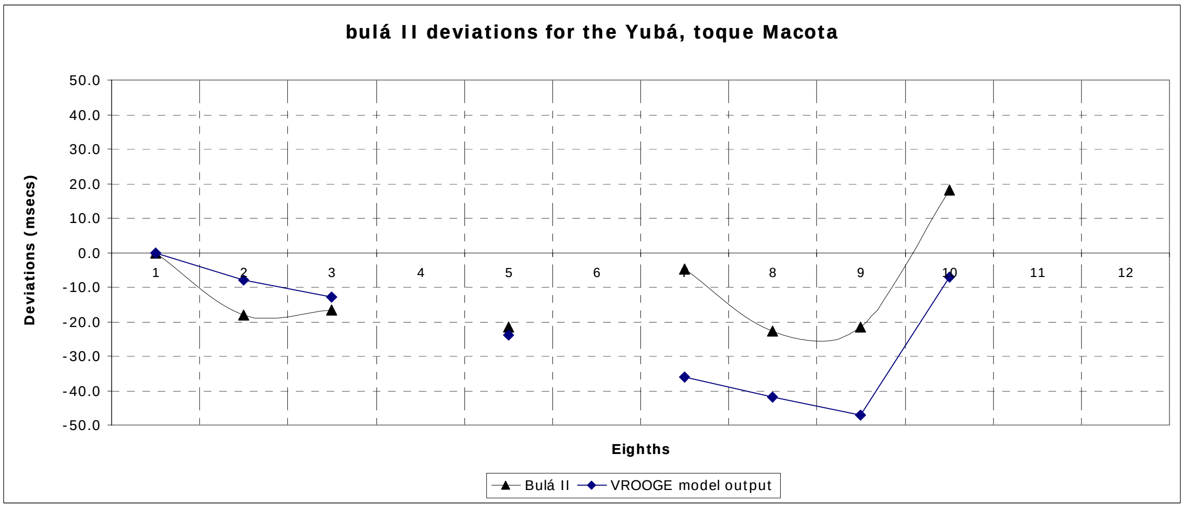


*Figure 7-1 VROOGE model output for the Bulá 2 part for the Macota toque of the Yubá dance*

*Figure 7-2 Covert Clock model results for the Bulá 2 part for the Macota toque of the Yubá dance. The model results are scaled by 0.29*

The model appears to reproduce the contour of some actual musical deviations, but is lacking in accuracy. It may be that scaling of the effect is not constant throughout the macroperiod, in human performers. In fact, such inconsistency is predicted by the strategy for timing cross rhythms of timing the fast hand to the internal clock trigger and the slow hand to appropriate delays following the central clock triggers - see (Summers et al., 1993), (Pressing et al., 1996), and section 4.2, above.

The software provides a simple and intuitive way to produce new timing deviations. A possible improvement might be the addition of a parameter to vary scaling over the macroperiod, within constraints. If some random variability was included in the control of that parameter, microtiming deviations produced by the model would vary to some extent with successive iterations of the macroperiod, giving a more human feel.

Shortcomings (as manifested in the discrepancies between the VROOGE model output, the theoretical application of the Covert Clock model, and the original data) may relate in part to a lack of clarity in how exactly the covert clock affects motor implementation. An essential unresolved question is whether only existing tactus triggers are shifted in time; or whether the trigger is moved to a different tatum count, which is shifted in time. Revision of the theory, model and software in relation to this issue might improve results.

# Chapter 8 Summary and conclusion

## 8.1 Impetus for research on microtiming and groove

I began research for this thesis with an interest in the connections between rhythm, the psychological present, and physical creativity – a kind of 'thinking with the body'. As a performing musician (a classical pianist), I had often had the experience of playing short passages with a harmonious integrity of expression which I had not consciously planned. These occurrences were marked by a fine subtlety of expression, particularly appropriate to the singularity of the current performance context, and by a strong subjective sense that the origin of creative expression in such instances was in the hands, rather than in the mind. In fact, I came to value and trust such occasions because the timing and pianistic touch was more finely balanced than I could achieve by conscious effort. Indeed, even the interpretation of the passages was usually a surprise!

This research, then, constitutes an effort to find the source of that felicity. I feel it has been a successful search – the identification of open loop motor planning in the context of planning within the space of the psychological present constitutes a concrete answer to the question posed, as well as suggesting possible topics for further investigation.

## 8.2 Summary of achievements

### 8.2.1 Clarification of the relationship between entrainment and planning processes in rhythm

The relation of entrainment processes and motor planning and implementation processes is made clear in the Covert Clock Theory. In normal entrainment, there exists the possibility of a system of hierarchical clocks at harmonic frequencies, which reinforce and stabilise each other. However, only one clock constitutes the tactus, which provides triggers for motor programs.

The Covert Clock Theory identifies motor programs at the tactus level as being under *open loop* control. The clear differentiation between timing processes relating to central clocks (entrainment) and timing processes relating to motor implementation is an original contribution of this research.

In the theory, the introduction of the covert clock itself sabotages any system of harmonic, hierarchical entraining clocks, reducing these to only the tactus clock. The central concept of the Covert Clock Theory – that a cross rhythm which is consonant with the tactus only at macroperiod boundaries serves to displace tactus trigger times and therefore the timing of the open loop motor programs triggered by the tactus – has not previously been suggested in the literature.

The Covert Clock Theory identifies the macroperiod as defining the extent of the psychological present for incremental planning purposes. As such, it provides an explanation for the clear marking of the macroperiod start in musics from a variety of cultures.

The available microtiming data has been analysed in ways which are unique to this investigation. Comparison of microtiming in successive iterations of the same pattern has been carried out for the first time, to my knowledge (in The Funky Drummer analysis). Direct comparison of microtiming for different parts of the same pattern is also a new approach. Two significant features which have appeared from analysis, not otherwise noted in the literature, are: that for different parts played at the same time by a single musician exhibit the same microtiming deviations; and that parts played for the same pattern by different musicians at the same time exhibit different microtiming deviations, although they appear to be different amplitudes of the same contour.

### 8.2.2 Integration of entrainment processes with microtiming deviation processes

A central problem of microtiming deviations is how a stable global tempo is achieved in conjunction with the local fluctuations of microtiming deviation. The Covert Clock Theory and VROOGE model are unique in integrating the period and phase correction processes of entrainment with microtiming deviation processes. Period correction is modelled as a central timing process, while phase correction becomes a motor implementation process, encompassing functional anticipation. Functional anticipation accommodates the latency inherent between initiating an action (such as a drum stroke) and hearing the result at the correct time, and may be associated with the performance of microtiming deviations – see below.

## 8.3 Outstanding issues for the current topic

There are a number of issues which this investigation raises which are directly relevant to the current topic.

- The precise relation between functional anticipation and microtiming deviations. If the tactus is assumed to remain invariant during groove, how in fact does the covert clock or cross rhythm work to create microtiming deviations? A possible explanation is that the covert clock timing point is used as a reference to vary the amount of functional anticipation in a controlled way, so producing microtiming deviations in terms of motor implementation of the rhythm, while leaving the tactus undisturbed.

- The possible duration of open loop planning, and whether some musical practices (eg. N Indian classical music) seek to extend this duration in performance.

- The relation between planning at the macroperiod level and open loop planning at the sub-tactus level. Repetition of the macroperiod, with covert clock cross rhythm, provides a repeatable framework in which subtactus planning can take place. However, some variation appears between iterations of the macroperiod. How this is dealt with has not been explored in depth, although some brief discussion of the issue in relation to The Funky Drummer is found at **5.2.2**.

- Exactly what subdivisions and rhythms are possible, and in what combinations, at the open loop motor program level. There are indications in the literature (Friberg & Sundström, 2002; Repp et al., 2002; Shaffer, 1981; Shaffer, 1985; Sternberg & Knoll, 1994) that the complexity of rhythms in individual motor programs is limited, although timing can be very precise.

## 8.4 Broader issues to be addressed

### 8.4.1 Issues for psychological research

Studies of rhythm production and perception in the psychological literature rarely distinguish between timing processes relating to entrainment, and those relating to motor implementation at the sub-tactus level. Experimental findings for rhythm timing might be clarified by taking into account the upper frequency limit of entrainment (ie. the lower limit for period of entrainment clocks).

Other broad issues, either for separate investigation or to be taken into account in experimental studies of rhythm, include: the range of possible rhythmic patterns, and timing resolution at sub-tactus level; the duration of open loop planning; and to what inputs open loop planning might respond.

### 8.4.2 Issues for musicological research

Similar issues might usefully be explored in musicological studies: the range of possible rhythmic patterns, and timing resolution at sub-tactus level; the duration of open loop planning; and how these relate and interact with the music of any individual culture.

In addition, the currently very limited range of microtiming data needs to be expanded, and preferably provided in a uniform format for easy comparison. A particularly useful resource would be microtiming data for successive iterations of the macroperiod within one performance (as with the data for The Funky Drummer (Freeman & Lacey, 2001) analysed in this thesis). Comparison of the iterations would enable more definitive conclusions to be drawn about the musical and psychological processes microtiming deviations. Similarly, a thoughtfully-collected body of data would allow comparison of different performances of the same work, by the same performer and by different performers, and cross-cultural comparisons of superficially similar rhythms.

The theories of entrainment and timing control underpinning the Covert Clock Theory of microtiming have not been tested with asymetrical and 'limping' meters found in Eastern Europe and Asia.

### 8.4.3 Issues for music technology

While a full implementation of the Covert Clock Theory might prove too clumsy for everyday use in music-making technology, the basic principles would be relatively easy to apply. Essentially, macroperiod boundaries remain undeviated. Within the macroperiod, a scheme is applied which is consistent for each musical example. A few timing points within the macroperiod are identified – these will be individually adjusted to be either slightly late or slightly early. Tatum timing between the identified timing points is adjusted so that subdivisions are even.

The amount of timing deviation should be globally adjustable: in this way, different parts in the same pattern can participate in the same microtiming scheme, while showing different amplitudes of the scheme. Random variation of the amplitude should be allowed, within a user-set range, to provide variation between successive iterations of the deviation scheme.

The advantage of a microtiming deviation tool as described would be that the user would be able to easily generate a variety of deviation schemes at will, without having to individually set tatum deviation parameters. Musically plausible schemes could be quickly generated and varied. Schemes closely approximating data drawn from recorded musical examples could be provided as parameter presets of the software.

# References

Alén, O. (1995). Rhythm as Duration of Sounds in *Tumba Francesca*. *Ethnomusicology, 39*(1), 55-71.

Anku, W. (2000). Circles and Time: A Theory of Structural Organization of Rhythm in African Music. *Music Theory Online, 6*(1).

Arom, S. (1991). *African polyphony and polyrhythm* (M. Thom, B. Tuckett & R. Boyd, Trans.). Cambridge: Cambridge University Press.

Arom, S. (1994). Intelligence in traditional music. In J. Khalfa (Ed.), *What is Intelligence?* Cambridge: Cambridge University Press.

Aschersleben, G. (2002). Temporal Control of Movements in Sensorimotor Synchronization. *Brain and Cognition, 48*, 66-79.

Aschersleben, G., & Prinz, W. (1995). Synchronizing actions with events: The role of sensory informmation. *perceptions & Psychophysics, 57*, 305-317.

Aschersleben, G., & Prinz, W. (1997). Delayed auditory feedback in synchronization. *Journal of Motor Behavior, 29*, 35-46.

Battel, G. U., Bresin, R., De Poli, G., & Vidolin, A. (1994). *Neural networks vs. Rules system: evaluation test of automatic performance of musical scores.* Paper presented at the Proceedings ICMC, San Francisco.

Becking, G. (1928). *Der musikalische Rhythmus als erkenntnisquelle*. Augsburg: Benno Filser; cited in Waadeland, 2001.

Beran, J., & Mazzola, G. (2000). Timing Microstructure in Schumann's "Träumerei" as an Expression of Harmony, Rhythm, and Motivic Structure in Music Performance. *Computers and Mathematics with Applications, 39*, 99-130.

Bilmes, J. (1992). *A Model for Musical Rhythm.* Paper presented at the ICMC.

Bilmes, J. (1993a). *Techniques to Foster Drum Machine Expressivity.* Paper presented at the Proceedings ICMC.

Bilmes, J. (1993b). *Timing is of the Essence: Perceptual and Computational Techniques for Representing, Learning, and Reproducing Expressive Timing in Percussive Rhythm.* Unpublished Masters Thesis, MIT.

Bresin, R. (1998). Artificial Neural Networks Based Models for Automatic Performance of Musical Scores. *Journal of New Music Research, 27*(3), 239-270.

Bresin, R. (2000). *Rule-Based Emotional Coloring of Music Performance.* Paper presented at the Proceedings ICMC, San Francisco.

Bresin, R., De Poli, G., & Vidolin, A. (1992). *Symbolic and sub-symbolic rules system for real time score performance.* Paper presented at the Proceedings ICMC, San Francisco.

Brown, J. (1986). In The Jungle Groove: Polydor.

Canazza, S., De Poli, G., Di Federico, R., & Drioli, C. (1999). *Expressive Processing of Audio and MIDI Performances in Real Time.* Paper presented at the Proceedings ICMC, San Francisco.

Canazza, S., De Poli, G., Drioli, C., Roda, A., & Zamperini, F. (2000). *Real-time morphing among different expressive intentions in audio playback.* Paper presented at the Proceedings ICMC, San Francisco.

Canazza, S., De Poli, G., Roda, A., & Vidolin, A. (1997). *Analysis by synthesis of the expressive intentions in musical performance.* Paper presented at the Proceedings ICMC, San Francisco.

Canazza, S., Roda, A., & Orio, N. (1999). *A Parametric Model of Expressiveness in Musical Performance Based on Perceptual and Acoustic Analyses.* Paper presented at the Proceedings ICMC, San Francisco.

Cannazza, S., De Poli, G., Di Sanzo, G., & Vidolin, A. (1998). *A model to add expressiveness to automatic musical performance.* Paper presented at the Proceedings ICMC, San Francisco.

Cholakis, E. (1999a). Clyde Stubblefield: The Funky Drummer. CD ROM of DNA Groove Templates (Version Serial No. 5007M): Numerical Sound.

Cholakis, E. (1999b). *Numerical sound: archeological sound engineering*. Retrieved August 26, 2003

Cholakis, E. (2003a). DNA Beat Templates process. Personal communication.

Cholakis, E. (2003b). Getz/Gilberto O Grande Amor. Personal communication.

Cholakis, E. (2003c). *Swing Analysis of 16 Jazz Drummers*

Chowning, J. M. (1973). The synthesis of complex audio spectra by means of frequency modulation. *Journal of the Audio Engineering Society, 21*, 526-534.

Clarke, E. F. (1985). Some aspects of rhythm and expression in performances of Erik Satie;s "Gnossienne No. 5". *Music Perception, 2*, 299-328.

Clayton, M. (1997). Le metre et le tal dans la musique de l'Inde du Nord. *Cahiers de Musiques traditionelles, 10*, 169-189 http://www.open.ac.uk/Arts/music/nclayton/metre.htm; trans. Georges Goormaghtigh.

Clynes, M. (1977). *Sentics, the touch of the emotions*. New York: Anchor Press/Doubleday; cited in Waadeland, 2001.

Clynes, M. (1983). Expressive Microstructure in Music, linked to Living Qualities. In J. Sundberg (Ed.), *Studies of Music Performance* (Vol. 39). Stockholm: Royal Swedish Academy of Music.

Clynes, M. (1986a). Generative principles of musical thought. Integration of microstructure with structure. *Communication and Cognition AI, 3*, 185-233; cited in Waadeland, 2001.

Clynes, M. (1986b). When time is music. In J. Evans & M. Clynes (Eds.), *Rhythm in Psychological, Linguistic, and Musical Processes*. Springfield, IL: Charles C. Thomas.

Clynes, M. (1987). What can a musician learn about microstructure from newly discovered microstructure principles (PM and PAS)? In A. Gabrielsson (Ed.), *Action and Perception in Rhythm and Music* (Vol. 55): Royal Swedish Academy of Music.

Clynes, M. (1992). *Time-forms, Nature's generators and communicators of emotion.* Paper presented at the IEEE International Workshop on Robot and Human Communication, Tokyo.

Clynes, M. (1995). Microstructural musical linguistics: composers' pulses are liked most by the best musicians. *Cognition, 55*, 269-310.

Clynes, M., & Walker, J. (1982). Neurobiologic functions of rhythm, time, and pulse in music. In M. Clynes (Ed.), *Music, Mind and Brain*. New York: Plenum Press.

Desain, P., & Honing, H. (1991). Towards a calculus for expressive timing in music. *Computers in Music Research, 3*, 43-20.

Desain, P., & Honing, H. (1992). Time functions best as functions of multiple times. *Computer Music Journal, 16*(2), 17-34.

Desain, P., & Honing, H. (1993). Tempo Curves Considered Harmful. *Contemporary Music Review, 7*(2), 123-138.

Desain, P., & Honing, H. (1996). *Physical motion as a metaphor for timing in music: the final ritard.* Paper presented at the Proceedings of ICMC.

Desain, P., & Honing, H. (2002). *Rhythmic stability as explanation of category size.* Paper presented at the 7th International Conference on Music Perception and Cognition, Sydney.

Desain, P., & Honing, H. (2003). The formation of rhythmic categories and metric priming. *Perception, 32*(3), 341-365.

Drake, C. (1993). Reproduction of musical rhythms by children, adult musicians, and adult nonmusicians. *Perception and Psychophysics, 53*, 25-33.

Drake, C. (1998). Psychological Processes Involved in the Temporal Organization of Complex Auditory Sequences: Universal and Acquired Processes. *music Perception, 16*(1), 11-26.

Drake, C., Jones, M. R., & Baruch, C. (2000). The development of rhythmic attending in auditory sequences: attunement, referent period, focal attending. *Cognition, 77*, 251-288.

Dunlap, K. (1910). Reactions on rhythmic stimuli, with attempt to synchronise. *Psychological Review, 17*, 399-416; cited in Scott, 1998.

Elsner, J. (1990). Der Rhythmus insiraf. Zum Problem quantitativer Rhythmik. In O. Elschek (Ed.), *Rhythmik und Metrik in traditionellen Musikkulturen*. Bratislava;

cited in Seifert et al, 1995: Verlag der Slovakischen Akademie der Wissenschaften.

Epstein, D. (1988). Tempo Relations in Music: A Universal? In I. Rentschler, B. Herzberger & D. Epstein (Eds.), *Beauty and the Brain: Biological Aspects of Aesthetics*. Basel: Birkhäuser Verlag.

Fraisse, P. (1967). Le seul différential de durée dans une suite réguliere d'intevalles. *Année Psychologique, 1*, 43-49.

Fraisse, P. (1982). Rhythm and Tempo. In D. Deutsch (Ed.), *The Psychology of Music*. New York: Academic Press, Inc.

Fraisse, P., Pichot, P., & Clairouin, G. (1949). Etude comparée des oligophrènes et des enfants normaux. *Journal de Psychologie Normale et Pathologique, 42*, 309-330; cited in Fraisse 1982.

Freeman, P., & Lacey, L. (2001). The Funky Drummer: Rhythmic Timing Analysis. *Personal communication*.

Freeman, P., & Lacey, L. (2002). *Swing and groove: Contextual rhythmic nuance in live performance.* Paper presented at the 7th International Conference on Music Perfception and Cognition, Sydney.

Friberg, A. (1995). *A Quantitative Rule System for Musical Expression.* Unpublished Doctoral dissertation, Royal Institute of Technology, Sweden.

Friberg, A., & Bresin, R. (1997). *Automatic Musical Punctuation: A Rule System and a Neural Network Approach.* Paper presented at the Proceedings of KANSEI - The Technology of Emotion AIMI International Workshop,, Genoa, Italy.

Friberg, A., Bresin, R., Fryden, L., & Sundberg, J. (1998). Musical Punctuation on the Microlevel: automatic Identification and Performance of Small Melodic Units. *Journal of New Music Research, 27*(3), 271-292.

Friberg, A., & Sundström, A. (2002). Swing Ratios and Ensemble Timing in Jazz Performance: Evidence for a Common Rhythmic Pattern. *Music Perception, 19*(3), 333-349.

Friedberg, J. (2000). *Cultural Change in Traditional Dances in Florina, Greece. A Pilot Study.* Paper presented at the Dance and History, 14th International Congress on Dance Research, Athens.

Frischeisen-Kohler, I. (1933). *Das Personlische Tempo. Eine erbliogische Untersuchung*. Leipzig: Thieme; cited in Fraisse 1982.

Gabrielsson, A. (1982). Perception and performance of musical rhythm. In M. Clynes (Ed.), *Music, mind and brain*. New York: Plenum Press.

Gordon, P. C., & Meyer, D. E. (1987). Hierarchical representation of spoken syllable order. In A. Allport, D. G. Mackay, W. Prinz & E. Scheerer (Eds.), *Language Perception and Production: Relationships between Listening, Speaking, Reading and Writing*. London: Academic Press.

Harrel, T. W. (1937). Factors influencing preference and memory for auditory rhythm. *Journal of General Psychology, 17*, 63-104; cited in Fraisse 1982.

Hary, D., & Moore, G. (1985). Temporal tracking and synchronization strategies. *Human Neurobiology, 4*, 73-76.

Hary, D., & Moore, G. (1987). Synchronizing Human Movement with an External Clock Source. *Biological Cybernetics, 56*, 305-311.

Helmuth, L., & Ivry, R. (1996). When two hands are better than one: Reduced timing variability during bimanual movements. *Journal of Experimental Psychology: Human Perception and Performance, 22*, 278-293.

Hirata, K., & Hiraga, R. (2002). *Ha-Hi-Hun: Performance Rendering System of High Controllability.* Paper presented at the Proceedings of ICAD, Rencon Workshop.

Hirata, K., Hiraga, R., & Aoyagi, T. (2000). *Next Generation Performance Rendering - Exploiting Controllability.* Paper presented at the Proceedings ICMC, San Francisco.

Hodges, R. (1992). Drum Is the Ear of God: Africa's Inner World of Music. *Material for Thought, 13*.

Honing, H. (1992). *Expresso, a strong and small editor for expression.* Paper presented at the Proceedings ICMC.

Honing, H. (2001). From Time to Time: The Representation of Timing and Tempo. *Computer Music Journal, 25*(3).

Ivry, R. B., & Richardson, T. C. (2002). Temporal Control and Coordination: The Multiple Timer Model. *Brain and Cognition, 48*, 117-132.

Ivry, R. B., Richardson, T. C., & Helmuth, L. L. (2002). Improved Temporal Stability in Multieffector Movements. *Journal of Experimental Psychology: Human Perception and Performance, 28*(1), 72-92.

Iyer, V., Bilmes, J., Wright, M., & Wessel, D. (1997). *A novel representation for Rhythmic Structure.* Paper presented at the Proceedings ICMC.

Iyer, V. S. (1998). *Microstructures of Feel, Macrostructures of Sound: Embodied Cognition in West African and African-American Musics.* Unpublished Thesis, University of California, Berkeley.

Jones, M. R., & Boltz, M. (1989). Dynamic attending and responses to time. *Psychological Review, 96*, 459-491.

Keele, S. W. (1968). Movement control in skilled motor performance. *Psychological Bulletin, 70*, 387-403.

Kelso, J. A. S. (1984). Phase transitions and ciritical behavior in human bimanual coordination. *American Journal of Physiology: Regulatory, Integrative and Comparative, 246*, R1000-R1004.

Kelso, J. A. S., Holt, K. G., Rubin, P., & Kugler, P. N. (1981). Patterns of human interlimb coordination emerge from the properties of non-linear, limit-cycle oscillatory processes: Theory and data. *Journal of Motor Behavior, 13*, 226-261.

Klapp, S. T., Anderson, W. G., & Berrian, R. W. (1973). Implicit speech in reading, reconsidered. *Journal of Experimental Psychology: Human Perception and Performance, 8*, 368-374; cited in MacKay, 1987.

Kowal, S., O'Connell, D., & Sabin, E. (1975). Development of temporal patterning and vocal hesitations in spontaneous narratives. *Journal of Psycholinguistic Research, 4*, 195-207.

Krumhansl, C. L. (2000). Rhythm and Pitch in Music Cognition. *Psychological Bulletin, 126*(1), 159-179.

Ladzekpo, C. K. (1995). *Foundation Course in African Dance-Drumming*. Retrieved 31/5/03, 2003, from http://www.cnmat.berkeley.edu/~ladzekpo/Foundation.html

Large, E. W., Fink, P., & Kelso, J. A. S. (2002). Tracking simple and complex sequences. *Psychological Research, 66*, 3-17.

Large, E. W., & Jones, M. R. (1999). The Dynamics of Attending: How People Track Time-Varying Events. *Psychological Review, 106*(1), 119-159.

Large, E. W., & Kolen, J. F. (1994). Resonance and the perception of musical meter. *Connection Science, 6*, 177-208.

Lerdahl, F., & Jackendoff, R. (1981). On the theory of grouping and meter. *Musical Quarterly, 25*, 45-90.

Lerdahl, F., & Jackendoff, R. (1983). *A Generative Theory of Tonal Music*. Cambridge, Mass.: MIT Press.

London, J. (2000). *Hierarchical representations of complex meters*. Retrieved April 20, 2005, from http://www.people.carleton.edu/~jlondon/hierarchical_representatio.htm

MacKay, D. G. (1987). Constraints on theories of sequencing and timing in language perception and production. In A. Allport, D. G. Mackay, W. Prinz & E. Scheerer (Eds.), *Language Perception and Production: Relationships between Listening, Speaking, Reading and Writing*. London: Academic Press.

Madison, G. (2000). Properties of Expressive Variability Patterns in Music. *Journal of New Music Research, 29*(4), 335-356.

Madison, G., & Merker, B. (2002). On the limits of anisochony in pulse attribution. *Psychological Research, 66*, 201-207.

Magill, J. M., & Pressing, J. L. (1997). Asymmetric Cognitive Clock Structures in West African Rhythms. *Music Perception, 15*(2), 189-222.

Mates, J. (1994a). A model of synchronization of motor acts to a stimulus sequence I. Timing and error corrections. *Biological Cybernetics, 70*(5), 463-473.

Mates, J. (1994b). A model of synchronization of motor acts to a stimulus sequence II. stability analysis, error estimation and simulations. *Biological Cybernetics, 70*(5), 474-484.

Mates, J., & Aschersleben, G. (2000). Sensorimotor synchronization: The influence of temporally displaced auditory feedback. *Acta Psychologica, 104*, 29-44.

Mates, J., Müller, U., Radil, T., & Pöppel, E. (1994). Temporal integration in sensorimotor synchronisation. *Journal of Cognitive Neuroscience, 6*, 332-340; cited in Pöppel 1996.

Mates, J., Radil, T., & Pöppel, E. (1992). Cooperative tapping: Time control under different feedback conditions. *Perception & Psychophysics, 52*, 691-704.

Mazzola, G., & Zahorka, O. (1994a). *The RUBATO Performance Workstation on NEXTSTEP.* Paper presented at the Proceedings ICMC, San Francisco.

Mazzola, G., & Zahorka, O. (1994b). Tempo Curves Revisited: Hierarchies of Performance Fields. *Computer Music Journal, 18*(1), 40-52.

Michon, J. (1967). *Timing in temporal tracking*. Assen, NL: Van Gorcum.

Mishima, J. (1951-1952). Fundamental research on the constancy of "mental tempo". *Japanese Journal of Psychology, 22*, 27-28; cited in Fraisse 1982.

Moelants, D. (2002). *Preferred tempo reconsidered.* Paper presented at the 7th international Conference on Music Perception and Cognition, Sydney.

Monahan, C. B. (1993). Parallels Between Pitch and Time and How They Go Together. In T. J. Tighe & W. J. Dowling (Eds.), *Psychology and music: The understanding of melody and rhythm*. Hillsdale, N.J.: Lawrence Erlbaum Associates, Inc.

O'Boyle, D. J., & Clarke, V. L. (1996). On the source of the negative synchronization error during temporal-tracking performance. *Brain Research Association Abstracts, 13*(40).

Paillard, J. (1948). Quelques donnees psychophysiologique relatives au declenchement de la commande motrice. *Annee Psychologique, 46-47*, 28-47.

Palmer, C. (1989). Mapping musical thought to musical performance. *Journal of Experimental Psychology, 15*(12), 331-346.

Palmer, C. (1996). Anatomy of a performance: sources of musical expression. *Music Perception, 13*, 433-454.

Palmer, C. (1997). Music Performance. *Annual Review of Psychology, 48*, 115-138.

Palmer, C., & Pfordresher, P. Q. (2003). Incremental planning in sequence production. *Psychological Review, 110*(4).

Palmer, C., & van de Sande, C. (1995). Range of planning in skilled music performance. *Journal of Experimental Psychology: Human Perception and Performance, 21*, 947-962.

Parncutt, R. (1994). A Perceptual Model of Pulse Salience and Metrical Accent in Musical Rhythms. *Music Perception, 11*(4), 409-464.

Pöppel, E. (1983). Musik-erleben und Zeit-struktur. In E. H. Gombrich & H. Petsche (Eds.), *Auge macht Bild - Ohr macht Klang - hirn macht Welt*. Wien, cited in Seifert et al, 1995: Franz Deuticke.

Pöppel, E. (1988). *Mindworks, Time and Conscious Experience*. Boston: Harcourt Brace Jovanich; cited in Pöppel 1996.

Pöppel, E. (1996). Reconstruction of subjective time on the basis of hierarchically organized processing system. In M. A. Pastor & J. Artieda (Eds.), *Time, Internal Clocks and Movement*. Amsterdam: Elsevier.

Povel, D.-J. (1981). Internal representation of simple temporal patterns. *Journal of Experimental Psychology: Human Perception and Performance, 7*, 3-18.

Povel, D.-J. (1984). A Theoretical Framework for Rhythm Perception. *Psychological Research, 45*, 315-337.

Povel, D.-J., & Collard, R. (1982). Structural factors in patterned finger tapping. *Acta Psychologica, 52*, 107-123; cited in Keele, 1987.

Povel, D.-J., & Essens, P. (1985). Perception of temporal patterns. *Music Perception, 2*, 411-440; cited in Summers & Burns, 1990.

Pressing, J. (1983). Cognitive isomorphisms in pitch and rhythm in world musics: West Africa, the Balkans, thailand, and Western tonality. *Studies in Music, 17*, 38-61.

Pressing, J. (1998a). Error Correction Processes in Temporal Pattern Production. *Journal of Mathematical Psychology, 42*, 63-101.

Pressing, J. (1998b). Referential Behaviour Theory: A framework for multiple perspectives on motor control. In J. P. Piek (Ed.), *Motor Control and Human Skill: A Multidisciplinary Approach* (pp. 357-384). Champaign, Illinois: Human Kinetics.

Pressing, J. (1999). The Referential Dynamics of Cognition and Action. *Psychological Review, 106*(4), 714-747.

Pressing, J. (2002). Black Atlantic rhythm: Its Computational and Transcultural Foundations. *Music Perception, 19*(3), 285-330.

Pressing, J., & Jolley-Rogers, G. (1997). Spectral properties of human cognition and skill. *Biological cybernetics, 76*, 339-347.

Pressing, J., Summers, J., & Magill, J. (1996). Cognitive Multiplicity in Polyrhythmic Pattern Performance. *Journal of Experimental Psychology: Human Perception and Performance, 27*(5), 1127-1148.

Puckette, M. (1991). Combining Event and Signal Processing in the MAX Graphical Programming Environment. *Computer Music Journal, 15*(3), 58-67.

Rasch, R. A. (1988). Timing and synchronization in ensemble performance. In J. A. Sloboda (Ed.), *Generative Processes in Music: The Psychology of Performance, Improvisation, and composition*. Oxford: Clarendon Press.

Repp, B. H. (1990). Patterns of expressive timing in performances of a Beethoven minuet by 19 famous pianists. *Journal of the Acoustical Society of America, 88*, 622-641.

Repp, B. H. (1992). Diversity and commonality in music performance - an analysis of timing microstructures in Schumann's Traumerei. *Journal of the Acoustical Society of America, 92*(5), 2546-2568.

Repp, B. H. (1998). A microcosm of musical expression: I. Quantitative analysis of pianists' timing in the initial measures of Chopin's Etude in E major. *Journal of the Accoustical Society of America, 104*, 1085-1100.

Repp, B. H. (2000). Compensation for subliminal timing perturbations in perceptual-motor synchronization. *Psychological Research, 63*(2), 106-128.

Repp, B. H. (2001a). Phase Correction, Phase Resetting and Phase Shifts After Subliminal Timing Perturbations in Sensorimotor Synchronisation. *Journal of Experimental Psychology: Human Perception and Performance, 27*(3), 600-621.

Repp, B. H. (2001b). Processes underlying adaptation to tempo changes in sensorimotor synchronisation. *HUMAN MOVEMENT SCIENCE, 20*, 277-312.

Repp, B. H. (2002a). Automaticity and Voluntary Control of Phase Correction Following event Onset Shifts in Sensorimotor Synchronisation. *Journal of Experimental Psychology: Human Perception and Performance, 28*(2), 410-430.

Repp, B. H. (2002b). *Effects of metrical structure on phase resetting in sensorimotor synchronization.* Paper presented at the 7th International conference on Music Perception and Cognition, Sydney.

Repp, B. H. (2002c). Phase correction in sensorimotor synchronization: Nonlinearities in voluntary and involuntary responses to perturbations. *Human Movement Science, 21*(1), 1-37.

Repp, B. H., Windsor, W. L., & Desain, P. (2002). Effects of Tempo on the Timing of Simple Musical Rhythms. *Music Perception, 19*(4), 565-593.

Richman, B. (2000). How Music fixed "Nonsense" into Significant Formulas: On Rhythm, Repetition, and Meaning. In N. L. Wallin, B. Merker & S. Brown (Eds.), *The Origins of Music*. Cambridge, Mass: MIT Press.

Rimoldi, H. J. A. (1951). Personal tempo. *Journal of Abnormal and Social Psychology, 46*, 283-303; cited in Fraisse 1982.

Rosenbaum, D. A. (1985). Motor programming: A review and scheduling theory. In H. H, U. Kleinbeck & K.-M. Schmidt (Eds.), *Motor behavior: Programming, control, and acquisition*. Berlin; cited in Rosenbaum, 1991: Springer-Verlag.

Rosenbaum, D. A., Inhoff, A. W., & Gordon, A. M. (1984). Choosing between movement sequences: A hierarchical editor model. *Journal of Experimental Pscyhology: General, 113*, 372-393; cited in Summers & Burns, 1990.

Rosenbaum, D. A., Kenny, S., & Derr, M. A. (1983a). Hierarchical control of rapid movement sequences. *Journal of Experimental Psychology: Human Perception and Performance, 9*, 86-102; cited in Rosenbaum, 1991.

Rosenbaum, D. A., Kenny, S. B., & Derr, M. A. (1983b). Hierarchical control of rapid movement sequences. *Journal of Experimental Psychology: Human Perception and Performance, 9*, 86-102; cited in Keele. 1987.

Saue, S. (2000). Implementing rhythmic frequency modulation. In C. H. Waadeland (Ed.), *Rhythmic Movements and Moveable Rhythms, Appendix II* (pp. 252-276; cited in Waadeland,, 2001). Trondheim: Department of Musicology, Norwegian University of Science and Technology.

Schleidt, M., Eibl-Eibesfeldt, I., & Pöppel, E. (1987). A universal constant in temporal segmentation of human short-term behaviour. *Naturwissenschaften, 74*, 289-290; cited in Pöppel 1996.

Schmidt, R. A. (1975). A schema theory of discrete motor skill learning. *Psychological review, 82*, 225-260.

Seifert, U., Olk, F., & Schneider, A. (1995). On Rhythm Perception: Theoretical Issues, Empirical Findings. *Journal of New Music Research, 24*, 164-195.

Semjen, A., & Ivry, R. B. (2001). The Coupled Oscillator Model of Between-Hand Coordination in Alternate-Hand Tapping: A Reappraisal. *Journal of Experimental Psychology: Human Perception and Performance, 27*(2), 251-265.

Semjen, A., Schulze, H.-H., & Vorberg, D. (2000). Timing precision in continuation and synchronisation tapping. *Psychological Research, 63*, 137-147.

Semjen, A., Vorberg, D., & Schulze, H.-H. (1998). Getting synchronized with the metronome: Comparisons between phase and period correction. *Psychological Research, 61*, 44-55.

Shaffer, L. H. (1980). Analysing piano performance: a study of concert pianists. In G. E. Stelmach & J. Requin (Eds.), *Tutorial in motor behavior* (pp. 443-455): North-Holland.

Shaffer, L. H. (1981). Performances of Chopin, Bach and Bartok: studies in motor programming. *Cognitive Psychology, 13*, 327-376.

Shaffer, L. H. (1985). Timing in Action. In J. A. Michon & J. L. Jackson (Eds.), *Time, Mind and Behavior*. Berlin: Springer-Verlag.

Shaffer, L. H., Clarke, E. F., & Todd, N. P. (1985). Metre and rhythm in piano playing. *Cognition, 20*, 61-77.

Shove, P., & Repp, B. H. (1995). Musical motion and performance: theoretical and empirical perspectives. In J. Rink (Ed.), *The Practice of Performance: Studies* *in Musical Interpretation* (pp. 55-83Cambridge University Press; cited in Waadeland, 2001). Cambridge: Cambridge University Press.

Sievers, E. (1924). *Ziele un Wege der Schallanalyse*. Heidelberg: Carl Winter's Universsitäts-buchhandlung. Cited in Waadeland, 2001.

Snyder, B. (2000). *Music and Memory*. Cambridge, Massachusetts: The MIT Press.

Stange-Elbe, J., & Mazzola, G. (1998). *"Cooking a Canon with RUBATO" Performance Aspects of J.S. Bach's "Kunst der Fuge".* Paper presented at the Proceedings ICMC, San Francisco.

Stern, W. (1900). Das psychisch Tempo. In *Uber psychologie der individuellen differenzen*. Leipzig; cited in Fraisse 1982: Barth.

Sternberg, S., & Knoll, R. L. (1994). Perception, Production and Imitation of Time Ratios by Skilled Musicians. In R. Aiello & J. A. Sloboda (Eds.), *MUSICAL PERCEPTIONS*. New York, Oxford: Oxford University Press.

Sternberg, S., Monsell, S., Knoll, R. L., & Wright, C. E. (1978). The latency and duration of rapid movement sequences: Comparisons of speech and typewriting. In G. E. Stelmach (Ed.), *Information Processing in Motor Control and Learning*. New York; cited in MacKay, 1987: Academic Press.

Summers, J. J., & Burns, B. D. (1990). Timing in Human Movement Sequences. In R. A. Block (Ed.), *Cognitive models of psychological time*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Summers, J. J., Rosebaum, D. A., Burns, B. D., & Ford, S. K. (1993). Production of Polyrhythms. *Journal of Experimental Psychology: Human Perception and Performance, 19*(2), 416-438.

Sundberg, J., Friberg, A., & Fryden, L. (1991). Common secrets of Musicians and Listeners: An analysis-by-synthesis study of Musical Performance. In P. Howell, R. West & I. Cross (Eds.), *Representing Musical Structure* (pp. 161-197). London: Academic Press.

Thaut, M., & Miller, R. (1994). Multiple synchronization strategies in tracking of rhythmic auditory stimulation. *Proceedings of the Society for Neuroscience, 146*(11; cited in Thaut et al, 1998).

Thaut, M., Tian, B., & Azimi-Sadjadi, M. (1998). Rhythmic finger tapping to cosine-wave modulated metronome sequences: Evidence of subliminal entrainment. *HUMAN MOVEMENT SCIENCE, 17*, 839-863.

Thaut, M. H., & Kenyon, G. P. (2003). Rapid motor adaptations to subliminal frequency shifts during syncopated rhythmic sensorimotor synchronization. *Human Movement Science, In press*.

Thaut, M. H., Miller, R. A., & Schauer, L. M. (1998). Multiple synchronization strategies in rhythmic sensorimotor tasks: phase vs period correction. *Biological Cybernetics, 79*, 241-250.

Timmers, R., Ashley, R., Desain, P., & Heijink, H. (2000). The Influence of Musical Context on Tempo Rubato. *Journal of New Music Research, 29*(2), 131-158.

Todd, N. (1985). A model of expressive timing in tonal music. *Music Perception, 3*, 33-51.

Todd, N. (1989). A Computational Model of Rubato. *Contemporary Music Review, 3*(1).

Todd, N. P. M. (1992a). *The communication of self-motion in musical expression.* Paper presented at the International Workshop on Man-Machine studies in Live Performance, Pisa.

Todd, N. P. M. (1992b). The dynamics of dynamics: A model of musical expression. *Journal of the Acoustical Society of America, 9*(6), 3540-3550; cited in Waadeland, 2001.

Todd, N. P. M. (1992c). *Music and motion: A personal view.* Paper presented at the Fourth Rhythm Workshop: Rhythm Perception and Production, Bourges.

Todd, N. P. M. (1999). Motion in Music: A Neurobiological Perspective. *Music Perception, 17*(1), 115-126.

Todd, N. P. M., O'Boyle, D., & Lee, C. S. (1999). A Sensory-Motor Theory of Rhythm, Time Perception and Beat Induction. *Journal of New Music Research, 28*(1), 5-28.

Trilsbeek, P., Desain, P., & Honing, H. (2001). *Spectral analysis of timing profiles of piano performances.* Paper presented at the Proceedings of the International Computer Music Conference, San Francisco.

Truslit, A. (1938). *Gestaltung und Bewegung in der Musik*. Berlin-Lichterfelde: Chr. Friedrich Vieweg; cited in Waadeland, 2001.

Turner, F., & Pöppel, E. (1988). Metered Poetry, the Brain, and Time. In I. Rentschler, B. Herzberger & D. Epstein (Eds.), *Beauty and the Brain: Biological Aspects of Aesthetics*. Basel: Birkhäuser Verlag.

Various. (2000). Women of Rembetika, CD [CD]. Cambridge, Mass.: Rounder Records Corp.

Vollrath, M., Kazenwadel, J., & Krüger, H.-P. (1992). A universal constant in temporal segmentation of human speech. *Naturwissenschaften, 79*, 479-480.

Vorberg, D., & Hambuch, R. (1978). On the Temporal Control of Rhythmic Performance. In J. Requin (Ed.), *Attention and PErformance VII*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Vorberg, D., & Wing, A. M. (1996). Modelling variability and dependence in timing. In H. Heuer & S. W. Keele (Eds.), *Handbook of perception and action: Vol 3. Motor skills* (pp. 181-262). London: Academic Press.

Vos, P., Mates, J., & van Kruysbergen, N. (1995). The perceptual center of a stimulus as the cue for synchronisation to a metronome: Evidence from asynchronies. *Quarterly Journal of Experimental Psychology, 48A*, 1024-1040; cited in Pöppel 1996.

Vos, P. G. (1976). *Identification of metre in music* (No. 76 ON 06): University of Nijmegen; cited in Seifert et al, 1995.

Waadeland, C. H. (2001). "It Don't Mean a Thing If It Ain't Got That Swing" - Simulating Expressive Timing by Modulated Movements. *Journal of New Music Research, 30*(1).

Waadeland, C. H., & Saue, S. (1999). *Computer Implementation of Rhythmic Frequency Modulation in Music.* Paper presented at the Proceedings 99 Digital Audio Effects Workshop, Trondheim.

Waterman, C. A. (1998). The Alberts Collection. In *The Arthur S. Alberts Collection: More Tribal, Folk, and Café Music of West Africa, CD liner notes* (Vol. RCD 10401): 360o Productions.

Welford, A. T. (1974). On the sequencing of action. *Brain Research, 71*, 383-392.

Wessel, D., Wright, M., & Khan, S. A. (1998). *Preparation for Improvised Performance in Collaboration with a Khyal Singer.* Paper presented at the Proceedings ICMC.

Widmer, G. (1994). *Learning Expression at Multiple Structural Levels.* Paper presented at the Proceedings ICMC, San Francisco.

Widmer, G. (1996). Learning Expressive Performance: The Structure-Level Approach. *Journal of New Music Research, 25*, 179-205.

Widmer, G. (2000). *Large-scale Induction of Expressive Performance Rules: First Quantitative results.* Paper presented at the Proceedings ICMC, San Francisco.

Widmer, G. (2001). *Inductive Learning of general Robust and Local Expression Principles.* Paper presented at the Proceedings ICMC, San Francisco.

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

Wing, A. M., & Kristofferson, A. B. (1973b). The timing of interresponse intervals. *Perception & Psychophysics, 13*.

Wright, M. (2002). *Problems and prospects for intimate and satisfying sensor-based control of computer sound.* Paper presented at the Sensors and Input for Multimedia Systems (SIMS), Santa Barbara, CA.

Yamanishi, J., Kawato, M., & Suzuki, R. (1980). Two coupled oscillators as a model for the coordinated finger tapping by both hands. *Biological Cybernetics, 37*, 219-225.

# Appendices

# Appendix A Models of expressive timing in music

This appendix summarises aspects of the current research literature on models of expressive timing in music, which were not appropriate to the literature review on models of microtiming deviation. Most of the musical expression literature focuses on methods for the analysis of Western notated art music and the modelling of its performance (with the notable exceptions of (Bilmes, 1992), (Bilmes, 1993a), (Bilmes, 1993b) (Iyer et al., 1997) and (Wessel, Wright, & Khan) and includes a variety of performance parameters besides timing deviation; for example, loudness and articulation.

Since my own research deals only with timing deviation in traditional African drumming, literature dealing with the following issues which are not directly relevant to the present topic is summarised briefly in this appendix:

- Parameters of analysis and performance other than timing deviation; interaction of different parameters .
- Analysis using hyper-meter and grouping rules, similar to those originally proposed by Lerdahl and Jackendoff in (Lerdahl & Jackendoff, 1983). As discussed in the section in this thesis on musicology, African drumming does not conform to a metrical hierarchy beyond the basic time cycle.
- Separation of the performance into a notated (or notatable) score, and performed deviations from the score.
- Adding expressive deviations to sound files through Fourier analysis and resynthesis.

## A.1 The relation of musical structure to deviations

### A.1.1 Significance of musical structure for expressive deviations

The interdependence of temporal structure (rhythm, meter, or phrase structure), global tempo, and timing deviations is pointed out by Clynes (Clynes, 1983) and by Honing (Honing, 2001) Timing deviations help in communicating temporal structure to the listener, and at different tempi other structural levels of the music are emphasised and the expressive timing adapted accordingly.

Various levels of temporal structure, and the compositional means of articulating these, have been found to affect timing in performance. Possibly the smallest unit which can be deemed 'structural' are the small melodic units identified by Friberg et al in (Friberg et al., 1998). At the next highest level, deviations which relate to metre are identified and described Clynes in (Clynes, 1995). Harmonic patterns are found to be recognised through temporal deviation in performance by Palmer (Palmer, 1996) and by Sundberg et al (Sundberg et al., 1991). The common practice of slowing towards the end of a musical phrase or group is identified by Todd (Todd, 1985), Palmer (Palmer, 1989), and Repp (Repp, 1990,1992). The hierarchic depth of the group or phrase in question is considered in relation to the amount of slowing at the end of the phrase by Shaffer (Shaffer, 1980) and Todd (Todd, 1985). An overlay of variation on deviations consistent with these kinds of analysis was found by Timmers et al (Timmers et al., 2000), when the same melodic theme was performed in different musical contexts (eg. with different accompaniment) by professional pianists.

Evidence for phrase structure as a determinant of performance deviations has been found by Trilsbeek et al (Trilsbeek, Desain, & Honing, 2001) through spectral analysis of timing deviations of different pianists playing transcriptions of Beatles songs, at a variety of tempi.

Timmers et al (Timmers et al., 2000) analysed MIDI data captured from performances by pianists of a melodic theme (from Brahms *Variations on an Original Theme)* in different musical settings, including the melody alone, with block chords, with counter-melody, and the full setting. Different settings were found to produce different timing deviations in each pianist, and there was more correlation of timing deviation between performers for some settings, and less for others. The authors warn against generalising from this study without further testing, however it seems clear at least that different musical setting can affect timing deviation choices in performance.

## A.2 Application of structural information to models of expressive deviation in music

Models which make use of structural information to generate expressive deviations include work undertaken by Desain and Honing (Desain & Honing, 1993), where Todd's model of rubato linked to phrase structure (Todd, 1989) is discussed and applied to a theme and variations for piano by Beethoven. This work is extended by

Honing in (Honing, 1992), where expression is defined in terms of a performance (rather than a score) and its structural description, so that deviations of local tempo (for example at the beat level) are relative to the tempo of the next highest structure in the performance (in this case, the bar).

Widmer (Widmer, 1994), (Widmer, 1996) applies a multiple-level structural analysis to music in a model designed to learn principles of musical dynamics (variations in loudness) and timing variation. Mazzola and Zahorka (Mazzola & Zahorka, 1994b) explore the application of hierarchical tempo variation (and its interaction with articulation) to piano pieces by Czerny and Chopin. More extensive work, applying hierarchical structure information to variations of onset timing, dynamics, articulation, and pitch (variations on score pitch, such as intonation and glissandi) is found in Mazzola and Zahorka (Mazzola & Zahorka, 1994b; Stange-Elbe & Mazzola, 1998).

Hierarchical time-span reduction of musical structure, based on principles drawn from Lerdahl and Jackendoff (Lerdahl & Jackendoff, 1983), is the input to the model described by Hirata et al in (Hirata & Hiraga, 2002; Hirata, Hiraga, & Aoyagi, 2000), which utilises the information to identify structurally important notes, and to apply agogic and dynamic stresses appropriately. Similarly, Bresin (Bresin, 1998,2000) proposes a model which takes information about the global structure of a piece of music as input, to produce a performance with the desired affect.

Analysis of the kinds of cyclic (or ‘synchronic’) music to which my model relates does not yield groupings of meter into hypermetric structures (please refer to note 2, above).

## A.3 Models of analysis and generation of expressive deviation

Analysis for the purposes of generative models expressive deviation must achieve one or both of two interlinked tasks. One task is to find expressive deviations where they occur – for timing deviations this is always in relation to a score, either explicitly available (as a written score on paper, or a quantised MIDI file) or implicitly recoverable from the performance (where rhythmic patterns can be transcribed into western notation by a competent musician). Other kinds of variation, such as loudness or articulation, depend more on local context. Since my model deals only with timing information, I will generally ignore other kinds of deviations.

A second task for analysis is to make sense of the deviations found, in terms of some identifiable property of the piece and its performance, or both. Thus one or more of: articulation of structural boundaries, or categorisation of musical entities as either ornaments or melodic notes, or local rules relating to the shape and rhythm of the melody, or rules which generate expressive affect, may apply depending on the model and the musical context.

Analysis of musical works can be approached through either human analysis or machine analysis. Human analysis can be achieved by analysis of the written score; analysis by hand of recordings of performances; or analysis-by-synthesis, where parameter values of a generative model of expressive deviations are manipulated until the performance sounds right to a competent listener.

Machine analysis of onset timing deviations is found in Trilsbeek et al (Trilsbeek et al., 2001) who use spectral analysis of deviation data to extract structural boundaries as marked by deviations in performance. MIDI data for a number of different performances (3 performances each by 12 different pianists) were obtained of specific scores at a particular tempo. Eighth notes were interpolated in the melody (both linear and spline methods were tried) where required to give a consistent (although temporally deviated) eighth note pulse. Inter-onset intervals for the onsets (including interpolated notes) were found, and the means of each inter-onset interval in the score over the 36 performances were found. These data were centred around zero and used as input to a Fourier transform. Peaks in the frequency domain output by the Fourier transform correspond to a timing fluctuation found in common across all performances. In fact, peaks were found at boundaries between phrases, as would be expected from a human analysis of the score. Interestingly, the peaks found in the analysis of performances by a group of jazz musicians were smaller than those in performances by a group of classical musicians: jazz players kept more to the beat. Clearly, methods such as this rely on a given, or recoverable score. Inter-onset deviations only have meaning in relation to an undeviated score, which is not available in my case (please refer to note 3, above).

An analysis-by-synthesis approach is described in Desain and Honing (Desain & Honing, 1993), where tempo changes are applied to a performance, and details of the tempo curve adjusted at various levels of the temporal hierarchy, in an incremental approach to extracting workable tempo transformation rules for musical entities commonly found in music from the European Classical period, such as grace notes and phrase final lengthenings.

Canazza et al (Canazza et al., 1997; Canazza, Roda et al., 1999) apply a method of analysis-by-synthesis in which computer-generated performances use parameter values which are interpolated between parameter values previously found to produce performances which correlated well with sensorial evaluation adjectives judged appropriate by listening subjects. The interpolated values were found to have intermediate expressive characteristics. This model is extended by Canazza and others (Canazza et al., 2000; Canazza, Roda et al., 1999) to analyse and resynthesise digital audio files, with loudness, tempo, note length, attack time and attack spectral information all independently subject to appropriate rules in resynthesis. An important difference of this model from those such as Trilsbeek et al (Trilsbeek et al., 2001) is a concern with expression of affect or emotion in music performance, in addition to a rendering which takes into account structural information in order to produce a natural, musical-sounding performance. Canazza et al use the terms *nominal performance* for an exact, mechanical reproduction, and *neutral performance* for a 'literal human performance of the score without any expressive intention or stylistic choice' (Cannazza et al., 1998). Hierarchical structural information is therefore included as an input to the model, as indeed are the *nominal performance* data (corresponding to score information) and the *neutral performance* data (Cannazza et al., 1998).

An attempt to extract commonality between different strategies of variation which are effective in expressing emotion has been made by Madison (Madison, 2000). Professional keyboard musicians recorded performances using a MIDI keyboard of various unaccompanied melodies played naturally, and so as to communicate anger, fear, happiness and sadness. Inter-onset interval data was converted to a proportional duration, relative to a deadpan performance at the same average tempo, and duration (onset to offset interval) was converted to articulation, defined as the duration relative to inter-onset interval. Key velocity was used as a loudness value. Principal component analysis was applied to the resultant data to separate typical variability patterns (more likely to be structure-related) from expressive variability patterns (those intended to express emotion). Variation found in the natural performances was subtracted, and variation per melody and specified emotion was averaged across performers. Timing patterns were then analysed in terms of their "roughness" – a more persistent pattern is smoother; and durational contrast – more durational contrast means that nominally shorter notes are played shorter and nominally longer notes are played longer than notated, while less durational contrast means the opposite. Roughness and durational contrast were also analysed for articulation – in this case, sharper contrast meant that smaller proportions of short notes sounded, and softened contrast meant that larger proportions of short notes were sounded. Roughness and loudness contrast were analysed for loudness. The study found that all these factors play a role in emotional expression, separated from the typical variability related to melodic structure (Madison, 2000).

An example of score-based, rather than performance-based analysis is found in the work of Bresin and others (Battel, Bresin, De Poli, & Vidolin, 1994; Bresin et al., 1992; Friberg & Bresin, 1997; Friberg et al., 1998). Both a rule-based and a neural network-based approach are described.  The development of the rules is credited to work previously undertaken by Sundberg (Sundberg et al., 1991) and Friberg (Friberg, 1995) using analysis-by-synthesis, and extended by other researchers (Bresin et al., 1992). Several different systems are described.  The input to the developed rule system is simply the coded music score, containing in some cases pitch and duration information, and a harmonic analysis made by hand (Friberg & Bresin, 1997; Friberg et al., 1998). The rules work on a context of at most five melody notes, adding weights to each note according to subrules, to ultimately identify points where timing commas are to be inserted. Relevant musical features are appoggiaturas, melodic leaps, relative duration of notes in the context, and tempo. Additional subrules handle interactions between potential comma locations.  The Artificial Neural Network was given similar input, and trained to produce time and loudness deviations in line with those produced by the rules.  Evaluations of the various systems found that both symbolic rule and neural network performances (alone) were given better subjective ratings than deadpan performances (Battel et al., 1994); that neural networks were able to generalise a specific musical style quite well from a very small training set, and  to apply deviations in real time, for example to a melody generated in real time (Bresin et al., 1992); and that neural nets (unsurprisingly) make different decisions than human performers, due to information such as metric structure or song lyric which is available to the human but not to the net (Friberg & Bresin, 1997; Friberg et al., 1998). More recently, Bresin and Friberg have explored the application of deviation rules previously applied to the communication of musical structure to produce emotionally coloured performances (Bresin, 2000).

The “RUBATO” performance workstation developed by Mazzola and others (Mazzola & Zahorka, 1994a; Stange-Elbe & Mazzola, 1998) is modular software designed to take in a score as MIDI data, and produce an appropriately inflected performance, MIDI format. The score is analysed in terms of local meter and hierarchical metrical levels; tonal functions and harmonic tension; and “topological concepts of melodic shape and of similarity between these objects” – presumably related to motivic unity and development. The output of the analysis is a weighting on defined events in the score. A performance score is calculated as a function of the input score data and the weightings produced by the analysis module. Vector fields and calculus are then applied to achieve a nested hierarchy of variations (separately and interactively) in each of onset timing, articulation, loudness and intonation (this model is unusual in being applicable to violin performance). The application of analysis weights to specific performance parameters can be specified by the user in complex ways, including scaling of weight contributions and interaction between different performance parameters (Mazzola & Zahorka, 1994a). Processing of a two-voice canon from JS Bach’s “Kunst der Fuge” excluded ornaments (which were added back to the performance score after processing) and harmonic analysis (Stange-Elbe & Mazzola, 1998). The researchers found, with their model at least, that setting parameter values in a way designed to sonify the structural analysis, produced more successful performances than trying to emulate an expressive style (Stange-Elbe & Mazzola, 1998).

Widmer (Widmer, 1994,1996) developed a model which generates analyses of given MIDI scores, using multiple level structuring concepts as found in Lerdahl and Jackendoff (Lerdahl & Jackendoff, 1983), and produces performance rules from input performances of the score in terms of the analysed structure. In an initial version rules were produced at the note level in terms of each note’s analysed role in the structure. An extension of the model described in (Widmer, 1994,1996) generates expression curves relating to timing and dynamics over temporal spans associated with structural groupings. Central to this projection of multiple-level structural analysis onto performance is machine identification of expression curves in the input performance example. A kind of curve-fitting strategy is applied to the dynamics and tempo of the performance of structural units identified in the score analysis. The structural units are of varying scope, and can be in hierarchical and overlapping structure with each other. Each association of an expressive curve with a type of structure is passed to the learning phase of the model. A symbolic generalization strategy associates specific expressive curves with types of structural unit, and a numeric learning algorithm builds interpolation tables for weight offset and scaling of the curves (Widmer, 1994,1996). The application of the learned rules to new performances is hierarchical by temporal scope, and in this bears some resemblance to the model designed by Mazzola and Zahorka (Mazzola & Zahorka, 1994a,b). Later research by Widmer on a much larger performance input data set (Widmer, 2000,2001) produces rules at the note level once more.  Here, the model attempts to find partial rules for performance of different situation categories which can be defined for individual notes in a score.  The rules (relating to loudness, articulation, and inter-onset timing) should be simple, roust, and musically meaningful, but will not be applicable to all instances of the category.  The rules extracted were found to generalise reasonably well to match performances of different styles and by different performers (Widmer, 2000,2001).

## A.4 Relevance to this thesis

My research will be confined to drum rhythms, and the only deviation I model will be in inter-onset timing.  Loudness is an obvious addition of interest, but will be avoided for the sake of simplicity, at this stage of research.  Therefore, research discussed above relating to deviations in loudness and articulation, and any interaction between these and inter-onset timing deviations, will be ignored.

Similarly, expression of affect is not a goal of my research.  Whether the performances I analyse and the generative model I will produce are expressive of particular, or any emotions will not affect my approach or the outcome.

In the absence of a recoverable score of the performance (please refer to note 3, above), extraction or application of rules relating to smaller structural units at the level of about the ‘bar’ or less, are not possible, However, since meter of some kind is usually perceptible in various styles of African drumming, and since expert practitioners affirm the presence of meter (see the section on musicology in this thesis for a discussion), it seems reasonable to expect that there may be systematic timing deviations in relation to metric structure, and possible to some sub-structures of the meter.  The analysis-by-synthesis model I propose will take this idea as a working hypothesis.

The generation of hierarchical timing curves in performance found in (Mazzola & Zahorka, 1994b) is therefore directly relevant to my work, and will be examined in more detail in the section of this thesis concerned with models of timing deviation which are directly applicable to my model.  Other research discussed in this appendix and relating to hierarchical timing deals with analysis and extraction of rules used to determine where and how to apply timing deviations. These are not considered relevant, both because they deal mainly with Western art music, or at least music which shares some of its diachronic characteristics (Trilsbeek et al., 2001), and because analysis in each case depends on a given, or implicitly available score (see note 3 above). My interest is with details of the implementation of hierarchic timing deviation, rather than this approach to analysis.

Analysis-by-synthesis methods found in (Desain & Honing, 1993) are not relevant because they deal with score-based music, and in the case of [32, 33] are concerned with the expression of emotion.

# Appendix B Raw data

## B.1 The Funky Drummer

### B.1.1 The Funky Drummer break – onset time data provided courtesy of Peter Freeman (Freeman & Lacey, 2001)

| (“1” is a downbeat) | | Onset time (sec) |
|---|---|---|
| 1 | 22.3 B1 Hi-Hat 00 | 22.334 |
| 0 | 22.4 Hi-Hat 5 | 22.486 |
| 0 | 22.6 Hi-Hat 5 | 22.637 |
| 0 | 22.7 Hi-Hat 5 | 22.798 |
| 0 | 22.9 Snare 5 | 22.946 |
| 0 | 23.1 Hi-Hat 5 | 23.103 |
| 0 | 23.2 Hi-Hat 5 | 23.243 |
| 0 | 23.3 Hi-Hat 5 | 23.391 |
| 0 | 23.5 B3 Hi-Hat 5 | 23.506 |
| 0 | 23.6 snare hi-hat 5 | 23.671 |
| 0 | 23.8 Hi-Hat 00 | 23.826 |
| 0 | 23.9 Hi-Hat snare 5 | 23.971 |
| 0 | 24.1 B4 Snare 5 | 24.105 |
| 0 | 24.2 Hi-Hat 00 | 24.269 |
| 0 | 24.4 Hi-Hat 5 | 24.406 |
| 0 | 24.5 Hi-Hat snare 5 | 24.569 |
| 1 | 24.7 B1 Hi-Hat 00 | 24.711 |
| 0 | 24.8 Hi-Hat 00 | 24.867 |
| 0 | 25.0 Hi-Hat 00 | 25.013 |
| 0 | 25.1 Hi-Hat 00 | 25.168 |
| 0 | 25.3 Snare 00 | 25.317 |
| 0 | 25.4 Hi-Hat 00 | 25.472 |
| 0 | 25.6 Hi-Hat 00 | 25.613 |
| 0 | 25.7 Snare 00 | 25.766 |
| 0 | 25.9 Hi-Hat 00 | 25.917 |
| 0 | 26.0 Snare 00 | 26.052 |
| 0 | 26.1 Hi-Hat & James Brown 00 | 26.177 |
| 0 | 26.3 Don't Know 00 | 26.353 |
| 0 | 26.4 Snare 00 | 26.496 |
| 0 | 26.6 Kick 00 | 26.654 |
| 0 | 26.7 High-Hat? 00 | 26.793 |
| 0 | 26.9 Snare 00 | 26.951 |
| 1 | 27.0 B1 Hi-Hat 00 | 27.09 |
| 0 | 27.2 Hi-Hat 00 | 27.243 |
| 0 | 27.3 Kick 00 | 27.38 |
| 0 | 27.5 Hi-Hat 00 | 27.545 |
| 0 | 27.6 Snare beat2 00 | 27.691 |
| 0 | 27.8 Hi-Hat 00 | 27.853 |
| 0 | 27.9 Hi-Hat 00 | 27.989 |
| 0 | 28.1 Hi-Hat 00 | 28.15 |
| 0 | 28.2 Hi-Hat 00 | 28.289 |

| | | |
|---|---|---|
| 0 | 28.4 Snare 00 | 28.425 |
| 0 | 28.5 Hi-Hat 00 | 28.58 |
| 0 | 28.7 Hi-Hat, Snare 00 | 28.732 |
| 0 | 28.8 Snare 00 | 28.872 |
| 0 | 29.0 Kick 00 | 29.03 |
| 0 | 29.1 Hi-Hat 00 | 29.169 |
| 0 | 29.3 Snare 00 | 29.327 |
| 1 | 29.4 B1 Hi-Hat and JB 00 | 29.471 |
| 0 | 29.6 Hi-Hat 00 | 29.626 |
| 0 | 29.7 Hi-Hat 00 | 29.77 |
| 0 | 29.9 Hi-Hat 00 | 29.927 |
| 0 | 30.0 Snare 00 | 30.074 |
| 0 | 30.2 Hi-Hat 00 | 30.23 |
| 0 | 30.3 JB 00 | 30.359 |
| 0 | 30.5 Snare roll? 00 | 30.522 |
| 0 | 30.6 Kick 00 | 30.672 |
| 0 | 30.8 Snare 00 | 30.81 |
| 0 | 30.9 Hi-Hat 00 | 30.954 |
| 0 | 31.1 Hi-Hat 00 | 31.102 |
| 0 | 31.2 Snare 00 | 31.245 |
| 0 | 31.3 Hi-Hat 00 | 31.398 |
| 0 | 31.5 JB "ain't it" 00 | 31.529 |
| 0 | 31.696 Snare | 31.696 |
| 1 | 31.8 B1 Hi-Hat 00 | 31.843 |
| 0 | 31.9 Hi-Hat 00 | 31.994 |
| 0 | 32.1 High-Hat 00 | 32.143 |
| 0 | 32.2 Hi-Hat 00 | 32.299 |
| 0 | 32.446 Snare | 32.446 |
| 0 | 32.6 Hi-Hat 00 | 32.609 |
| 0 | 32.7 Hi-Hat 00 | 32.748 |
| 0 | 32.8 Snare 00 | 32.898 |
| 0 | 33.0 High-Hat 00 | 33.049 |
| 0 | 33.1 Snare 00 | 33.192 |
| 0 | 33.3 Hi-Hat, Kick 00 | 33.335 |
| 0 | 33.4 Snare 00 | 33.486 |
| 0 | 33.6 Snare 00 | 33.626 |
| 0 | 33.7 Hi-Hat 00 | 33.789 |
| 0 | 33.9 JB 00 | 33.906 |
| 0 | 34.079 Hi-Hat - Perhaps | 34.079 |
| 1 | 34.220 Hi-Hat | 34.22 |
| 0 | 34.3 Hi-Hat 00 | 34.373 |
| 0 | 34.5 Hi-Hat 00 | 34.521 |
| 0 | 34.6 Hi-Hat 00 | 34.674 |
| 0 | 34.8 B2 Snare 00 | 34.822 |
| 0 | 34.9 Hi-Hat 00 | 34.979 |
| 0 | 35.1 Hi-Hat 00 | 35.119 |
| 0 | 35.2 Snare 00 | 35.277 |
| 0 | 35.3 Cymbal 00 | 35.398 |
| 0 | 35.5 Snare 00 | 35.557 |
| 0 | 35.7 Hi-Hat, Kick 00 | 35.713 |
| 0 | 35.8 Snare 00 | 35.86 |
| 0 | 35.9 Snare 00 | 35.991 |

| | | |
|---|---|---|
| 0 | 36.1 Hi-Hat 00 | 36.142 |
| 0 | 36.2 JB 00 | 36.28 |
| 0 | 36.4 Hi-Hat 00 | 36.442 |
| 1 | 36.5 B1 Hi-Hat 00 | 36.587 |
| 0 | 36.7 Hi-Hat 00 | 36.744 |
| 0 | 36.8 Hi-Hat 00 | 36.892 |
| 0 | 37.0 Hi-Hat 00 | 37.051 |
| 0 | 37.1 Snare 00 | 37.199 |
| 0 | 37.3 Hi-Hat 00 | 37.351 |
| 0 | 37.5 Hi-Hat 00 | 37.5 |
| 0 | 37.6 Snare 00 | 37.656 |
| 0 | 37.8 Hi-Hat 00 | 37.802 |
| 0 | 37.9 Snare 00 | 37.947 |
| 0 | 38.0 Hi-Hat 00 | 38.09 |
| 0 | 38.2 Snare 00 | 38.246 |
| 0 | 38.3 Snare 00 | 38.383 |
| 0 | 38.5 High-Hat 00 | 38.546 |
| 0 | 38.6 JB 00 | 38.681 |
| 0 | 38.836 Hi-Hat - about here | 38.836 |
| 1 | 38.9 B1 JB "funky" no hi-hat could be found 00 | 38.969 |
| 0 | 39.1 Hi-Hat 00 | 39.136 |
| 0 | 39.2 Hi-Hat 00 | 39.284 |
| 0 | 39.4 Hi-Hat 00 | 39.438 |
| 0 | 39.5 Snare 00 | 39.588 |
| 0 | 39.7 Hi-Hat 00 | 39.737 |
| 0 | 39.8 Hi-Hat 00 | 39.881 |
| 0 | 40.0 JB "A" 00 | 40.016 |
| 0 | 40.1 Hi-Hat 00 | 40.174 |
| 0 | 40.3 Rim Shot 00 | 40.308 |
| 0 | 40.4 JB "Two" 00 | 40.435 |
| 0 | 40.603 Hi-Hat - should be here | 40.603 |
| 0 | 40.7 Hi-Hat 00 | 40.708 |
| 0 | 40.8 Kick 00 | 40.888 |
| 0 | 41.0 Hi-Hat 00 | 41.041 |
| 0 | 41.1 Kick 00 | 41.197 |
| 1 | 41.3 B1 Kick and Band 00 | 41.338 |

### B.1.2 The Funky Drummer break – calculated deviations

Deviations are found by subtracting ‘actual onset times’ from ‘nominal onset times’; where nominal onset times are calculated from a mean tatum period of 148msec. The mean tatum period is calculated by divding the duration of the whole break by the number of tatums in the break.

| | Nominal onset times | Actual onset times | Deviations |
|---|---|---|---|
| 22.3 B1 Hi-Hat 00 | 22.334 | 22.334 | 0.000 |
| 22.4 Hi-Hat 5 | 22.482 | 22.486 | 0.004 |
| 22.6 Hi-Hat 5 | 22.631 | 22.637 | 0.006 |
| 22.7 Hi-Hat 5 | 22.779 | 22.798 | 0.019 |

| | | | |
|---|---|---|---|
| 22.9 Snare 5 | 22.928 | 22.946 | 0.018 |
| 23.1 Hi-Hat 5 | 23.076 | 23.103 | 0.027 |
| 23.2 Hi-Hat 5 | 23.225 | 23.243 | 0.018 |
| 23.3 Hi-Hat 5 | 23.373 | 23.391 | 0.018 |
| 23.5 B3 Hi-Hat 5 | 23.522 | 23.506 | -0.016 |
| 23.6 snare hi-hat 5 | 23.670 | 23.671 | 0.001 |
| 23.8 Hi-Hat 00 | 23.819 | 23.826 | 0.007 |
| 23.9 Hi-Hat snare 5 | 23.967 | 23.971 | 0.004 |
| 24.1 B4 Snare 5 | 24.116 | 24.105 | -0.011 |
| 24.2 Hi-Hat 00 | 24.264 | 24.269 | 0.005 |
| 24.4 Hi-Hat 5 | 24.413 | 24.406 | -0.007 |
| 24.5 Hi-Hat snare 5 | 24.561 | 24.569 | 0.008 |
| 24.7 B1 Hi-Hat 00 | 24.710 | 24.711 | 0.002 |
| 24.8 Hi-Hat 00 | 24.858 | 24.867 | 0.009 |
| 25.0 Hi-Hat 00 | 25.006 | 25.013 | 0.007 |
| 25.1 Hi-Hat 00 | 25.155 | 25.168 | 0.013 |
| 25.3 Snare 00 | 25.303 | 25.317 | 0.014 |
| 25.4 Hi-Hat 00 | 25.452 | 25.472 | 0.020 |
| 25.6 Hi-Hat 00 | 25.600 | 25.613 | 0.013 |
| 25.7 Snare 00 | 25.749 | 25.766 | 0.017 |
| 25.9 Hi-Hat 00 | 25.897 | 25.917 | 0.020 |
| 26.0 Snare 00 | 26.046 | 26.052 | 0.006 |
| 26.1 Hi-Hat & James Brown 00 | 26.194 | 26.177 | -0.017 |
| 26.3 Don't Know 00 | 26.343 | 26.353 | 0.010 |
| 26.4 Snare 00 | 26.491 | 26.496 | 0.005 |
| 26.6 Kick 00 | 26.640 | 26.654 | 0.014 |
| 26.7 High-Hat? 00 | 26.788 | 26.793 | 0.005 |
| 26.9 Snare 00 | 26.937 | 26.951 | 0.014 |
| 27.0 B1 Hi-Hat 00 | 27.085 | 27.090 | 0.005 |
| 27.2 Hi-Hat 00 | 27.233 | 27.243 | 0.010 |
| 27.3 Kick 00 | 27.382 | 27.380 | -0.002 |
| 27.5 Hi-Hat 00 | 27.530 | 27.545 | 0.015 |
| 27.6 Snare beat2 00 | 27.679 | 27.691 | 0.012 |
| 27.8 Hi-Hat 00 | 27.827 | 27.853 | 0.026 |
| 28.2 Hi-Hat 00 | 27.976 | 27.989 | 0.013 |
| 28.4 Snare 00 | 28.124 | 28.150 | 0.026 |
| 28.5 Hi-Hat 00 | 28.273 | 28.289 | 0.016 |
| 28.7 Hi-Hat, Snare 00 | 28.421 | 28.425 | 0.004 |
| 28.8 Snare 00 | 28.570 | 28.580 | 0.010 |
| 29.0 Kick 00 | 28.718 | 28.732 | 0.014 |
| 29.1 Hi-Hat 00 | 28.867 | 28.872 | 0.005 |
| 29.3 Snare 00 | 29.015 | 29.030 | 0.015 |
| 29.4 B1 Hi-Hat and JB 00 | 29.164 | 29.169 | 0.005 |
| 29.6 Hi-Hat 00 | 29.312 | 29.327 | 0.015 |
| 29.7 Hi-Hat 00 | 29.461 | 29.471 | 0.011 |
| 29.9 Hi-Hat 00 | 29.609 | 29.626 | 0.017 |
| 30.0 Snare 00 | 29.757 | 29.770 | 0.013 |
| 30.2 Hi-Hat 00 | 29.906 | 29.927 | 0.021 |

|  |  |  |  |
|---|---|---|---|
| 30.3 JB 00 | 30.054 | 30.074 | 0.020 |
| 30.5 Snare roll? 00 | 30.203 | 30.230 | 0.027 |
| 30.6 Kick 00 | 30.351 | 30.359 | 0.008 |
| 30.8 Snare 00 | 30.500 | 30.522 | 0.022 |
| 30.9 Hi-Hat 00 | 30.648 | 30.672 | 0.024 |
| 31.1 Hi-Hat 00 | 30.797 | 30.810 | 0.013 |
| 31.2 Snare 00 | 30.945 | 30.954 | 0.009 |
| 31.3 Hi-Hat 00 | 31.094 | 31.102 | 0.008 |
| 31.5 JB "ain't it" 00 | 31.242 | 31.245 | 0.003 |
| 31.696 Snare | 31.391 | 31.398 | 0.007 |
| 31.8 B1 Hi-Hat 00 | 31.539 | 31.529 | -0.010 |
| 31.9 Hi-Hat 00 | 31.688 | 31.696 | 0.008 |
| 32.1 High-Hat 00 | 31.836 | 31.843 | 0.007 |
| 32.2 Hi-Hat 00 | 31.984 | 31.994 | 0.010 |
| 32.446 Snare | 32.133 | 32.143 | 0.010 |
| 32.6 Hi-Hat 00 | 32.281 | 32.299 | 0.018 |
| 32.7 Hi-Hat 00 | 32.430 | 32.446 | 0.016 |
| 32.8 Snare 00 | 32.578 | 32.609 | 0.031 |
| 33.0 High-Hat 00 | 32.727 | 32.748 | 0.021 |
| 33.1 Snare 00 | 32.875 | 32.898 | 0.023 |
| 33.3 Hi-Hat, Kick 00 | 33.024 | 33.049 | 0.025 |
| 33.4 Snare 00 | 33.172 | 33.192 | 0.020 |
| 33.6 Snare 00 | 33.321 | 33.335 | 0.014 |
| 33.7 Hi-Hat 00 | 33.469 | 33.486 | 0.017 |
| 33.9 JB 00 | 33.618 | 33.626 | 0.008 |
| 34.079 Hi-Hat - Perhaps | 33.766 | 33.789 | 0.023 |
| 34.220 Hi-Hat | 33.915 | 33.906 | -0.009 |
| 34.3 Hi-Hat 00 | 34.063 | 34.079 | 0.016 |
| 34.5 Hi-Hat 00 | 34.211 | 34.220 | 0.009 |
| 34.6 Hi-Hat 00 | 34.360 | 34.373 | 0.013 |
| 34.8 B2 Snare 00 | 34.508 | 34.521 | 0.013 |
| 34.9 Hi-Hat 00 | 34.657 | 34.674 | 0.017 |
| 35.1 Hi-Hat 00 | 34.805 | 34.822 | 0.017 |
| 35.2 Snare 00 | 34.954 | 34.979 | 0.025 |
| 35.3 Cymbal 00 | 35.102 | 35.119 | 0.017 |
| 35.5 Snare 00 | 35.251 | 35.277 | 0.026 |
| 35.7 Hi-Hat, Kick 00 | 35.399 | 35.398 | -0.001 |
| 35.8 Snare 00 | 35.548 | 35.557 | 0.009 |
| 35.9 Snare 00 | 35.696 | 35.713 | 0.017 |
| 36.1 Hi-Hat 00 | 35.845 | 35.860 | 0.015 |
| 36.2 JB 00 | 35.993 | 35.991 | -0.002 |
| 36.4 Hi-Hat 00 | 36.142 | 36.142 | 0.000 |
| 36.5 B1 Hi-Hat 00 | 36.290 | 36.280 | -0.010 |
| 36.7 Hi-Hat 00 | 36.439 | 36.442 | 0.003 |
| 36.8 Hi-Hat 00 | 36.587 | 36.587 | 0.000 |
| 37.0 Hi-Hat 00 | 36.735 | 36.744 | 0.009 |
| 37.1 Snare 00 | 36.884 | 36.892 | 0.008 |
| 37.3 Hi-Hat 00 | 37.032 | 37.051 | 0.019 |

| | | | |
|---|---|---|---|
| 37.5 Hi-Hat 00 | 37.181 | 37.199 | 0.018 |
| 37.6 Snare 00 | 37.329 | 37.351 | 0.022 |
| 37.8 Hi-Hat 00 | 37.478 | 37.500 | 0.022 |
| 37.9 Snare 00 | 37.626 | 37.656 | 0.030 |
| 38.0 Hi-Hat 00 | 37.775 | 37.802 | 0.027 |
| 38.2 Snare 00 | 37.923 | 37.947 | 0.024 |
| 38.3 Snare 00 | 38.072 | 38.090 | 0.018 |
| 38.5 High-Hat 00 | 38.220 | 38.246 | 0.026 |
| 38.6 JB 00 | 38.369 | 38.383 | 0.014 |
| 38.836 Hi-Hat - about here | 38.517 | 38.546 | 0.029 |
| 38.9 B1 JB "funky" no hi-hat could be found 00 | 38.666 | 38.681 | 0.015 |
| 39.1 Hi-Hat 00 | 38.814 | 38.836 | 0.022 |
| 39.2 Hi-Hat 00 | 38.962 | 38.969 | 0.007 |
| 39.4 Hi-Hat 00 | 39.111 | 39.136 | 0.025 |
| 39.5 Snare 00 | 39.259 | 39.284 | 0.025 |
| 39.7 Hi-Hat 00 | 39.408 | 39.438 | 0.030 |
| 39.8 Hi-Hat 00 | 39.556 | 39.588 | 0.032 |
| 40.0 JB "A" 00 | 39.705 | 39.737 | 0.032 |
| 40.1 Hi-Hat 00 | 39.853 | 39.881 | 0.028 |
| 40.3 Rim Shot 00 | 40.002 | 40.016 | 0.014 |
| 40.4 JB "Two" 00 | 40.150 | 40.174 | 0.024 |
| 40.603 Hi-Hat - should be here | 40.299 | 40.308 | 0.009 |
| 40.7 Hi-Hat 00 | 40.447 | 40.435 | -0.012 |
| 40.8 Kick 00 | 40.596 | 40.603 | 0.007 |
| 41.0 Hi-Hat 00 | 40.744 | 40.708 | -0.036 |
| 41.1 Kick 00 | 40.893 | 40.888 | -0.005 |
| 41.3 B1 Kick and Band 00 | 41.041 | 41.041 | 0.000 |
| | 41.190 | 41.197 | 0.007 |
| | 41.338 | 41.338 | 0.000 |

### B.1.3 The Funky Drummer – covariance of the rate of change of bar-first deviations relative to bar-end deviations from the previous bar

Start deviation difference from deviation at start of previous bar

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 0.004 | 0.006 | -0.003 | 0.002 | -0.008 | 0.007 | 0.005 | | |

End of previous bar deviation difference from end of bar before that

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 0.007 | 0.001 | -0.006 | 0.008 | -0.012 | 0.019 | 0.014 | | |

Start deviation difference over end of previous bar deviation difference

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 0.538 | | 0.538 | 0.200 | 0.680 | 0.351 | 2.308 | ***MEAN*** | ***0.385*** |

## B.1.4 The Funky Drummer – difference between odd and even deviations

| **Bar 1** | **Bar 2** | **Bar 3** | **Bar 4** | **Bar 5** | **Bar 6** | **Bar 7** | **Bar 8** |
|---|---|---|---|---|---|---|---|
| 0.004 | 0.008 | 0.005 | 0.007 | 0.003 | 0.005 | 0.009 | 0.019 |
| 0.013 | 0.007 | 0.017 | 0.009 | 0.008 | 0.005 | 0.011 | 0.006 |
| 0.009 | 0.007 | 0.014 | 0.008 | 0.015 | 0.009 | 0.004 | 0.001 |
| 0.000 | 0.005 | 0.013 | 0.015 | 0.002 | 0.010 | 0.008 | -0.013 |
| 0.017 | -0.013 | -0.012 | -0.010 | -0.005 | 0.011 | -0.003 | -0.014 |
| -0.003 | 0.028 | 0.004 | 0.000 | 0.003 | -0.001 | 0.008 | 0.020 |
| 0.016 | 0.010 | 0.010 | 0.005 | 0.015 | 0.003 | 0.015 | 0.032 |
| 0.015 | 0.010 | 0.010 | 0.019 | 0.025 | 0.014 | 0.007 | 0.008 |
| 0.067 | 0.058 | 0.057 | 0.049 | 0.062 | 0.052 | 0.055 | 0.055 |

## B.1.5 The Funky Drummer – application of the Covert Clock model

Column C shows onset times for the application of the Covert Clock model to nominal onset times of The Funky Drummer break: the first 5 tatums are fit into first beat of a ¾ bar of same length as the normal 4/4 bar of the break, the next 6 into the second beat; and the last 5 into the third beat.

| A | B | C | D | E | F |
|---|---|---|---|---|---|
| Tatum no | Nominal onset time (msec) | Onset time according to Covert Clock | Deviations (difference between nominal and Covert Clock onsets) | Deviations with 12msec swing added to even-numbered tatum onsets | Deviations with 12msec swing added to even-numbered tatum onsets, scaled by 0.5 |
| 1 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| 2 | 0.148 | 0.158 | 0.010 | 0.022 | 0.011 |
| 3 | 0.297 | 0.317 | 0.020 | 0.020 | 0.010 |
| 4 | 0.445 | 0.475 | 0.030 | 0.042 | 0.021 |
| 5 | 0.594 | 0.633 | 0.040 | 0.040 | 0.020 |
| 6 | 0.742 | 0.792 | 0.049 | 0.061 | 0.031 |
| 7 | 0.891 | 0.924 | 0.033 | 0.033 | 0.016 |
| 8 | 1.039 | 1.056 | 0.016 | 0.028 | 0.014 |
| 9 | 1.188 | 1.188 | 0.000 | 0.000 | 0.000 |
| 10 | 1.336 | 1.320 | -0.016 | -0.004 | -0.002 |
| 11 | 1.485 | 1.452 | -0.033 | -0.033 | -0.016 |
| 12 | 1.633 | 1.584 | -0.049 | -0.037 | -0.019 |
| 13 | 1.782 | 1.742 | -0.040 | -0.040 | -0.020 |
| 14 | 1.930 | 1.900 | -0.030 | -0.018 | -0.009 |
| 15 | 2.079 | 2.059 | -0.020 | -0.020 | -0.010 |
| 16 | 2.227 | 2.217 | -0.010 | 0.002 | 0.001 |

## B.2 Tumba francesca values

### B.2.1 Yubá

Values drawn from (Alén, 1995)

**Catá, toque macota, Yubá: 1150-1200msec phrase length**

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| mean deviation (%) | 0.3 | 0.4 | 0.4 | 0.5 | 0.4 | 0.3 | 0.3 | 0.4 |
| Av value of standardised values (%) | 16.4 | 8.0 | 16.6 | 8.7 | 17.3 | 8.0 | 8.4 | 16.6 |
| Cumulative distribution of standardised average values (%) | 16.4 | 24.4 | 41.0 | 49.7 | 67.0 | 75.0 | 83.4 | 100.0 |
| Notation | quarter | eighth | quarter | eighth | quarter | eighth | eighth | quarter |

**Bulá 1, Macota, Yubá**

| | | | | | | |
|---|---|---|---|---|---|---|
| mean deviation (%) | 0.4 | 0.5 | 0.5 | 0.4 | 0.5 | 0.6 |
| Av value of standardised values (%) | 15.9 | 16.3 | 17.8 | 15.8 | 10.1 | 24.1 |
| Cumulative distribution of standardised average values (%) | 15.9 | 32.2 | 50.0 | 65.8 | 75.9 | 100.0 |
| Notation | two eighths | quarter | quarter | quarter | eighth | dotted quarter |

**Bulá 2, Macota, Yubá**

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| mean deviation (%) | 0.3 | 0.4 | 0.5 | 0.3 | 0.3 | 0.4 | 0.5 | 0.4 |
| Av value of standardised values (%) | 6.8 | 8.4 | 16.3 | 18.1 | 6.8 | 8.4 | 11.6 | 23.6 |
| Cumulative distribution of standardised average values (%) | 6.8 | 15.2 | 31.5 | 49.6 | 56.4 | 64.8 | 76.4 | 100.0 |
| Notation | eighth | eighth | quarter | quarter | eighth | eighth | eighth | dotted quarter |

Application of the Covert Clock model to the Yubá, toque Macota, Bulá 1

**Bulá 1, Macota, Yubá**

| Tatum number | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Nominal onset times (msec) | 0.0 | 100.0 | 200.0 | 300.0 | 400.0 | 500.0 | 600.0 | 700.0 | 800.0 | 900.0 | 1000.0 | 1100.0 |
| Covert clock onset time (5/4) | 0.0 | | | 240.0 | | 480.0 | | | 720.0 | 960.0 | | |
| Resulting deviation | 0.0 | | | -60.0 | | -20.0 | | | -80.0 | 60.0 | | |
| Deviation scaled by .17 | 0.0 | | | -10.0 | | -3.3 | | | -13.3 | 10.0 | | |

Application of the Covert Clock model to the Yubá, toque Macota, Bulá 2

**Bulá 2, Macota, Yubá**

| Tatum number | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Nominal onset times (msec) | 0.0 | 100.0 | 200.0 | 300.0 | 400.0 | 500.0 | 600.0 | 700.0 | 800.0 | 900.0 | 1000.0 | 1100.0 |
| Covert clock onset time (5/4) | 0.0 | | | 240.0 | | 480.0 | | | 720.0 | 960.0 | | |
| Resulting deviation | 0.0 | | | -60.0 | | -20.0 | | | -80.0 | 60.0 | | |
| Deviation scaled by 2.9 | 0.0 | | | -10.0 | | -3.3 | | | -13.3 | 10.0 | | |

VROOGE model output for the Yubá, toque Macota, Bulá 2

**Bulá 2, Macota, Yubá**

| Tatum number | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Nominal onset times | 0.0 | 100.0 | 200.0 | 300.0 | 400.0 | 500.0 | 600.0 | 700.0 | 800.0 | 900.0 | 1000.0 | 1100.0 |
| VROOGE model onset times | 0.0 | 92.0 | 187.0 | | 376.0 | | 564.0 | 658.0 | 753.0 | 893.0 | | |
| VROOGE deviation values | *0.0* | *-8.0* | *-13.0* | | *-24.0* | | *-36.0* | *-42.0* | *-47.0* | *-7.0* | | |
| Actual deviation values | 0.0 | -18.0 | -16.8 | | -21.6 | | -4.8 | -22.8 | -21.6 | 18.0 | | |

## B.2.2 Mason

Values drawn from (Alén, 1995)

**Catá, Mason**

| | | | | | |
|---|---|---|---|---|---|
| mean deviation (%) | 0.5 | 0.6 | 0.6 | 0.7 | 0.4 |
| Av value of standardised values (%) | 12.6 | 24.6 | 13.2 | 25.9 | 23.7 |
| Cumulative distribution of standardised average values (%) | 12.6 | 37.2 | 50.4 | 76.3 | 100.0 |
| Notation | eighth | quarter | eighth | quarter | quarter |

**Bulá, Mason**

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| mean deviation (%) | 0.4 | 0.4 | 0.3 | 0.4 | 0.4 | 0.4 | 0.3 | 0.3 | 0.4 |
| Av value of standardised values (%) | 12.5 | 7.1 | 11.2 | 7.6 | 11.2 | 12.2 | 11.9 | 14.9 | 11.4 |
| Cumulative distribution of standardised average values (%) | 12.5 | 19.6 | 30.8 | 38.4 | 49.6 | 61.8 | 73.7 | 88.6 | 100 |
| Notation | triplet quarter + eighth | | triplet quarter + eighth | | triplet eighths | | | eighth | eighth |

Application of the Covert Clock model to the Mason, Tambora

**Tambora, Mason**

| Tatum number | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Nominal onset times (msec) | 0.0 | 100.0 | 200.0 | 300.0 | 400.0 | 500.0 | 600.0 | 700.0 | 800.0 | 900.0 | 1000.0 | 1100.0 |
| Covert clock onset time (5/4) | 0.0 | | 240.0 | 240.0 | | | 480.0 | 720.0 | 720.0 | 960.0 | 960.0 | |
| Resulting deviation | 0.0 | | 40.0 | -60.0 | | | -120.0 | 20.0 | -80.0 | 60.0 | -40.0 | |
| Deviation scaled by 2.9 | 0.0 | | | -13.3 | | | -26.7 | | -17.8 | | -8.9 | |

Application of the Covert Clock model to the Mason, Bulá

**Bulá, Mason**

| Tatum number | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Nominal onset times (msec) | 0.0 | 100.0 | 200.0 | 300.0 | 400.0 | 500.0 | 600.0 | 700.0 | 800.0 | 900.0 | 1000.0 | 1100.0 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Covert clock onset time (5/4) | 0.0 | 240.0 | 240.0 | 480.0 | 720.0 | 720.0 | 960.0 | 960.0 |
| Resulting deviation | 0.0 | 40.0 | -60.0 | -120.0 | 20.0 | -80.0 | 60.0 | -40.0 |
| Deviation scaled by 1.0 | 0.0 | | -60.0 | -120.0 | | -80.0 | | -40.0 |

## B.3 O Grande Amor

Onset timing for O Grande Amor (8 bars only), from the Stan Getz and Joao Gilberto album Getz/Gilberto; data courtesy of Ernest Cholakis

***Getz/Gilberto***

| | |
|---|---|
| tempo (bpm) | 64.9139969 |
| tactus period (ms) | 924.299887 |
| tactus period (samples) | 40761.625 |
| tatum (sixteenth) period (ms) | 231.074972 |
| tatum (sixteenth) period (samples) | 10190.4062 |

***the original tempo is 64.91399693***

***The Number of Beats is 32***

***The Pulse Division is 4 (16th notes)***

***The Loop Start Point in Samples 0***

***The Loop End Point in samples 1304372***

| **Timing in samples (44100Hz sampling rate) of each hi hat event** | **Nominal attack times (samples at 44.1kHz)** | **Deviation (samples at 44.1kHz)** | **Tactus deviation (samples at 44.1 kHz)** | **Bar downbeat deviation (samples at 44.1kHz)** | **Deviation (msec)** | **Tactus deviation (msec)** | **Bar downbeat deviation (msec)** |
|---|---|---|---|---|---|---|---|
| ***490*** | 490 | 0 | 0 | 0 | 0 | 0 | 0 |
| ***10668*** | 10680.40625 | -12.40624957 | | | -0.2813209 | | |
| ***21370*** | 20870.8125 | 499.1875009 | | | 11.3194445 | | |
| ***31630*** | 31061.21875 | 568.7812513 | | | 12.897534 | | |
| ***41848*** | 41251.625 | 596.3750017 | 596.375002 | | 13.5232427 | 13.5232427 | |
| ***51858*** | 51442.03125 | 415.9687522 | | | 9.43239801 | | |
| ***61820*** | 61632.4375 | 187.5625026 | | | 4.25311797 | | |

*(cont)* Onset timing for O Grande Amor (8 bars only), from the Stan Getz and Joao Gilberto album Getz/Gilberto; data courtesy of Ernest Cholakis

| Timing in samples (44100Hz sampling rate) of each hi hat event | Nominal attack times (samples at 44.1kHz) | Deviation (samples at 44.1kHz) | Tactus deviation (samples at 44.1 kHz) | Bar downbeat deviation (samples at 44.1kHz) | Deviation (msec) | Tactus deviation (msec) | Bar downbeat deviation (msec) |
|---|---|---|---|---|---|---|---|
| ***72071*** | 71822.84375 | 248.156253 | | | 5.62712592 | | |
| ***82740*** | 82013.25 | 726.7500034 | 726.750003 | | 16.4795919 | 16.4795919 | |
| ***92972*** | 92203.65625 | 768.3437539 | | | 17.4227609 | | |
| ***103111*** | 102394.0625 | 716.9375043 | | | 16.2570863 | | |
| ***113439*** | 112584.4687 | 854.5312547 | | | 19.377126 | | |
| ***123429*** | 122774.875 | 654.1250052 | 654.125005 | | 14.8327666 | 14.8327666 | |
| ***133745*** | 132965.2812 | 779.7187556 | | | 17.6806974 | | |
| ***144089*** | 143155.6875 | 933.312506 | | | 21.1635489 | | |
| ***154558*** | 153346.0937 | 1211.906256 | | | 27.4808675 | | |
| ***165039*** | 163536.5 | 1502.500007 | 1502.50001 | 1502.50001 | 34.0702949 | 34.0702949 | 34.0702949 |
| ***174924*** | 173726.9062 | 1197.093757 | | | 27.1449832 | | |
| ***185243*** | 183917.3125 | 1325.687508 | | | 30.0609412 | | |
| ***195562*** | 194107.7187 | 1454.281258 | | | 32.9768993 | | |
| ***205498*** | 204298.125 | 1199.875009 | 1199.87501 | | 27.2080501 | 27.2080501 | |
| ***215706*** | 214488.5312 | 1217.468759 | | | 27.6070013 | | |

*(cont)* Onset timing for O Grande Amor (8 bars only), from the Stan Getz and Joao Gilberto album Getz/Gilberto; data courtesy of Ernest Cholakis

| **Timing in samples (44100Hz sampling rate) of each hi hat event** | **Nominal attack times (samples at 44.1kHz)** | **Deviation (samples at 44.1kHz)** | **Tactus deviation (samples at 44.1 kHz)** | **Bar downbeat deviation (samples at 44.1kHz)** | **Deviation (msec)** | **Tactus deviation (msec)** | **Bar downbeat deviation (msec)** |
|---|---|---|---|---|---|---|---|
| ***226063*** | 224678.9375 | 1384.062509 | | | 31.3846374 | | |
| ***235315*** | 234869.3437 | 445.6562599 | | | 10.1055841 | | |
| ***246094*** | 245059.75 | 1034.25001 | 1034.25001 | | 23.4523812 | 23.4523812 | |
| ***256658*** | 255250.1562 | 1407.843761 | | | 31.9238948 | | |
| ***266818*** | 265440.5625 | 1377.437511 | | | 31.2344107 | | |
| ***277395*** | 275630.9687 | 1764.031262 | | | 40.0007089 | | |
| ***287231*** | 285821.375 | 1409.625012 | 1409.62501 | | 31.964286 | 31.964286 | |
| ***297503*** | 296011.7812 | 1491.218762 | | | 33.8144844 | | |
| ***308267*** | 306202.1875 | 2064.812513 | | | 46.8211454 | | |
| ***317808*** | 316392.5937 | 1415.406263 | | | 32.0953801 | | |
| ***328496*** | 326583 | 1913.000014 | 1913.00001 | 1913.00001 | 43.3786851 | 43.3786851 | 43.3786851 |
| ***338824*** | 336773.4062 | 2050.593764 | | | 46.4987248 | | |
| ***349102*** | 346963.8125 | 2138.187515 | | | 48.4849777 | | |
| ***359413*** | 357154.2187 | 2258.781265 | | | 51.2195298 | | |
| ***370042*** | 367344.625 | 2697.375016 | 2697.37502 | | 61.1649663 | 61.1649663 | |

*(cont)* Onset timing for O Grande Amor (8 bars only), from the Stan Getz and Joao Gilberto album Getz/Gilberto; data courtesy of Ernest Cholakis

| Timing in samples (44100Hz sampling rate) of each hi hat event | Nominal attack times (samples at 44.1kHz) | Deviation (samples at 44.1kHz) | Tactus deviation (samples at 44.1 kHz) | Bar downbeat deviation (samples at 44.1kHz) | Deviation (msec) | Tactus deviation (msec) | Bar downbeat deviation (msec) |
|---|---|---|---|---|---|---|---|
| ***379763*** | 377535.0312 | 2227.968766 | | | 50.5208337 | | |
| ***390121*** | 387725.4375 | 2395.562516 | | | 54.3211455 | | |
| ***400241*** | 397915.8437 | 2325.156267 | | | 52.7246319 | | |
| ***410708*** | 408106.25 | 2601.750017 | 2601.75002 | | 58.996599 | 58.996599 | |
| ***421341*** | 418296.6562 | 3044.343768 | | | 69.0327385 | | |
| ***431485*** | 428487.0625 | 2997.937518 | | | 67.9804426 | | |
| ***442103*** | 438677.4687 | 3425.531269 | | | 77.676446 | | |
| ***452474*** | 448867.875 | 3606.125019 | 3606.12502 | | 81.7715424 | 81.7715424 | |
| ***462231*** | 459058.2812 | 3172.718769 | | | 71.9437363 | | |
| ***472628*** | 469248.6875 | 3379.31252 | | | 76.6284018 | | |
| ***482569*** | 479439.0937 | 3129.90627 | | | 70.9729313 | | |
| ***493179*** | 489629.5 | 3549.500021 | 3549.50002 | 3549.50002 | 80.4875288 | 80.4875288 | 80.4875288 |
| ***503627*** | 499819.9062 | 3807.093771 | | | 86.3286569 | | |
| ***513440*** | 510010.3125 | 3429.687522 | | | 77.7706921 | | |
| ***523702*** | 520200.7187 | 3501.281272 | | | 79.3941332 | | |

*(cont)* Onset timing for O Grande Amor (8 bars only), from the Stan Getz and Joao Gilberto album Getz/Gilberto; data courtesy of Ernest Cholakis

| Timing in samples (44100Hz sampling rate) of each hi hat event | Nominal attack times (samples at 44.1kHz) | Deviation (samples at 44.1kHz) | Tactus deviation (samples at 44.1 kHz) | Bar downbeat deviation (samples at 44.1kHz) | Deviation (msec) | Tactus deviation (msec) | Bar downbeat deviation (msec) |
|---|---|---|---|---|---|---|---|
| ***533590*** | 530391.125 | 3198.875022 | 3198.87502 | | 72.5368486 | 72.5368486 | |
| ***543948*** | 540581.5312 | 3366.468773 | | | 76.3371604 | | |
| ***553517*** | 550771.9375 | 2745.062523 | | | 62.2463157 | | |
| ***563799*** | 560962.3437 | 2836.656274 | | | 64.3232715 | | |
| ***573731*** | 571152.75 | 2578.250024 | 2578.25002 | | 58.4637194 | 58.4637194 | |
| ***584179*** | 581343.1562 | 2835.843775 | | | 64.3048475 | | |
| ***594313*** | 591533.5625 | 2779.437525 | | | 63.0257942 | | |
| ***603909*** | 601723.9687 | 2185.031275 | | | 49.5471945 | | |
| ***614286*** | 611914.375 | 2371.625026 | 2371.62503 | | 53.7783453 | 53.7783453 | |
| ***624337*** | 622104.7812 | 2232.218776 | | | 50.6172058 | | |
| ***634571*** | 632295.1875 | 2275.812527 | | | 51.6057262 | | |
| ***644941*** | 642485.5937 | 2455.406277 | | | 55.6781469 | | |
| ***654970*** | 652676 | 2294.000028 | 2294.00003 | 2294.00003 | 52.0181412 | 52.0181412 | 52.0181412 |
| ***665460*** | 662866.4062 | 2593.593778 | | | 58.8116503 | | |
| ***675217*** | 673056.8125 | 2160.187528 | | | 48.9838442 | | |

*(cont)* Onset timing for O Grande Amor (8 bars only), from the Stan Getz and Joao Gilberto album Getz/Gilberto; data courtesy of Ernest Cholakis

| Timing in samples (44100Hz sampling rate) of each hi hat event | Nominal attack times (samples at 44.1kHz) | Deviation (samples at 44.1kHz) | Tactus deviation (samples at 44.1 kHz) | Bar downbeat deviation (samples at 44.1kHz) | Deviation (msec) | Tactus deviation (msec) | Bar downbeat deviation (msec) |
|---|---|---|---|---|---|---|---|
| 685622 | 683247.2187 | 2374.781279 | | | 53.8499156 | | |
| 695801 | 693437.625 | 2363.375029 | 2363.37503 | | 53.5912705 | 53.5912705 | |
| 705592 | 703628.0312 | 1963.96878 | | | 44.5344394 | | |
| 715734 | 713818.4375 | 1915.56253 | | | 43.4367921 | | |
| 725494 | 724008.8437 | 1485.156281 | | | 33.6770132 | | |
| 736279 | 734199.25 | 2079.750031 | 2079.75003 | | 47.1598646 | 47.1598646 | |
| 746347 | 744389.6562 | 1957.343781 | | | 44.3842127 | | |
| 756393 | 754580.0625 | 1812.937532 | | | 41.1096946 | | |
| 766743 | 764770.4687 | 1972.531282 | | | 44.7286005 | | |
| 776634 | 774960.875 | 1673.125033 | 1673.12503 | | 37.9393431 | 37.9393431 | |
| 786749 | 785151.2812 | 1597.718783 | | | 36.2294509 | | |
| 796708 | 795341.6875 | 1366.312534 | | | 30.9821436 | | |
| 806773 | 805532.0937 | 1240.906284 | | | 28.1384645 | | |
| 817598 | 815722.5 | 1875.500034 | 1875.50003 | 1875.50003 | 42.5283455 | 42.5283455 | 42.5283455 |
| 827823 | 825912.9062 | 1910.093785 | | | 43.3127842 | | |
| 837600 | 836103.3125 | 1496.687535 | | | 33.9384929 | | |
| 848127 | 846293.7187 | 1833.281286 | | | 41.5710042 | | |

*(cont)* Onset timing for O Grande Amor (8 bars only), from the Stan Getz and Joao Gilberto album Getz/Gilberto; data courtesy of Ernest Cholakis

| Timing in samples (44100Hz sampling rate) of each hi hat event | Nominal attack times (samples at 44.1kHz) | Deviation (samples at 44.1kHz) | Tactus deviation (samples at 44.1 kHz) | Bar downbeat deviation (samples at 44.1kHz) | Deviation (msec) | Tactus deviation (msec) | Bar downbeat deviation (msec) |
|---:|---:|---:|---:|---:|---:|---:|---:|
| ***857584*** | 856484.125 | 1099.875036 | 1099.87504 | | 24.940477 | 24.940477 | |
| ***867727*** | 866674.5312 | 1052.468787 | | | 23.8655054 | | |
| ***878445*** | 876864.9375 | 1580.062537 | | | 35.8290825 | | |
| ***888202*** | 887055.3437 | 1146.656287 | | | 26.0012764 | | |
| ***898970*** | 897245.75 | 1724.250038 | 1724.25004 | | 39.0986403 | 39.0986403 | |
| ***909269*** | 907436.1562 | 1832.843788 | | | 41.5610836 | | |
| ***919079*** | 917626.5625 | 1452.437539 | | | 32.9350916 | | |
| ***929437*** | 927816.9687 | 1620.031289 | | | 36.7354034 | | |
| ***939443*** | 938007.375 | 1435.62504 | 1435.62504 | | 32.5538558 | 32.5538558 | |
| ***949747*** | 948197.7812 | 1549.21879 | | | 35.1296778 | | |
| ***960268*** | 958388.1875 | 1879.81254 | | | 42.6261347 | | |
| ***970159*** | 968578.5937 | 1580.406291 | | | 35.8368773 | | |
| ***980650*** | 978769 | 1881.000041 | 1881.00004 | 1881.00004 | 42.6530622 | 42.6530622 | 42.6530622 |
| ***990635*** | 988959.4062 | 1675.593792 | | | 37.9953241 | | |
| ***1000128*** | 999149.8125 | 978.1875422 | | | 22.1811234 | | |
| ***1010136*** | 1009340.219 | 795.7812926 | | | 18.0449273 | | |
| ***1020619*** | 1019530.625 | 1088.375043 | 1088.37504 | | 24.6797062 | 24.6797062 | |

*(cont)* Onset timing for O Grande Amor (8 bars only), from the Stan Getz and Joao Gilberto album Getz/Gilberto; data courtesy of Ernest Cholakis

| Timing in samples (44100Hz sampling rate) of each hi hat event | Nominal attack times (samples at 44.1kHz) | Deviation (samples at 44.1kHz) | Tactus deviation (samples at 44.1 kHz) | Bar downbeat deviation (samples at 44.1kHz) | Deviation (msec) | Tactus deviation (msec) | Bar downbeat deviation (msec) |
|---|---|---|---|---|---|---|---|
| ***1030944*** | 1029721.031 | 1222.968793 | | | 27.7317187 | | |
| ***1040927*** | 1039911.437 | 1015.562544 | | | 23.0286291 | | |
| ***1051197*** | 1050101.844 | 1095.156294 | | | 24.8334761 | | |
| ***1061531*** | 1060292.25 | 1238.750045 | 1238.75004 | | 28.0895702 | 28.0895702 | |
| ***1071460*** | 1070482.656 | 977.3437952 | | | 22.1619908 | | |
| ***1081687*** | 1080673.062 | 1013.937546 | | | 22.9917811 | | |
| ***1092380*** | 1090863.469 | 1516.531296 | | | 34.3884648 | | |
| ***1102071*** | 1101053.875 | 1017.125047 | 1017.12505 | | 23.06406 | 23.06406 | |
| ***1112224*** | 1111244.281 | 979.7187969 | | | 22.2158457 | | |
| ***1122518*** | 1121434.687 | 1083.312547 | | | 24.5649104 | | |
| ***1132823*** | 1131625.094 | 1197.906298 | | | 27.1634081 | | |
| ***1143338*** | 1141815.5 | 1522.500048 | 1522.50005 | 1522.50005 | 34.5238106 | 34.5238106 | 34.5238106 |
| ***1152820*** | 1152005.906 | 814.0937987 | | | 18.4601768 | | |
| ***1163308*** | 1162196.312 | 1111.687549 | | | 25.2083344 | | |
| ***1173399*** | 1172386.719 | 1012.2813 | | | 22.9542245 | | |
| ***1183618*** | 1182577.125 | 1040.87505 | 1040.87505 | | 23.6026088 | 23.6026088 | |
| ***1193827*** | 1192767.531 | 1059.4688 | | | 24.0242358 | | |

*(cont)* Onset timing for O Grande Amor (8 bars only), from the Stan Getz and Joao Gilberto album Getz/Gilberto; data courtesy of Ernest Cholakis

| Timing in samples (44100Hz sampling rate) of each hi hat event | Nominal attack times (samples at 44.1kHz) | Deviation (samples at 44.1kHz) | Tactus deviation (samples at 44.1 kHz) | Bar downbeat deviation (samples at 44.1kHz) | Deviation (msec) | Tactus deviation (msec) | Bar downbeat deviation (msec) |
|---|---|---|---|---|---|---|---|
| ***1203508*** | 1202957.937 | 550.0625508 | | | 12.4730737 | | |
| ***1213715*** | 1213148.344 | 566.6563012 | | | 12.8493492 | | |
| ***1223698*** | 1223338.75 | 359.2500517 | 359.250052 | | 8.14625968 | 8.14625968 | |
| ***1233856*** | 1233529.156 | 326.8438021 | | | 7.41142408 | | |
| ***1243539*** | 1243719.562 | -180.5624475 | | | -4.0943866 | | |
| ***1253639*** | 1253909.969 | -270.968697 | | | -6.1444149 | | |
| ***1263868*** | 1264100.375 | -232.3749466 | -232.37495 | | -5.2692732 | -5.2692732 | |
| ***1273758*** | 1274290.781 | -532.7811962 | | | -12.081206 | | |
| ***1284087*** | 1284481.187 | -394.1874457 | | | -8.9384908 | | |
| ***1293723*** | 1294671.594 | -948.5936953 | | | -21.510061 | | |

# Appendix C Paper read at the 7th International Conference on Cognitive Science, Sydney, July 2003

# Do discontinuous temporal windows in metrical music point to open-loop motor control? A musicological and cross-cultural analysis of rhythmic timing planning.

**Andrew McGuiness (andymcguiness@yahoo.com.au)**
Australian Centre for the Arts & Technology
Australian National University, Canberra ACT 0200

## Abstract

Metrical music is commonly perceived as segmented into phrases at durations above the level of the tactus, up to temporal intervals of approximately 3 secs. These segments are normally contiguous and non-overlapping. Examples of similar temporal segmentation are found in human utterances and in human mother-infant interactions. Standard reference beat patterns such as the Cuban clave and African bell patterns are in this order of duration, and have rotation symmetry only at the level of the whole pattern. The start point of the patterns is either musically inherent or culturally understood. Various researchers have found the existence of musically meaningful deviations of less than 60ms from quantised rhythm, a level of timing control not possible through closed-loop kinaesthetic feedback control,which has a cycle time in the order of 100ms. I propose that reference beat patterns provide a guide to performers in a group as to the start point of a recurring open-loop timing control interval.

## Introduction

This paper examines the evidence for the existence of open-loop control mechanisms for the perception and production of rhythms in metrical music over durations of ~3000ms.

For convenience, the “moderate tempo” pulse identified by (Parncutt, 1994) and others (Fraisse, 1982; Moelants, 2002) is here referred to as the tactus and its period is assumed to vary around ~600msec, with lower and upper limits roughly half and double that period respectively. Tactus subdivisions are discussed, and the paper also deals with finer-grained timing control, below the level of tactus subdivisions.

A review of motor theory relevant to control at tactus and tactus subdivision levels is followed by a review of Liberman and Mattingly’s motor theory of speech perception. (Liberman & Mattingly, 1985) The fine motor control at tolerances of less than 60msec required for speech is related to systematic timing deviations from strict meter, found in a variety of musics. (Beran & Mazzola, 2000; Bilmes, 1993b; Freeman & Lacey, 2002; Stange-Elbe & Mazzola, 1998)

Temporal segmentation at the ~3 second level in human communication is identified with the psychological, or subjective, present (Turner & Pöppel, 1988). An important feature of the subjective present as identified here is its discontinuous, non-overlapping nature. Finally, a musicological discussion of period segmentation in traditional Central African and West African drumming, in Afro-Cuban music, and in North Indian music is related to the subjective present.

A model is proposed in which the subjective present is seen as a window in which fine temporal control of motor actions can be projected via a feed forward model, and carried out without reference to peripheral feedback.

## Motor preplanning

Motor programs were defined by Keele, writing in 1968, (S. W. Keele, 1968) as “a set of muscle commands that are structured before a movement begins, and that allows the entire sequence to be carried out uninfluenced by peripheral feedback”. Peripheral feedback may include both sensory and kinaesthetic feedback.

(Freund, 1989) showed that only hand movements with a frequency of 2 HZ (500ms intervals) are under visual control, with faster movements resulting from “subconscious automatic processes controlled by the cerebellum.” (Seifert, Olk, & Schneider, 1995) Wing suggested that delayed tapping adjustments to small temporal perturbations (between –10msec and +50msec) in delayed feedback for 350msec periods indicated that normal operation of isochronous tapping was open loop. Wing concluded that repetitive response timing was normally under open loop control, but that subjects did process auditory feedback associated with the responses on some occasions. (Wing, 1977, 1980)

### Kinaesthetic control at the tactus subdivision level

Keele reviewed experimental studies that tested the use of kinaesthetic feedback. Subjects with high kinaesthetic sensitivity became progressively better at the Two-hand Coordination Task than subjects with low kinaesthesic sensitivity. Kinaesthetic responses to a sudden increase in tension opposing a movement were in the order of .16 sec. (S. W. Keele, 1968)

Adams made a case for kinaesthetic feedback involvement in fast movements, such as the fastest movements made by a pianist's fingers, citing an EMG study by Sears and Davis (Sears & Davis, 1968) where EMG response in the respiratory muscles to a change of pressure load on the lungs showed a latency in the 50-60msec range. (Adams, 1976) It therefore appears that while sensory feedback may not be required for sequences of actions, peripheral kinaesthetic feedback may be used for control, and sequences may not be controlled exclusively by strictly open loop motor programs.

### The separate timing control theory

Later conceptions of motor program theory, taking account of growing evidence against a feedback-based account of timing, do not deny that feedback (including proprioceptive feedback) plays an essential role in movement control, but suggest that feedback-response associations are not the mechanism used to time sequences of movements. (Summers & Burns, 1990) Summers & Burns argue for a separate representation for timing in motor programs, on the basis of studies showing that relative timing within a movement series remains constant across changes in the speed and size of action.

Analyses by Summers of learned keypress sequences (Summers & Burns, 1990), and by Terzuolo and Viviani of the timing of keystrokes in typing and in handwriting (Terzuolo & Viviani, 1980; Viviani & Terzuolo, 1980) also supported separate encoding of timing; as did a study of speech by (MacKay & Bowman, 1969).

## Timing below the tactus-subdivision level

William Calvin, in several articles and books (Calvin, 1983, 1993), has argued that the evolution of accurate throwing for hunting (rather than threat-throwing, as practised by chimpanzees) involved a rise in the facility for offline planning of novel sequences of actions. (Calvin, 1983, 1993) Of particular interest is the fine control of timing required for accurate throwing. Calvin calculated that, to hit a rabbit-sized animal at 4 metres using an overarm throwing technique, release of the projectile needs to occur within an 11msec window. At 8 metres, the angle subtended by the target is halved and the launch velocity must be doubled to reach the target, so the release window required for accuracy shrinks to only a few milliseconds. (Calvin, 1983)

Humans can perceive the serial order of sounds in a familiar word such as "sand" at rates of 20msec per segment (R.M. Warren & Warren, 1970); with training (800 repetitions) listeners were able to identify the order of non-speech sounds (hiss-vowel-buzz-tone) with durations of less than 20msec. (R.M. Warren, 1974)

Shaffer et al examined durations of notes in "equal-note" groups (according to the notation), averaged over repeated performances by the same pianist of a piano piece by Erik Satie. Within-group differences were of the order of 50msec-100msec; Shaffer states that differences in form between different 4-note groups in the piece were large(Timmers, Ashley, Desain, & Heijink, 2000)r than differences between repetitions in different performances of the same group. (Shaffer, E.F.Clarke, & N.P.Todd, 1985)

### The motor theory of speech perception

In their 1985 paper on a revised motor theory of speech perception, Liberman & Mattingly (Liberman & Mattingly, 1985) marshal evidence for the idea that speech perception distinguishes motor commands, rather than auditory events. They make a claim a specialised phonetic module that prevents listeners from hearing speech signals as ordinary sound, but translates auditory patterns into perceptions of the underlying phonetic gestures. (Liberman & Mattingly, 1985) Thus, perception of the distinction of the syllable ba versus pa would not (in the motor theory) depend on a general auditory ability to perceive temporal disparity as such, as suggested by (Pisoni, 1977), but would indicate unmediated perception of a gesture involving relative timing of vocal tract opening and start of laryngeal vibration. (Lisker & Abramson, 1964) Liberman & Mattingly argue that, since the distal object of perception in speech is neuromuscular processes giving rise to the phonetic gesture, perception must be carried out by analogy with neuromuscular processes within the listener; that is, the phonetic module is characterised by a link between perception and production. (Liberman & Mattingly, 1985)

It is clear from Liberman & Mattingly's examples that a fine-grained temporal control is required for speech at a level below that of the tactus subdivision, and down to the 10msec region. A noteworthy feature of perception in this region of temporal difference is that the perception comes whole: the perceiver can distinguish between percepts which depend on small temporal differences for their difference, but do not have access to the acoustic properties which trigger one percept rather than another.

Broadbent & Ladefoged found that untrained listeners could not initially discriminate the order of pairs of sounds with onsets separated by 150msec. With practice, they were able to discriminate order at interonset intervals of 30msec – not by perceiving order, but by a distinguishing each pattern by its perceived 'quality'. (Broadbent & Ladefoged, 1959)

Warren et al found subjects could distinguish between same and different orders of three synthesized vowels with item durations ranging from 10msec through to 5000msec, and with accuracy at all durations. At durations below 30msecs, the sequence of vowels were heard as temporal compounds, and distinguished by their timbre. (Richard M. Warren, 1993)

## Microtiming – expressive deviation of rhythm in music

Various authors have undertaken studies of expressive timing variation in Western notated art music. Typically, this involves both local tempo changes as well as systematic variation of tactus subdivision durations that do not affect tempo. (Beran & Mazzola, 2000; Bilmes, 1993b; Freeman & Lacey, 2002; Stange-Elbe & Mazzola, 1998) Less work appears to have been undertaken on systematic timing deviation in groove music, where tempo, if not precisely invariant, does not vary locally. (Arom, 1991; Bilmes, 1993a; Iyer, 1998; Ladzekpo, 1995) Freeman & Lacey's work (Freeman & Lacey, 2002) on jazz swing in recordings by individual performers has identified consistent delays on specific beats of the bar that fall within 30msec windows. Bilmes applied computer beat extraction techniques to a multitrack recording of a Cuban drum ensemble, and found the variance for attacks of three of the drums, in relation to a reference timing pattern. Analysis of the deviations showed repeating patterns of deviation, mainly in a range of approximately 20msec-80msec ahead or behind the smallest tactus subdivision. (Bilmes, 1993a, 1993b)

## Forward models for physiological motor control

Keele & Summers put forward a theory of motor programs in which "a central representation of a motor sequence […] can in the absence of error be initiated and carried out without subsequent stimulation from kinaesthetic feedback." (Steven W. Keele & Summers, 1976)

Miall & Wolpert give one of the potential uses of forward models as distal supervised learning – where the goal and outcome of a movement are defined in task-related coordinates. (Miall & Wolpert, 1996) In terms of a physiological theory of rhythm, the coordinates would be in dimensions of both space and time, and might reference, for example, a particular point to strike on the drum head, at a particular point in the meter. Time coordinates could be given in terms of an hierarchical meter involving a reference pattern and tactus subdivisions (see the discussion of metrical reference patterns, below).

An important feature of the model I propose is that forward models of motor control may be used to achieve fine timing control of rhythm at timing resolutions below the tactus subdivision level, where neither sensory nor kinaesthetic feedback can provide online feedback.

## The psychological present

Pöppel defined the "subjective present" as the temporal integrating level where temporal gestalt perception appears. (E. Pöppel, 1983)

Turner and Pöppel state that a human speaker will pause for a few milliseconds at about three second intervals, in order to determine syntax and wording of the next three seconds. They assert that meter in metered poetry functions to synchronise the speaker with the hearers' "so that each person's three-second present is in phase with the others". (Turner & Pöppel, 1988)

Dissanayake, in her survey of studies of mother-infant early interaction noted that mothers and babies participate in coactive (occurring almost simultaneously) and alternating kinesic patterns of emotionally expressive behaviours of face and body. (Dissanayake, 2000)

Pöppel provided a review of evidence for temporal segmentation at the 3 second level. Speakers, independently of language or age, tend to construct closed verbal utterances up to about 2 to 3 seconds. (Kowal, O'Connell, & Sabin, 1975; E. Pöppel, 1988; Vollrath, Kazenwadel, & Krüger, 1992) The duration of spontaneous intentional acts falls in a temporal window of approximately 3 seconds. (Schleidt, Eible-Eibesfeldt, & Pöppel, 1987) Pöppel concludes that there is a general temporal segmentation mechanism, and he terms the single states of 3-second segmentation "states of being conscious" (STOBCON, each representing "a mental island of activity distinctly separated from the temporally neighbouring ones". (Ernst Pöppel, 1996)

## Period segmentation in traditional Central African music

An important device in the organization of central African rhythm is the production of macroperiods through the superposition of two or more periods of different length, in a ratio to each other of (for example) 2:3 or 3:4. The macroperiod in this case is defined by the coincidence of the beginning of all periods, which occurs at a tactus count of their lowest common denominator. (Arom, 1991) The periods themselves comprise groupings of subdivisions of the tactus (which may or may not make up a whole number of tactus pulses). How long them, are the macroperiods defined by these different devices?

The rhythmic cells for the dance of the Dakpa male initiation rites carry over two pulses, and Arom gives the tempo as 112 bpm, making the longest period defined only about 1070msec. However, Arom gives several two- and three-part examples drawn from Aka pygmy music of macroperiods four pulses long, and six three-part examples which are eight pulses long – even at a faster tempo of 160 bpm, these macroperiods would be 3000msec long.

## The Derler system

The Derler system was developed as a formal theory by Alfons Dauer, Rudolf Derler and Michael Pfob (Dauer, 1988, 1989). Taking the number of fastest pulses in a recurring cycle (ie. subdivisions of the tactus) as the form number, its subset of the number of actual onsets in the cycle gives the beat number. The set of all possible positions of the onsets in one cycle is called the rhythm set. (Seifert et al., 1995) A form number of five and a beat number of three allows 10 possible distributions of the three beats on the five pulses, or 10 elements in the rhythm set. The 10 elements of the form-5, beat-3 rhythm set are shown in Table 1.

Rhythms A, C, F, G and J in Table 1 form an equivalence set in the Derler system. Rhythms B, D, E, H and I form the only other equivalence set for form number 5 with beat number 3. (Seifert et al., 1995)

Table 1. Drawn from a figure provided in (Seifert et al., 1995).
Form number = 5; beat number = 3
**x** represents a realized beat; **.** represents an empty fastest pulse.

| | | | | | |
|---|---|---|---|---|---|
| **A** | **x** | **x** | **x** | **.** | **.** |
| **B** | **x** | **x** | **.** | **x** | **.** |
| **C** | **x** | **x** | **.** | **.** | **x** |
| **D** | **x** | **.** | **x** | **x** | **.** |
| **E** | **x** | **.** | **x** | **.** | **x** |
| **F** | **x** | **.** | **.** | **x** | **x** |
| **G** | **.** | **x** | **x** | **x** | **.** |
| **H** | **.** | **x** | **x** | **.** | **x** |
| **I** | **.** | **x** | **.** | **x** | **x** |
| **J** | **.** | **.** | **x** | **x** | **x** |

## Bell patterns and other reference patterns

The circular nature of the time cycle in Central and West African drumming is asserted by (Anku, 2000). The cycle is defined by a recurrent bell pattern, which however may not necessarily be realised in all musical situations. Only one element of the equivalence class from the appropriate rhythm set is used by a particular culture, while other cultures will utilise a different element from the same equivalence set. Different elements of the same equivalence class, as used by three different cultures, are shown in Table 2.

Table 2. Drawn from figures given in (Anku, 2000).

<table>
<tr><th>Region</th><th>Culture</th><th>Bell pattern<br>(equivalence class element)</th></tr>
<tr><td>Ghana</td><td></td><td><b>x . x . x x . x . x . x</b></td></tr>
<tr><td></td><td>Anlo Ewe people</td><td>| | | |</td></tr>
<tr><td>Central Africa</td><td></td><td><b>x . x x . x . x . x x .</b></td></tr>
<tr><td></td><td>Bemba people</td><td>| | | |</td></tr>
<tr><td>Nigeria</td><td></td><td><b>x . x . x . x x . x . x</b></td></tr>
<tr><td></td><td>Yoruba people</td><td>| | | |</td></tr>
</table>

The only difference between the patterns shown in figure 2 is their relation to the tactus (shown beneath the bell rhythm), and the point in the pattern at which they start. Once a starting point for the pattern is fixed (by cultural convention), the expert listener can always unambiguously identify the starting point of the whole pattern – the macroperiod – from hearing the realised pattern.

The standard clavé pattern of Afro-Cuban music functions in a similar way (Figure 3). Here, although cultural agreement still applies, the starting point can be argued to be perceptually determined through experimentally determined onset salience and grouping effects. Only the first onset and the last onsets in the pattern actually fall on a tactus. The second of the two realised beats at the end of the pattern separated by a distance of two fastest pulses will tend to be heard with an accent, as the second of the two is followed by a longer empty interval. The first of those two beats might otherwise be heard as the beginning of the group, as it follows the longest empty interval (3 fastest pulses). Instead, the other group of 3 empty pulses that follows the two end beats produces a grouping effect, so that the next beat is heard as the beginning of the pattern. The clavé, like the West African bell patterns shown in Table 2, exhibits no rotational symmetry, except at the level of the whole pattern. It functions to effectively define the start and duration of a macroperiod, in the same way as African bell patterns.

Table 3. Cuban clavé pattern

<table>
<tr><td></td><td><b>x . . x . . x . . . x . x . . .</b></td></tr>
<tr><td>tactus</td><td>| | | |</td></tr>
</table>

Similarly, the first beat of the tal in North Indian music does not need to be marked with a special gesture, despite its structural importance in the music, because it is determined unambiguously by the asymmetrical nature of the clap pattern - musicians and musically-educated listeners know where the first beat lies. The tal as a whole, however, is always explicit. (Clayton, 1997)

The clavé, the West African bell patterns, and the tal all seem designed to synchronise the phase of the "specious present" for musicians and listeners alike, in a similar way to the universal 10-syllable line of metered verse, noted by Turner and Pöppel above.

## Summary

The physical embodiment of rhythm in movement and dance to groove music (including traditional African drumming, and African diasporic musics) has been discussed by many authors. (Arom, 1991; Blacking, 1973; Iyer, 1998; Ladzekpo, 1995) Pressing (Pressing, 2002) suggest that stimulation to movement assists memory by providing multiple codings of timing (auditory and motor). I propose that a common forward planning mechanism for physiological control of timing may be at work in rhythmic behaviours which segment time into discontinuous periods of approximately 3 seconds (the "psychological present"); and that it is mechanism which provides the control of timing required for both fluent coarticulation of phonemes in speech, and the systematic microtiming deviations found in groove musics.

## Acknowledgments

This paper was written with the support of the Australian Centre for the Arts & Technology. Dr Kate Stevens of the Macarthur Auditory Research Centre provided comments on early drafts. Conference attendance was supported by ACAT.

## References

Adams, J. A. (1976). Issues for a Closed-loop Theory of Motor Learning. In G. E.Stelmach (Ed.), *Motor Control. Issues and Trends*. New York: Academic Press.

Anku, W. (2000). Circles and Time: A Theory of Structural Organization of Rhythm in African Music. *Music Theory Online, 6*(1).

Arom, S. (1991). *African polyphony and polyrhythm* (M. Thom, B. Tuckett & R. Boyd, Trans.). Cambridge: Cambridge University Press.

Beran, J., & Mazzola, G. (2000). Timing Microstructure in Schumann's "Träumerei" as an Expression of Harmony, Rhythm, and Motivic Structure in Music Performance. *Computers and Mathematics with Applications, 39*, 99-130.

Bilmes, J. (1993a). *Techniques to foster Drum Machine Expressivity.* Paper presented at the Proceedings ICMC.

Bilmes, J. (1993b). *Timing is of the Essence: Perceptual and Computational Techniques for Representing, Learning, and Reproducing Expressive Timing in Percussive Rhythm.* Unpublished Masters Thesis, MIT.

Blacking, J. (1973). *How musical is man?* Seattle: University of Washington Press.

Broadbent, D. E., & Ladefoged, P. (1959). Auditory perception of temporal order. *Journal of the Acoustical Society of America, 31*, 1539-1540; cited in Warren, 1993.

Calvin, W. H. (1983). A Stone's Throw and its Launch Window: Timing Precision and its Implications for Language and Hominid Brains. *Journal of Theoretical Biology, 104*, 121-135.

Calvin, W. H. (1993). The unitary hypotheis: A common neural circuitry for novel manipulations, language, plan-ahead, and throwing? In K. R. Gibson & T. Ingold (Eds.), *Tools, Language, and Cognition in Human Evolution*. Cambridge, Mass: Cambridge University Press.

Clayton, M. (1997). Le metre et le tal dans la musique de l'Inde du Nord. *Cahiers de Musiques traditionelles, 10*, 169-189 http://www.open.ac.uk/Arts/music/nclayton/metre.htm; trans. Georges Goormaghtigh.

Dauer, A. M. (1988). Derler 1: ein System zur klassifikation von Rhythmen. Musiktheoretische und musikhistorische Aspekte. *Jazz Research, 20*, 117-154, cited in Seifert, U.W. et al 1995.

Dauer, A. M. (1989). Das Derler-system under der Jazzrhythmus. Berichtung und Ergänzung. *Jazz Research, 21*, 137-151, cited in Seifert, U.W. et al 1995.

Dissanayake, E. (2000). Antecedents of the Temporal Arts in Early Mother-Infant Interaction. In N. L. Wallin, B. Merker & S. Brown (Eds.), *The Origins of Music*. Cambridge, Massachusets: MIT Press.

Fraisse, P. (1982). Rhythm and Tempo. In D. Deutsch (Ed.), *The Psychology of Music*. New York: Academic Press, Inc.

Freeman, P., & Lacey, L. (2002). *Swing and groove: Contextual rhythmic nuance in live performance.* Paper presented at the 7th International Conference on Music Perfception and Cognition, Sydney.

Freund, F. (1989). Handmotorik und musikalisches Lernen. In H. Petsche (Ed.), *Musik - gehirn - Spiele. Beiträge zum vierten Herbert von Karajan-Symposium*. Basel; cited in Seifert et al, 1995: Birkhäuser.

Iyer, V. S. (1998). *Microstructures of Feel, Macrostructures of Sound: Embodied Cognition in West African and African-American Musics.* University of California, Berkeley.

Keele, S. W. (1968). Movement control in skilled motor performance. *Psychological Bulletin, 70*, 387-403.

Keele, S. W., & Summers, J. J. (1976). The Structure of Motor Programs. In G. E. Stelmach

(Ed.), *Motor Control. Issues and Trends*. New York: Academic Press.

Kowal, S., O'Connell, D., & Sabin, E. (1975). Development of temporal patterning and vocal hesitations in spontaneous narratives. *Journal of Psycholinguistic Research, 4*, 195-207.

Ladzekpo, C. K. (1995). *Foundation Course in African Dance-Drumming*. Retrieved 31/5/03, 2003

Liberman, A. M., & Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition, 21*, 1-36.

Lisker, L., & Abramson, A. (1964). A cross-language study of voicing in initial stops: Acoustical measurement. *Word, 20*, 384-422; cited in Liberman & Mattingly, 1985.

MacKay, D. G., & Bowman, R. W. (1969). On producing the meaning in sentences. *American Journal of Psychology, 82*, 23-39; cited in MacKay, 1982.

Miall, R. C., & Wolpert, D. M. (1996). Forward Models for Physiological Motor Control. *Neural Networks, 9*(8), 1265-1279.

Moelants, D. (2002). *Preferred tempo reconsidered.* Paper presented at the 7th international Conference on Music Perception and Cognition, Sydney.

Parncutt, R. (1994). A Perceptual Model of Pulse Salience and Metrical Accent in Musical Rhythms. *Music Perception, 11*(4), 409-464.

Pisoni, D. B. (1977). Identification and discrimination of the relative onset of two component tones: Implications for the perception of voicing in stops. *Journal of the Accoustical Society of America, 61*, 1352-1361; cited in Liberman & Mattingly, 1985.

Pöppel, E. (1983). Musik-erleben und Zeit-struktur. In E. H. Gombrich & H. Petsche (Eds.), *Auge macht Bild - Ohr macht Klang - hirn macht Welt*. Wien, cited in Seifert et al, 1995: Franz Deuticke.

Pöppel, E. (1988). *Mindworks, Time and Conscious Experience*. Boston: Harcourt Brace Jovanich; cited in Pöppel 1996.

Pöppel, E. (1996). Reconstruction of subjective time on the basis of hierarchically organized processing system. In M. A. Pastor & J. Artieda (Eds.), *Time, Internal Clocks and Movement*. Amsterdam: Elsevier.

Pressing, J. (2002). Black Atlantic rhythm: Its Computational and Transcultural Foundations. *Music Perception, 19*(3), 285-330.

Schleidt, M., Eible-Eibesfeldt, I., & Pöppel, E. (1987). A universal constant in temporal segmentation of human short-term behaviour. *Naturwissenschaften, 74*, 289-290; cited in Pöppel 1996.

Sears, T. A., & Davis, J. N. (1968). *Annals of the New York Academy of Science, 155*, 183-190; cited in Adams, 1976.

Seifert, U., Olk, F., & Schneider, A. (1995). On Rhythm Perception: Theoretical Issues, Empirical Findings. *Journal of New Music Research, 24*, 164-195.

Shaffer, L. H., E.F.Clarke, & N.P.Todd. (1985). Metre and rhythm in piano playing. *Cognition, 20*, 61-77.

Stange-Elbe, J., & Mazzola, G. (1998). *"Cooking a Canon with RUBATO" Performance Aspects of J.S. Bach's "Kunst der Fuge".* Paper presented at the Proceedings ICMC, San Francisco.

Summers, J. J., & Burns, B. D. (1990). Timing in Human Movement Sequences. In R. A. Block (Ed.), *Cognitive models of psychological time*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Terzuolo, C. A., & Viviani, P. (1980). Determinants and characteristics of motor patterns used for typing. *Neuroscience, 5*, 1085-1103; cited in rosenbaum, 1991.

Timmers, R., Ashley, R., Desain, P., & Heijink, H. (2000). The Influence of Musical Context on Tempo Rubato. *Journal of New Music Research, 29*(2), 131-158.

Turner, F., & Pöppel, E. (1988). Metered Poetry, the Brain, and Time. In I. Rentschler, B. Herzberger & D. Epstein (Eds.), *Beauty and the Brain: Biological Aspects of Aesthetics*. Basel: Birkhäuser Verlag.

Viviani, P., & Terzuolo, C. (1980). Space-time invariance in learned motor skills. In G. E. Stelmach & J. Requin (Eds.), *Tutorials in Motor Behavior*. Amsterdam: North-Holland Publishing Company.

Vollrath, M., Kazenwadel, J., & Krüger, H.-P. (1992). A universal constant in temporal segmentation of human speech. *Naturwissenschaften, 79*, 479-480.

Warren, R. M. (1974). Auditory temporal discrimination by trained listeners. *Cognitive Psychology, 6*, 237-256; cited in MacKay, 1987.

Warren, R. M. (1993). Perception of acoustic sequences: global integration versus temporal resolution. In S. McAdams & E. Bigand (Eds.), *Thinking in Sound - The Cognitive PSychology of Human Audition*. Oxford: Oxford University Press.

Warren, R. M., & Warren, R. P. (1970). Auditory illusions and confusions. *Scientific American, 223*, 30-36; cited in MacKay, 1987.

Wing, A. M. (1977). Perturbations of Auditory Feedback Delay and the Timing of Movement. *Journal of Experimental Psychology: Human Perception and Performance, 3*(2), 175-186.

Wing, A. M. (1980). The long and short of timing in response sequences. In G. E. Stelmach & J. Requin (Eds.), *Tutorials in Motor Behavior*. Amsterdam: North Holland Publishing Company.