In time with the music: The concept of entrainment and its significance for ethnomusicology

Martin Clayton, Rebecca Sager and Udo Will
Abstract

Entrainment, broadly defined, is a phenomenon in which two or more independent rhythmic processes synchronize with each other. To illuminate the significance of entrainment for various directions of music research and promote a nuanced understanding of the concept among ethnomusicologists, this publication opens with an exposition of entrainment research in various disciplines, from physics to linguistics and psychology, while systematically introducing basic concepts that are directly relevant to musical entrainment. Topics covered include consideration of self-synchrony and interpersonal synchrony in musical performance, humans’ innate propensities to entrain, the influence of cultural and personal factors on entrainment, the numerous functions of musical entrainment in individual health, socialization, and cultural identification, and a presentation of methodologies and analytical techniques. Finally, some case studies illustrating one methodological strand, that of chronometric analysis, exemplify how the application of the entrainment concept might lead to an understanding of music making and music perception as an integrated, embodied and interactive process.

1. Introduction

The aim of this publication is to stimulate research in ethnomusicology informed by the concept of entrainment, which describes the interaction and consequent synchronization of two or more rhythmic processes or oscillators. Entrainment as a concept has a considerable history - it was first identified by the Dutch physicist Christiaan Huygens in 1665 and has been applied widely in mathematics and in the physical, biological, and social sciences. It is a process that manifests in many ways, some of which involve human agency or cognition. Strangely though, it has had relatively little impact to date in studies of music, where it might be thought particularly relevant, and is only beginning to be seriously considered within ethnomusicology. This is, to our knowledge, the first publication to address the concept in detail from an ethnomusicological perspective.1

In music research, we have already seen an entrainment perspective adopted in the study of musical metre, particularly in the 1990s. Instead of looking for musical cues transmitted from performer to listener as the sole determinants of time and metre percepts, music psychologists have begun to apply an entrainment model in which rhythmic processes endogenous to the listener entrain to cues in the musical sound (Large and Kolen 1994). Although there is much to be done in this area, the entrainment model seems to reflect the cognitive processes much better than do previous models of metrical perception. Some recent work also points to new perspectives offered by the entrainment concept in the study of proto-musical behaviour in infants, and in the evolution of musical behaviour in the human species (Trevarthen 1999-2000, Merker 2000).

We believe that this concept could have a particularly significant impact if applied to ethnomusicological research because it offers a new approach to understanding music making and music perception as an integrated, embodied and interactive process, and can therefore shed light on many issues central to ethnomusicological thought. Entrainment may be central to an ethnomusicological orientation for which performance and listening are the focus of interest. Such a development is likely to be more productive

---

1 The authors of this paper have developed a collaborative approach through organising two panels at meetings of the European Seminar in Ethnomusicology, in Rauland, Norway (2001) and Druskininkai, Lithuania (2002). This article is partly a response to suggestions from some of those present, that ethnomusicologists would benefit from a detailed description of the entrainment concept, presented with methodological suggestions and examples.
if researchers share an understanding of what the concept implies (as well as what its limitations might be); we offer this contribution to colleagues in that spirit.

In later sections we discuss the relationship between entrainment and ethnomusicological research, and offer some suggestions regarding methods for investigation. Before that, however, it is important to establish what entrainment means and how other disciplines have characterised its main features. What follows comprises a summary overview of the concept of entrainment and its application in other fields, particularly the biological and social sciences (sections 2 and 3); some consideration of applications to date in music studies (section 4); some suggestions on the relevance of entrainment in ethnomusicology, with a preliminary consideration of possible methods (sections 5 and 6); and finally some case studies illustrating one methodological strand, that of chronometric analysis. Our focus is, of course, on music, but by discussing work in other fields we first aim to build up a more nuanced picture of entrainment and its ramifications, so that ethnomusicological research can build more productively on this base.

2. Entrainment basics

2.1 Introduction: the concept of entrainment

Entrainment describes a process whereby two rhythmic processes interact with each other in such a way that they adjust towards and eventually ‘lock in’ to a common phase and/or periodicity. There are two basic components involved in all instances of entrainment:

1) **There must be two or more autonomous rhythmic processes or oscillators.** Autonomy means that if the two oscillators are separated, i.e. if they do not interact, they must still be able to oscillate: the oscillations must be active processes requiring an internal source of energy. In other words, none of the oscillations (or the rhythmic processes) should be *caused* through the interaction. Resonance, for example, is not to be considered entrainment: if a tuning fork producing sound waves in a resonance box is removed, the oscillations in the box also cease. This is an important point, because it alerts us to the possibility that the mere observation of synchronized behaviour or synchronous variation in two variables does not necessarily imply entrainment.

2) **The oscillators must interact.** There is a variety of different forms of possible interactions or ‘coupling’, but in the majority of cases this interaction is weak, as demonstrated by Huygens’ clock example. Strong coupling, on the other hand, puts too strong a limitation on the oscillators, and they lose their ‘individuality’ (for example, consider how one’s arm movements would be limited if both hands held on to the same stick). As will become clear from the following discussion, identification and quantified description of the ‘weak interactions’ is one of the biggest challenges in entrainment research.

The tendency for rhythmic processes or oscillations to adjust in order to match other rhythms has been described in a wide variety of systems and over a wide range of time scales (i.e. periodicities): from fireflies illuminating in synchrony, through human individuals adjusting their speech rhythms to match each other in conversation, to sleep-wake cycles synchronizing to the 24-hour cycle of light and dark. Examples have been claimed from the fast frequency oscillations of brain waves to periods extending over many years, and in organisms from the simplest to the most complex as well as in the behaviour of inorganic materials and systems.

Obviously, the wide range of entrainment phenomena is not based on a single physical process. Rather, the concept of entrainment describes a shared tendency of a wide range of physical and biological
systems: namely, the coordination of temporally structured events through interaction. In principle, it is easy to see how entrainment is relevant to music. If an ethnomusicologist were asked to suggest a familiar example of a "temporally structured event", it is likely that music, or some social or ritual event mediated through music, would quickly come to mind. Examples of rhythmic coordination to and through music, such as a foot tapping to the beat of a song, are equally easy to think of. However, entrainment studies in ethnomusicology are not limited to a few apparently obvious examples, and musical applications have to take into account the wider context of entrainment in human biological and social functioning.

Examples of endogenous or naturally occurring rhythms within the human body include the heart beat, blood circulation, respiration, locomotion, eyes blinking, secretion of hormones, female menstrual cycles, and many others. It has been suggested, indeed, that all human movements are inherently rhythmic: Jones writes that “All human performance can be evaluated within a rhythmic framework” (1976:340), while Bernieri, Reznick and Rosenthal suggest that “…human behavior is understood to occur rhythmically and therefore can be described in terms of cycles, periods, frequencies, and amplitudes.” (1988:244). As we will see, these endogenous rhythmic processes may interact in many different ways, either within a single person or, in some cases, between individuals: both cases are relevant to ethnomusicology, although much of the material in this paper concerns interpersonal or social entrainment. Entrainment to and through music needs to be seen as a particular case of entrainment in social interaction, and its particular qualities explored - as indeed, we need to explore the specific possibilities for entrainment that different musical repertories or performances afford.

2.2 A short history of the entrainment concept

Huygens identified the phenomenon we know as entrainment as a result of his invention of the pendulum clock: he noticed that two such clocks, when placed on a common support, would synchronise with each other - even when the pendulum of one was deliberately disturbed, they would regain perfect synchrony within half an hour or so. Huygens’ description of, as he put it, ‘the sympathy of the clocks’ as well as the apt explanation he supplied, were considered a singular phenomenon for more than 200 years. Then, in the second half of the nineteenth century British physicist Lord Rayleigh described the synchronization of two similar but slightly different organ pipes and introduced the distinction between forced and maintained oscillations (Rayleigh, 1945). At the time these findings were not thought to be connected, but things took a different shape following Poincaré’s development of a new approach for dealing with complex, non-linear systems at the end of the nineteenth century. Following the advances made by Poincaré, it became possible to describe and understand the observations of Huygens and Rayleigh in terms of interacting non-linear systems. (In a linear system, changes in one variable produce predictable changes in a dependent variable. By contrast, in a non-linear system, small changes in one variable may cause large, erratic changes in a dependent variable.)

For more than 250 years after the time of Galileo and Newton, the sciences made considerable advances and it looked as if one day they might enable us to fully understand the world. This impression arose from the fact that classical physics had developed a method to analyse and describe the physical world

---

2 The first being ‘forced’ upon a system from outside, the latter being maintained autonomously by a system.

3 Instead of asking quantitative questions that could not be answered at his time, Poincaré approached complex systems by emphasizing qualitative questions and developing a geometric analysis of complex system behaviour.
in a simple way, namely in terms of sets of linear equations. This meant, given the initial conditions of a
system, we could predict its behaviour and deduce any of its future states – we could ‘understand’ it.
The only problem with this approach was that there seemed to exist a number of systems – Huygens’
pendulum being one of them - that could not simply be described by linear equations. The remedy taken
by classical physics was linearization: most non-linear systems can be described approximately by
linear equations if one only considers a limited range of their behaviour. (For instance, the behaviour of
a pendulum can be described quite well in linear terms if one only considers small movement
amplitudes.) The differences between the actual behaviour and the linear descriptions were thought to
be negligible. The linear approach was also successful because it allowed any complex problem to be
broken down into parts, each part to be solved and then the part solutions put together to obtain the
solution for the complex problem. The behaviour of the whole system could thus be described as the
sum of the behaviour of its parts. However, many systems in our world do not act that way. Mostly, when
parts of a system interfere, cooperate or compete, they produce non-linear interactions. At the beginning
of the last century it became obvious that there were not only considerable differences in the behaviour
of linear and non-linear systems – for example, the latter are not predictable – but that a severely biased
world view was implied here, that of a predictable linear world. It was now realized that in our actual
world, complex non-linear systems with chaotic behaviour are the rule, not the exception.

Following the new orientation in complex system dynamics introduced by Poincaré, mathematics and
physics in the first half of the last century became mainly concerned with non-linear oscillators and their
applications (such as in radio, radar, and later, lasers). In the 1920s, important theoretical groundwork
was laid by Appelton and van der Pol when they were able to show how the frequency of an oscillator
can be entrained or synchronized by a weak signal of a slightly different frequency. Finally, the invention
of computers also meant a considerable push for non-linear system theory. From the 1960s, the
computer allowed researchers to experiment with non-linear equations (e.g. to perform simulations) in
a way that was unthinkable before, thereby offering new insights into the behaviour of non-linear
systems.

In the 1970s the development in mathematics and physics of chaos theory had considerable impact on
several other disciplines such as biology (early studies concerned chaotic phenomena in biological
populations: for a review see May 1976), social sciences (Jantsch 1975; Allen et al. 1977) and
neurosciences. The field of synergetics, the theory that is concerned with how the cooperation of
individual components of a complex system leads to the formation of new macroscopic spatial, temporal
and functional structures, considerably informed our present understanding of the functioning of the
human brain (Haken 1983a, 1983b).

At least two aspects of the new view on the brain seem to be relevant in the context of this presentation:
a) most brain functions can best be described as cooperative, synchronized activity of large, distributed
ensembles of neurons, and b) a large part of this synchronized activity is of an oscillatory nature (Basar
1983; Nunez et al 1993). These autorhythmic oscillatory properties of neurons in the central nervous
system are a consequence of their electrochemical properties. The cooperative and oscillatory activities
of these neurons can be seen, amongst others, as the basis for the timing of sensory-motor coordination
(for a review: Llinas 1988). With these new views on the functioning of the brain, it seems most
promising to apply the concept of entrainment to the analysis of human interactions at the interpersonal
and social level as well. Indeed, such applications have been pursued in the social sciences in parallel
with the development of neuroscientific approaches, the mathematics of coupled oscillators, and
numerous other related strands of research. Many of the aspects most relevant for ethnomusicology are
discussed in this publication.
2.3 Physiological rhythms

Although Huygens (see above) worked on mechanical rather than biological systems, and much of the succeeding work on entrainment - both theoretical and applied - has been carried out within the fields of mathematics and physics, our main emphasis is on the entrainment of physiological rhythms, especially in human beings. One of the first applications of the entrainment model in biology is attributed to the French physicist de Mairan, who in 1729 first identified and studied photic entrainment - synchronisation to the cycle of light and dark - in plants (Chapple 1970). Apart from entrainment to such environmental cues, entrainment studies on various animal species have identified numerous examples of mutual entrainment between individuals, including crickets chirping in unison and fireflies flashing synchronously (Ancona and Chong 1996; Strogatz and Stewart 1993). Thus, although the capacity to entrain may be exploited in particular ways by the higher animals, and particularly by man, it seems to be evidenced in some way or other by all animal and plant species, and the ubiquity of the phenomenon points towards its importance in the world of living organisms.

In modern times, a lot of research has concentrated upon studying the nature of rhythmic processes in living organisms. The proponents of chronobiology, as it is called, point to the fact that biological or physiological rhythms appear to be essential to life itself (Aschoff 1979, Strogatz and Stewart 1993, Glass 1996). Menaker reports that cyanobacteria, simple organisms that originated at least 3 billion years ago, "have fully functional circadian clocks", which may give support to the suggestion that biological rhythms and their entrainment are fundamental to life in any form (2002:2). Some have gone so far as to characterize any organism "as a (loosely coupled) 'population of oscillators'" (Pittendrigh 1975, quoted by Warner 1988:68-9). A good deal of current medical research is concerned with the behaviour of endogenous physiological rhythms in humans (such as the variation of body temperature over the 24-hour cycle), and how the study of those rhythms might be further developed as a tool in the diagnosis of pathological states and ultimately lead to the development of new treatments (see Glass 1996). An important part of this work is the consideration of entrainment and in particular, identifying which physiological rhythms entrain to which stimuli, and under what conditions.

These studies, needless to say, reveal that the entrainment of biological rhythms is a complex subject. As we have seen, in order to model the behaviour of these systems mathematically we need to consider the entrainment of oscillators whose interactions are non-linear and potentially chaotic. That is, even if the relationship between two biological rhythms is taken to be ruled by a relatively simple mathematical equation (telling us what the outcome ought to be, given a particular configuration of endogenous rhythm, entraining rhythm, and intensity of interaction), since that mathematical equation is non-linear the output behaviour of the system can be unpredictable, or at least, predictable only when those conditions fall within a fairly narrow band. An illustration of this would be that the behaviour of our endogenous cardiac rhythm ought to be predictable when stimulated by a pacemaker, but if the period of the pacemaker were set outside a certain range, the behaviour of the cardiac rhythm would be for practical purposes unpredictable.

---

4 Beyond the introductory comments in section 1, the present article will not address applications in mathematics and physics, nor will it deal in detail with the mathematical modelling of entrainment in any detail, although we will need to make brief reference to some of the implications of the mathematics. The interested reader may want to consult the respective literature (e.g. Kelso, 1995; Pikovsky et al., 2001; Strogatz, 2000).

5 In what may have been the first scientific observation of this kind, the Dutch physician Engelbert Kaempfer, following his journey to Siam in 1680, described the synchronization of light emission in large populations of glowworms (Buck and Buck 1968).
The relationship between entrainment, the stability of biological rhythms and health is not easily summarised. There are many examples where relatively stable and entrained biological rhythms are associated with good health (the enhanced stability of the heart rate afforded by a pacemaker is an example of this), while conversely asynchrony and instability of rhythmic processes can be associated with pathologies. However, the situation overall is not quite so simple. Entrainment does not necessarily imply stability of biological rhythms, and in any case stability is not necessarily associated with good health. In the case of brain waves, for example, we have a different pairing: stable brain waves may indicate a pathology (epilepsy) while unstable waves may indicate a healthy state. Indeed, a certain amount of flexibility and dynamic equilibrium seems to be associated with health in many systems, as is a degree of "noise", or random variation in normal physiological rhythms (Glass 1996: 281). One might suppose a connection here with the well-known 'humanising function' in music sequencing software, where a random element is introduced: this 'noisy' behaviour is perceived as 'natural', whereas clean behaviour is perceived as 'artificial'. Similarly, there is a link here with Keil’s theory of 'participatory discrepancies,' which will be discussed below (see section 5.1).

Although we should be careful not to oversimplify, it does seem that there are healthier and less healthy, or pathological, ways in which to be entrained. If the healthy functioning of a system requires a certain degree of entrainment, then either a lack of entrainment, a weakening or even an excessive strengthening of entrainment could be associated with a change to a pathological state. Pathological states involving a disruption of 'normal' entrainment within or between individuals include conditions such as epilepsy, Parkinson’s, and autism. Autism, to take one particularly relevant example, is clinically defined in terms of "impairments affecting social interaction, communication and creativity": it can be associated with motor incoordination and co-occurs with abnormal circadian rhythms and "probably universally - a very poor intuitive sense of time" (Boucher 2001:111). Boucher suggests that people with autism suffer from deficits in their timing systems, a finding supported by Condon’s studies of the coordination of social interactions (e.g. Condon 1982; see below, section 3.2). Not surprisingly, then, music therapists have developed the use of rhythmic entrainment to successfully treat clients with autism (see below, section 4.3).

2.4 Circadian rhythms and entrainment to environmental stimuli

An important part of chronobiology, since de Mairan, is the study of entrainment to environmental stimuli. This can be characterized as asymmetrical entrainment, in that the individual cannot influence the entraining rhythm (e.g. the alternation of light and dark). In the case of asymmetrical entrainment, a body is 'forced' to adjust to externally set, cyclically varying conditions without being able to influence the latter. (Where one rhythm appears to be driving another, the former is referred to as the 'entraining rhythm', or Zeitgeber: the adjusting rhythm is sometimes termed the 'entrained' rhythm.)

Circadian rhythms of living organisms are examples of this type of entrainment, demonstrating the adaptive capabilities of living systems (e.g. Wang and Brown 1996; Zeng et al. 1996).

The importance of the circadian rhythm is not limited to the regular alternation of sleep and wakefulness: this same rhythm entrains a host of other physiological rhythms such as the daily temperature cycle and the endocrinal system. The detail of those systems need not concern us here: more interesting perhaps are attempts to model the entrainment to light cycles, which might have parallels with entrainment to auditory cues (see for instance the summary in Groos 1983, Moore 1990).

\[\text{In some contexts the more loaded terms 'master' and 'slave' rhythm are used.}\]
The mechanism of this entrainment process has been the subject of considerable research aiming to identify the entraining pathway; to determine whether the brain contains a single pacemaker or several linked oscillators; and to locate the pacemaker or pacemakers. Moore (1990) suggests that synchronisation of circadian rhythms depends on the entrainment of a pacemaker located in a part of the brain known as the suprachiasmic nuclei. The endogenous rhythm of this pacemaker is reset by exposure to light, this information being relayed via the optic nerve. The central pacemaker, in turn, entrains a number of other oscillators that regulate particular physiological systems, and maintains stable phase relations between these oscillators (Rusak 1990). A similar volume of research has yet to be carried out on entrainment to auditory cues, although recent research on motor performance and metrical perception employ models of multiple, linked, oscillators in the brain (see section 4.1) that are similar to models of entrainment to light cycles.

2.5 Ultradian rhythms and interpersonal entrainment

Chronobiologists are turning their attention increasingly to ultradian rhythms, i.e. those with periods of less than a day. The term 'ultradian' covers a wide range of possible periods, from milliseconds up to 12 hours, although it tends to refer particularly to the range from a few minutes up to a few hours: faster rhythms are sometimes referred to as supra-ultradian. There are of course numerous examples of ultradian biological rhythms, so it is not possible here to do more than cite a couple of examples and draw out some features of ultradian rhythm research that are of interest to ethnomusicologists. One of the important findings of chronobiology was the discovery of cycles of rapid eye movement (REM) in sleep, with a period in the order of 90 minutes. Slightly more controversial is the hypothesis that this 90-minute cycle continues throughout the day, in wakefulness as well as in sleep (Kleitman 1963), a pattern known as the Basic Rest-Activity Cycle.

An important feature of ultradian biological rhythms is that they are never simple, symmetrical patterns (the same is true, of course of the 24-hour patterns of light and dark). Patterns of activity and rest in humans, for instance, are not balanced. As Sing puts it, "more time is spent in activity than in the rest phase, and its exact period from one day to the next is very variable. Hence, predictability of exact circadian periodicity is not possible - one can only give probable periods between a narrow range of values." (Sing 1992:353).

An important difference may also be observed between circadian and ultradian rhythms: the former, which entrain to the environmental light-dark cycle, are the same for all species, whereas for many ultradian rhythms, period length varies with body mass (an illustration of this is that the heartbeat of a mouse is approximately 20 times faster than that of an elephant, and its maximal life span is roughly 20 times shorter; Gerkema 2002). Gerkema suggests, therefore, that circadian rhythms could be said to reflect "'objective', environmental time", while internal ultradian rhythms "reflect subjective, physiological time" (2002:212, citing Brody). Of course if one takes into account social interactions, different individuals do mutually entrain their "subjective, physiological rhythms", so the distinction is perhaps more appropriately expressed as that between different modes of entrainment than between the objective and the subjective. The two modes would be (a) asymmetrical entrainment to environmental cues at the circadian level, and (b) symmetrical entrainment to other individuals at the ultradian level. (As Gerkema goes on to point out, mutual entrainment or social synchrony is reported in a wide variety of species including "amoebas, the barnacle gees, voles, cattle and rhesus monkeys" (2002: 212, references omitted).) These observations nonetheless perhaps support the idea that entrainment may
relate phenomenologically to a sense of social belonging, or of one's subjectivity relating to "something larger": impressions that are frequently linked to musicking, among other activities.

2.6 Self-entrainment

Not all entrainment involves an external stimulus, either environmental or inter-personal. 'Self-entrainment' describes the case where two or more of the body's oscillatory systems, such as respiration and heart rhythm patterns, become synchronized. Port et al. (1996) have pointed out that humans and animals typically exhibit self-entrainment in their physical activity. By self-entrainment they mean that in actions by a complex body, a gesture by one part of the body tends to entrain gestures by other parts of the body. For example, arm movements in walking could - in principle - be independent from leg movements, but in fact they are not. It 'feels' much easier, is more harmonious, and less strenuous if the arms lock into the leg movements. A similar effect is reported for the locking of step and inhalation cycles in jogging (Bramble and Carrier 1983), or between respiration and heartbeat in high performance swimmers (Schäfer et al, cited by Glass 1996:280). The degree and kind of self-entrainment exhibited depends on the individual and the task being carried out. For some of the cases just mentioned, it could be argued that the coupling is due largely to a direct physical link between physical oscillators: legs and arms are mechanically connected via the trunk. However, entrainment does not imply a rigid mechanical coupling between oscillators. On the contrary, with rigid mechanical coupling, entrainment will be lost because the two oscillators lose their independence and form a new, unified system.

It has been a longstanding discussion whether or not multiple motor actions in humans are governed by a central clock. Recent research, however, seems to be in favour of a multiple timer theory in which each motor action is controlled by an independent timer (Ivry and Richardson 2002). The multiple timer model that Ivry and Richardson developed is a form of coupled oscillator model. The coordination of motor actions appears to be ensured by a neural gating mechanism that, at the same time, improves temporal stability of the actions involved. Ivry and colleagues have shown this for bimanual coordination as well as for experimental tasks involving repetitive movements of two hands and one foot (this is not unlike keeping time by clapping while tapping a foot). The difficulty people have in temporally uncoupling their limbs suggests that the gating process may reflect a fundamental constraint in human performance. At present, however, this constraint has only been identified in coordinated limb movements. It is not clear whether the same or a similar gating mechanism exists for all types of motor actions, for example between limb movement and speech acts. In any case, the fact that such gating

---

7 To borrow a verb from Small 1998.

8 Another related application of entrainment theory is Moore-Ede, Sulzman and Fuller's distinction of two clusters of ultradian rhythms, "System X and System Y" (discussed in McGrath and Kelly 1986: 82-83). Their System X describes a collection of oscillatory processes organized around core body temperature and REM sleep; System Y describes a second collection of oscillators, organized around the rest-activity cycle, and including skin temperature and slow-wave sleep (82). System Y seems to be located in the suprachiasmatic nuclei (see above), and is usually entrained to environmental circadian rhythms, while System X is located elsewhere and not so entrained. The different rhythms in each system seem to be strongly mutually entrained, while the two clusters entrain each other rather more loosely as well as asymmetrically, with System X influencing System Y more strongly than the converse.

9 Neural information for the motor actions passes through a 'gate' that opens only when simultaneously activated by the different oscillators involved.
mechanisms would hardly be able to constrain entrainment between actions located in different bodies justifies us in distinguishing between self-entrainment (within one body) and interpersonal entrainment. Self-entrainment in speaking has been studied by Port et al. (1996). They report the results of some experiments in which they test the relation between onsets of repeated phrases and the timing of stressed syllables within these phrases under various repetition rates. They suggest that ordinary human speech exhibits self-entrainment more or less whenever given the opportunity to do so, and that self-entrainment may be deeply revealing about the way in which time is handled in the nervous system of animals, for many purposes going well beyond the coordination of limbs. Below, in case study 2, we report the analysis of two cases of musical self-entrainment where a musician simultaneously sings and plays a rhythm accompaniment.

The motor system is not only responsible for producing a rhythm, but is also involved in the perception of rhythm: this allows us to understand in part why we experience a visceral response to rhythm. While not yet clearly understood, research is ongoing to determine "how the gestural experience of producing a sound in a physical world interacts with its perception" (Risset & Wessel 1999:142 fn 6, citing Cadoz, Lisowski and Florens 1984, 1993). Keil has referred to this phenomenon as 'kinaesthetic listening,' where music listeners experienced in performing music "[feel] the melodies in their muscles, [and imagine] what it might be like to play what they are hearing" (Keil 1995:10). Keil's comments have been corroborated with numerous examples (among them Blacking 1983:57, Sager 2002:198, Sundberg 1999:210). While this phenomenon has broad implications for ethnomusicological research and interpretation, its importance here is to emphasize the significance of entrainment in listening to music, and to suggest that the study of entrainment in listeners is as important as that in performers. (Needless to say, as ethnomusicologists we are aware that the distinction between “listener” and “participant” is not always easy to make in any case!)

2.7 Entrainment and brain waves

The term entrainment, in the specific sense of frequency-following, is also widely used in connection with brain waves. The technique of recording electrical activity of the human brain from the scalp originated in 1875 with observations of Richard Caton and was developed into electroencephalography by Hans Berger in the late 1920s. Berger found that the recorded activity can be described in terms of four frequency bands: delta (1-4 Hz), theta (4-8 Hz), alpha (8-12 Hz), and beta (14-30 Hz) waves.\(^{10}\)

---

\(^{10}\) The electrical activity that is recorded in electroencephalograms (EEGs) is largely attributable to postsynaptic potentials (PSPs; i.e. graded potentials produced by synaptic activities, that eventually lead to firing of neurons) in cell bodies and dendrites of cortical neurons (Lopes da Silva and Storm van Leeuwan 1978). Neurons of the human brain, the 'gray matter', come in two principal arrangements: layered they form a cortex and in non-layered agglomerations they form a nucleus. Two cortices, the cerebral and the cerebellar cortex form the surface layer of the human brain; nuclei are located beneath the cortex and in the brain stem. The columnar arrangement of neurons in the cerebral cortex facilitates summation of these potentials and their registration at the scalp. However, other geometric arrangements of neuronal assemblies can lead to extracellular attenuation or even cancellation and therefore not all activities of brain cells can be recorded in the EEG. The regular spontaneous EEG components are thought to be due to PSPs synchronized by discharges from deep nuclei (thalamus) and the degree of synchronicity is reflected in the amplitude and form of the EEG (Lopes da Silva 1991). If cortical activity is synchronous over a larger area it produces larger potentials (e.g. Cooper et al. 1965). Desynchronisation of the EEG and reduction of its amplitudes presumably reflects increased interaction of several neuronal sub-populations engaging in cooperative activities.
Despite persisting uncertainties about the actual neuronal processes that generate EEG (i.e., electroencephalogram) waves, and despite the fact that EEGs are incomplete records of neuronal activities in the brain, it has been obvious since Berger's days that EEGs reflect certain mental states. A dominance of low-amplitude beta waves (14-30 Hz) was observed in busy and alert states, whereas alpha waves (8-12 Hz) with larger amplitudes dominate in a relaxed, inattentive state. In the early days of EEG research it was also discovered that some of the alpha and beta waves could be synchronized — entrained — to the frequency of an external, bright strobe light stimulus. The English neurosurgeon Gray Walter was the first to report that at certain 'entrainment' frequencies of the external stimulus, his subjects would enter trance-like states where they began to experience deep peacefulness, dream-like visions, and other unexpected sensations (Walter 1953). Later it was discovered that not only strobe lights but also rhythmic noises could produce such effects. Although this phenomenon is still not fully understood, it does indicate that external stimuli are able to affect brain states.

The problems surrounding the interpretation of spontaneous EEGs probably explain why those few studies concerned with the relationship between external stimuli and brain waves often end up with general findings, for instance that pleasurable stimuli tend to increase the theta wave components (e.g. Walter 1953; Maulsby 1971; Walker 1977; Ramos and Corsi-Cabrera 1989). So, while studies of brain wave patterns via EEGs present some enticing research opportunities, the limitations mean that this research method has so far only provided rather general conclusions about the function of musical phenomena (as sources of external entraining frequencies) upon changes in mental states.

2.8 Degrees and phases of entrainment

Where two or more rhythmic processes or oscillators have the potential to interact, they do not synchronize automatically and instantaneously, and in some cases do not synchronize at all. Some of the factors determining whether two oscillators will entrain, and the different entrainment possibilities, are the following:

First, not all oscillators will entrain: it is generally observed that, for instance, the periodicities of autonomous oscillators need to be fairly close to each other for entrainment to happen (see Aschoff 1979: 6).

Second, oscillators may entrain more or less strongly, and more or less quickly (partially-entrained or transient rhythms are exhibited during the entrainment process: see Chapple 1970). The term entrainment does not only include cases of strict phase and frequency synchronization, and may cover a spectrum from weak to strong coupling. A rhythmic process may adjust towards the frequency or phase of another rhythmic process without attaining absolute synchronisation.

Third, we can distinguish two aspects of entrainment that need not necessarily co-occur. One is frequency or tempo entrainment, where the periods of the two oscillators adjust toward a consistent and systematic relationship (see Ancona and Chong 1996). The other is phase entrainment, or phase-locking: where two processes are phase-locked, focal points (such as a foot striking the floor when dancing) occur at the same moment (this is discussed in greater detail section 3.3 below).

Fourth, the phenomenon of entrainment is further complicated by the fact that oscillators may entrain in states other than exact synchrony. For instance, two entrained oscillators have two possible phase-locked states, namely synchrony and anti-synchrony: a normal human gait exhibits the latter, i.e. one foot comes up as the other goes down. The more oscillators are involved, the more phase-locked states are possible. These possibilities can be derived mathematically using the theory of coupled oscillators, and can be observed in natural phenomena: for instance, quadruped gaits closely resemble the natural
patterns of four-oscillator systems. Stewart and others deduce from this fact that "The most likely source of this concordance between nature and mathematics is in the architecture of the circuits in the nervous system that control locomotion" (Strogatz and Stewart 1996:73).

As Bluedorn puts it, then,

```
Entrainment is the process in which the rhythms displayed by two or more phenomena become synchronized, with one of the rhythms often being more powerful or dominant and capturing the rhythm of the other. This does not mean, however, that the rhythmic patterns will coincide or overlap exactly; instead, it means the patterns will maintain a consistent relationship with each other. (2002:149)
```

Such a “consistent relationship” could be, as we have seen, one of synchrony or anti-synchrony, or of some other definite relationship (e.g. one oscillator runs twice as fast as another). Moreover in real-life situations, it is often the case that two periodic processes lock frequency, but remain out of phase. The relationship between two oscillators can therefore be described as either lagging, synchronous, or leading. Where one oscillator either lags or leads the other, the difference in phase can be expressed in terms of phase angle. The period of oscillation is represented by the 360 degrees of a circle (this terminology is based on the representation of a periodic rhythm by a cyclical motion mapped onto a circular shape). Where two oscillations are in precise anti-synchrony the phase-angle difference is 180 degrees; if one is 1/4 out of phase with the other (a musical example would be when one musician is one beat out of phase with another in a four beat metre) the phase-angle difference is 90 degrees; and so on. A rhythm that ‘follows’ another is described as “lagging” and having a negative phase-angle difference; one which anticipates is called “leading” and has a positive phase-angle difference. Some entrainment studies assume that entrainment necessarily involves synchronisation of phase, but in real life systems this is not always the case. This concept and the terminology of phase difference can be useful in studies of musical entrainment, as will be demonstrated in case studies 2 and 3 below.

3. Further applications of the entrainment concept

3.1 Social psychology

Applications of the entrainment model also extend into a number of areas of social scientific research, including social psychology, linguistics and communication studies, and cognitive psychology. Social psychologists routinely employ the term entrainment, describing phenomena such as the synchronization of an individual’s activity cycles to his or her working hours. Some of the findings in this area are of interest to ethnomusicology, such as McGrath and Kelly’s application of entrainment to social psychology – referred to as the “social entrainment model” (McGrath and Kelly 1986: 83-84; see also Kelly 1988, Ancona and Chong 1996, and Warner 1988).

The social entrainment model rests on five major propositions:

1. That much of human behaviour, at physiological, psychological, and interpersonal levels, is temporal in character; that is, that the behaviour is regulated by processes that are cyclical, oscillatory, or rhythmical.

2. That these are endogenous rhythms; that is, that they are inherent in the life processes of the organisms involved.

3. That sets of such rhythms, within the individual, become mutually entrained to each other, hence that they come to act in synchrony in both phase and frequency or periodicity.
4. That the temporal patterns of individuals who are in interaction become mutually entrained to one another, that is, that they get in synchrony of phase and period.

5. That temporal patterns of behaviour of individuals and sets of individuals become collectively entrained to certain powerful external pacer events and entraining cycles. The former alter the phase or onset, and the latter alter the periodicity, of those endogenous and mutually entrained rhythms.

(McGrath and Kelly 1986: 83-84)

This model is useful in as much as it sets a precedent for bridging the gap between neurological and biological studies on the one hand, and social scientific research (in this case, social psychology) on the other. Their conclusion here is that:

During a period of social interaction, the members of [a] social system must work out a "negotiated temporal order" in which they adjust their activity patterns to coordinate with each other. Each member of the social system can be viewed as an oscillator (or as a set of loosely coupled and mutually entrained oscillators); that is, the person’s activity will show one or more patterns of alternation over time… The multiple independent cycles of activity of the members of a social system become coordinated with one another into a temporally patterned system of activity that is characterized by a dynamic equilibrium rather than by a fixed homeostatic pattern. (1986: 89-90)

We reflect below on some of the ways in which these ideas may be applied to ethnomusicology. For now it is enough to note that this is a precedent for regarding entrainment not only in physiological, but also in social terms. When we entrain to a piece of music this synchronization can also be analysed in terms of its social conditions and implications (whether relatively symmetrical, as between two or more performers, or asymmetrical, where a group responds to amplified recorded music). The degree of entrainment, and the negotiation of symmetrical entrainment between individuals, can be studied in terms of a negotiation of relative power between those individuals (i.e. if there is a significant imbalance of power or authority, the less powerful individual(s) may adjust their endogenous rhythms further and more readily than do the more powerful). An illustration of this will be found in case study 3.

3.2 Interaction and communication

Arguably the most significant areas of entrainment research for ethnomusicologists are those concerning interaction and communication between human individuals: this includes verbal and nonverbal communication and the relation between the two (roughly speaking, language and gesture). Considerable evidence from social scientific studies points to the rhythmic organisation of both verbal and gestural communication, to the mutual entrainment of speech and gesture, and to entrainment between the communicative rhythms of interacting individuals (Davis 1982; Montagu and Matson 1979). Research in the biological sciences also demonstrates the existence of inter-personal entrainment. For instance Schmidt et al. (1990) report that subjects tend to synchronize their leg movement with that of other subjects observed visually. This makes entrainment a promising concept for the study of musical performances, where exchange and communication takes place over some physical distance, via visual and auditory channels, not via direct mechanical ‘coupling’.

An early contribution to this strand of research was the work of Lenneberg, who made a case for rhythm as the “essential organizing principle” of language, and developed analyses which treated the rhythmic patterns of speech as a kind of “carrier signal” for what is normally considered linguistic meaning, analogous to the carrier signal of a radio wave (Lenneberg 1967). Subsequent researchers have identified periodicities in communicative rhythms on a number of different levels, from the alternation
of speech and silence in conversation down to the alternation between consonants and vowels in each phrase (see Jaffe and Anderson 1979).

An interesting finding in the study of speech rhythm that might suggest entrainment, is that in certain situations the periodic rhythms of speech continue across turn boundaries - in other words, a speaker may conform to precisely the speech rhythms of the preceding speaker (e.g. Couper-Kuhlen 1993, Auer et al 1999, Webb 1972). This is not the case for all communicative interaction, which raises the question under what circumstances this apparent synchrony will manifest, and what its presence or absence (or degree) indicates about interpersonal relationships.

A productive line of research concerns the coordination of these speech rhythms with other motor movements. Indeed, among the main interests in studies of verbal and nonverbal communication has been the relationship between speech-related and speech-independent gesture, the ways in which gestural communication may be like (or unlike) linguistic communication, how gesture is co-ordinated with speech (for an accessible textbook introduction to this subject, see Knapp and Hall 1997), or how gestures may even be at the origin of speech (McNeillage 1998).

Speech and gestures are apparently strongly coupled in adults. Coupling is so effective that when speakers stutter, the gestures tend to stop until the speech is recovered (Mayberry et al., 1998). In analysing the ontogenetic development of this interaction, Iverson and Thelen (1999) invoke the concept of coupled oscillators. They propose that hand and mouth activity are loosely coupled from birth, and this initial linkage (e.g. the Babkin reflex) seems to be established phylogenetically through the feeding system. Through rhythmic activity, and later in the life-cycle, through gesture, the arms gradually entrain the activity of the vocal apparatus.11 Mutual activation increases as vocal communication through words and phrases becomes more practised, leading to strong synchronization of speech and manual gestures for communicative purposes in adults.

Some, but by no means all, research on gestural communication has addressed the question of interaction - examining for instance, how an empathetic listener mimics a speaker’s posture and gesture patterns. A subset of these interaction studies directly addresses the question of interactional synchrony. Because the study of interactional rhythms is of the greatest significance for ethnomusicological approaches to entrainment, we will consider here some of the key researchers and their methods and theoretical orientations here.

The first major figure to carry out empirical timing studies of interaction was the American anthropologist Eliot D. Chapple, beginning in the 1930s (Chapple 1939, 1970). His initial studies were based on simple timings of "action" and "inaction" of participants in interaction (for instance, the durations participants in a conversation spent speaking and listening). On the basis of these studies, Chapple became convinced that normal social interactions were rhythmically - that is, periodically - organized. His book Culture and biological man, published in 1970, makes it clear that he saw these rhythmic processes as located on a continuum of naturally occurring rhythms from the very fast (brain waves, muscle fibre firing) to those with much longer periods (menstrual cycles, migration patterns) – i.e., from the ultradian to the infradian. He also discusses the synchronisation of these rhythms explicitly in terms of entrainment (1970:27).

The next generation of researchers was able to develop more sophisticated methodologies for the study of interactional rhythms and synchrony, in particular employing the analysis of sound film. The first to apply film analysis to the study of interactional synchrony in any detail was William Condon, who carried

---

11 Entrainment of vocal activity by manual activity is also suggested in some ethnomusicological studies (see discussion of case study 2 and Will, in press).
out a number of detailed cinematographic studies of interaction in the 1960s using methods developed by Ray Birdwhistell. 12 This "microanalysis", as he called it, yielded a number of important results: Condon reported that "microanalysis of sound films of listener behaviour led to the surprising and unsuspected observation that listeners move in precise synchrony with the articulatory structure of the speaker's speech" (Condon 1976:305), which he interpreted as evidence of "entrainment or stimulus tracking" (309). Although in some respects Condon's methodology raised suspicions of researcher bias (his method for dealing with the key problem of the segmentation of behaviour "constituted re-viewing a sound film of human communication over and over for many hours until forms of order began to be seen" [1976:288]), his findings that this synchrony could not be observed in subjects "with severe psychopathology or communication disorders" such as autism, dyslexia and schizophrenia (1982:61), and that the detailed features of the interactional synchrony varied with the ethnic origin of his subjects, do suggest that his work should not be casually dismissed. At one point, Condon even suggested that the periodicities he observed in behaviour were uncannily close to the frequency ranges of the different types of brain waves, and in light of this he coined the term "behavioural waves" (1986: 63-66). 18 Following in the footsteps of Chapple and Condon, among the most prominent researchers of the latter part of the 20th century (and into the 21st) were Adam Kendon and David McNeill. Kendon's studies concentrate on the coordination of speech with body movement, again employing sound film analysis (e.g. Kendon 1972, 1981, 1982). David McNeill has developed this approach further, working up a detailed theoretical model of the interrelationship between speech and gesture (concentrating in fact on a particular category of gesture, alternatively labelled "gesticulation"). McNeill argues that gesture and speech "arise from a single process of utterance formation. The utterance has both an imagistic side and a linguistic side" (1991:29). His approach thus differs from that of Birdwhistell (see footnote 12): where Birdwhistell looked for language-like structures in gesture, McNeill sees gesture as an imagistic complement of language. This view is apparently influenced by Kendon's distinction between digital and analogical encoding: "Words and other discrete symbol systems… convey their messages digitally. Actions, whether practical or expressive, convey their messages analogically" (Kendon 1981: 5). 19 This

12 Ray Birdwhistell pioneered sound film analyses of interaction: his particular approach is referred to as 'kinesics', and amounts to a quasi-linguistic study of nonverbal communication. He was convinced that there was no distinction between verbal and nonverbal communication - he refused to use the latter term - and worked to analyse gestural communication in the terms of structural linguistics. Thus, he attempted to define the basic units of gestural communication as 'kinemes' (analogous, in his view, to phonemes), and to show how they were built up in successively larger units of meaning. Birdwhistell's approach is unusual, since studies of verbal communication are more often seen as contributing to a critique of the Saussurian structural tradition, which saw linguistics as the study of the structural relationships between units of meaning, working up from the phoneme (Gumperz 1996).

13 It is also noteworthy that Condon is responsible for one of the few direct cases where the entrainment model has influenced ethnomusicological thinking - through his contact with Alan Lomax.

14 The other major figure in this area has been Paul Ekman - however, his work seems to be less directly relevant for entrainment studies. Ekman took a slightly different line, concentrating more on the relationship between nonverbal behaviour and feeling states (Weitz 1974). Ekman considered gestural communication as a rather disparate phenomenon and worked on the taxonomy of gesture types. He proposed four major categories of gesture: (i) emblems (movements tied to specific verbal meanings), (ii) body manipulators (head scratching, nose picking etc), (iii) illustrators (tied to the content or flow of speech, and (iv) emotional expressions (Ekman 1977: for a more extensive discussion see Ekman and Friesen 1969). The distinctions between different types and functions of gestures may explain some of the apparent differences between the orientations of different researchers discussed in this section. Although we would suggest this work is only indirectly relevant to entrainment studies, it is worth pointing out that Ekman's work is cited in several musicological works: see e.g. Blacking 1977 and Frith 1996: 215-217.
may be a significant observation for future entrainment studies - many musical behaviours may be thought better suited to description in analogical rather than digital terms (for instance, the continuous, flowing melodic figures used in much Indian classical music).

A useful critical discussion of some of these approaches can be found in Rebecca Warner’s article "Rhythm in social interaction" (1988). Warner discusses, with reference to the work of Chapple, Condon and others, concerns of observer bias as well as the possibility of statistical artefacts, proposing what she suggests are more reliable methods for studying synchrony in social interaction. One of her proposals is to treat rhythm in terms of recurrent cycles modelled as sinusoidal waveforms, rather than as alternations of activity and rest (square-tooth waveforms) or as sequences of events occurring at discrete time points. Interestingly, a similar suggestion is made in a specifically musical context by Robert Gjerdingen (1993): this will be discussed in more detail in section 4.1.

We noted above the need to take care in correlating entrainment and health. The same caution applies here in correlating types and degrees of social entrainment with emotional affect. However, there does seem to be some evidence of a correlation between entrainment and positive affect in communication. One study reported that women who are "notoriously skillful conversationalists, are particularly adept at modifying their endogenously generated speech rhythms to match those of their dialogic partners. Furthermore, the degree of rhythmic entrainment to an exogenous speech source is associated with positive social evaluation and interpersonal attraction." (Marcus, Welkowitz, Feldstein and Jaffe quoted by Jaffe and Anderson 1979:18-19). It seems reasonable to speculate that in music, as in verbal interactions, particular modes of entrainment are associated with desirable (or in some cases undesirable) types of emotional response or affect.

As noted above, rigid entrainment and precise synchrony in human rhythmic processes are not necessarily associated with health or positive affect: Warner et al. have shown experimentally that “moderately rhythmic social interactions [were] evaluated most positively” (1987:57, emphasis added). It seems that entrainment that is not too ‘perfect’ generally provides a more positive social experience than either entrainment that is too tightly coordinated, or not coordinated at all. What is considered a desirable tightness of entrainment seems to vary at a cultural level, and this knowledge is also an important aspect of any musical culture. Determining what is considered to be the optimum or preferred degree of entrainment in different cultures could serve as an important diagnostic of cultural style (e.g., see Waterman 1995:93).

3.3 Cognitive psychology

The entrainment concept has also had an important role in the development of theory in parts of the cognitive sciences. We will briefly consider some applications of entrainment in cognitive psychology that are relevant to music research: in particular, theories of attention.

Several cognitive psychologists hold that perception, attention, and expectation are all rhythmic processes subject to entrainment. In other words, even when a person is not speaking or performing music, but is only listening, their perceptions and expectations will be coordinated by their entrainment to what they hear. Entrainment is fundamental then, not just to coordinate with others, but even to

15 One suggestion is that the mathematical manipulation of timing data may generate the appearance of regularities or periodicities even where the data is completely random. Another factor here, which is discussed in more detail below, is that correlations between rhythmic processes do not in themselves prove entrainment.

In time with the music
perceive, react to, and enjoy music. Music, as an external oscillator entraining our internal oscillators, has the potential to affect not only our sense of time but also our sense of being in the world. Moreover, it is clear that people exercise a significant measure of self-control in negotiating musical entrainment. The cognitive psychology literature reviewed below addresses the relation of unconscious processes and individual agency as co-determinants of the entrainment.

The presentation in this section is based largely upon the work of Mari Riess Jones and her co-workers, as published between 1976 and 2002. Some basic assumptions about entrainment that we have already encountered above are quite similarly expressed in these studies. First, Jones assumes that human beings are inherently rhythmical, with “tunable perceptual rhythms” that can entrain to time patterns in the physical world. In other words, she postulates a propensity for an individual’s endogenous rhythms to synchronize with perceived and expected rhythmic processes. Second, entrainment occurs as both period and phase synchronize (this is a more restricted definition than we are proposing; cf sections 2.8, 3.1). Third, entrainment varies in degree (cf. section 2.8).

Another basic premise of Jones’ theory is the assumption that many of the time structures evident in real-world events are patterned in coherent and hierarchical ways (Jones 1976:353). According to Jones, a hierarchical time structure is one in which “the temporal distribution of markers [event onsets and relative stress] reveals nested time levels that are consistently related to one another... by ratio or additive time transformations” (Jones & Boltz 1989: 465). Examples of additive time transformations include gradual changes in velocity, such as musical tempo changes, or changes occurring with a shift in physical momentum, as when a person breaks into a run from a casual walk. Most musical meters are examples of simple ratio transformations between at least two distinct but nested time levels. “One is a referent time level, the beat period, and the other is a higher order period based on a fixed number of beat periods, the measure” (Jones & Boltz 1989: 467, see also Yetsin 1976). Nested time levels in music may extend upward from the smallest subdivisions of a beat, to a beat level, to a measure, to a phrase, to a period, to large order forms.

Perceived rhythms will also be interpreted according to learned knowledge structures (referred to as schemes or schemata in the cognitive sciences). For this reason, even a novel rhythm may be assessed as being like a familiar, already learned rhythm. In all, varying rhythmic contexts, stages of development, as well as physiological and psychological factors will cause people to have “different temporal experiences of the same event,” which means people will have different entrainment experiences even though they may be participating in the same musical performance (Jones & Boltz 1989: 471).

Jones argues that people have an initial bias to entrain to simple, coherent rhythms (Jones & Boltz 1989: 470, Jones 1976:341). For example, we might assume that in speech, people will entrain to the speaker’s “characteristic articulation rate,” or in music, to “the beat period or measure span”. This regular period to which we entrain, known as the “referent time level,” serves as an anchor that locks the listener to the speech or music event. When an entrained rhythm is part of a hierarchical rhythmic context, people can selectively shift or focus their attending energies to different levels of referent

---

16 It is a basic premise of cognitive theory that all experiences and sensations cause learning. This process of learning changes the physical structures of the body’s nervous system (e.g., seventy-five percent of the neo-cortex—which includes the music association areas of the human brain—is unassigned at birth and is assigned according to experience and learning; see e.g., Hodges 1996a and Deutsch 1999). Since the very neurological formation of the brain is shaped by individualized learning, it follows that no two people could have exactly the same physical brain - even if they have shared in the same public, musical discourses. This learning and the resulting physical changes in brain structure are one way individuals come to “embody” or internalize their society’s culture. This rather elegantly explains how individual music psychology and music culture overlap and are mutually inclusive—albeit imperfectly and always with individual variation.
periods in order to accomplish different goals. In a musical context, such goals might include keeping track of the unfolding formal structures, or conversely, noting fine subtleties in timbre or pitch. Conversely, such shifts of attention are less likely to happen in events of low temporal coherence (Jones & Boltz 1989). An example of an event of relatively low temporal coherence would be unmetered music, sometimes described as ‘free rhythm’. An implication of Jones’s theory could be that listeners to such music are less likely to shift their attention between different temporal levels (beat, measure, hypermeasure, section) than they are when listening to strongly metered music. "Free rhythm" genres are however very diverse, and the category may include several different forms of temporal organisation (Clayton 1996), so we should not be too quick to jump to conclusions in this case.

Experiments testing entrainment theories in the cognitive sciences have begun to address how the complex patterns of real-world events relate to the cognitive bias toward entraining to simple patterns and hierarchical rhythms. Although speech and music “fail to fit the simplest time hierarchies” (Jones & Boltz 1989: 468), entrainment seems to occur in spite of these complexities. When considering entrainment processes involved in the perception, expectation, or attention to musics with so-called ‘free’ rhythms (e.g., what occurs in the improvisatory, solo introductions of Indian alap or Turkish taksim), the theory of “centering” suggests that even where there is no explicit pulse (and therefore no coherent or hierarchical rhythmic structure), a listener might entrain to what s/he perceives as the ‘centered’ or median period length (see especially Barnes and Jones 2000:290). Barnes and Jones also state that multiple expectancies may be generated by oscillators that entrain independently to different referent levels of an auditory signal; the expectations generated by these different oscillators may either coincide or conflict with each other. Jones and her colleagues’ work shows entrainment to be a flexible process that can adapt and accommodate many ranges of rhythmic complexity and coherence found in the real world. Such a model of entrainment allows for a better understanding of music as a communicative and socially interactive process.

In terms of the mechanics of entrainment in human cognition, Jones theorizes that there are three primary stages: (1) perception, which primes the listener to form expectations; if expectations are met, (2) synchronization; and if expectations are not met, (3) adjustment or assimilation. Perception and the priming of expectations are nearly instantaneous occurrences. Cues from events unfolding around the attender are taken as indicators of where to focus attentional energies in order to ‘catch’ upcoming events. Anticipation of future events is facilitated by the presence of highly coherent (i.e., regularly patterned) temporal events, such as a steady beat (see Jones and Boltz 1989:466). Synchronization follows priming and occurs as our expectations are met. As such, synchronization is itself a verification of the correctness of our expectations. If our expectations do not match what happens next, then synchronization has not occurred. It should be noted, however, that the discrepancies between our expectations and the actual unfolding of events can cause arousal, that in turn heightens attention and results in learning.

The disparities between actual event onsets and the expected onsets first cause a “time estimation response” in which the new time relationship is evaluated. Next, the disparity triggers an adaptive response. Adjustments in expectations and the focus of attention are made in an effort to accommodate the unexpected, new time relationship. Such an adjustment will either be made in terms of where a

---

17 Jones & Boltz (1989: 471) argue that this kind of focal attending “…is a tacit, how-to skill that is implicitly acquired.” Even though they have found that people rarely can verbalize their use of such a rhythm generator, nor can they report which time levels they have focused upon or shifted between, their experiments have shown the presence of this type of attending. In a musical context, such goals might be keeping track of the unfolding structures of a large order form, or conversely, noting the subtleties of the slightest variations in timbre or expressive timing.
phase begins, or in terms of a change in the period length. The phase of an internal oscillator (e.g. an attentional rhythm) is more susceptible to rapid adjustments, whereas changes in period occur relatively slowly (cf. McGrath and Kelly's social entrainment model, summarized in section 3.1). The ease with which an oscillator’s phase adjusts in comparison to the stability of the oscillator's period may explain why many temporary changes in musical timing, such as rubato or breaths taken between phrases, do not disrupt the underlying sense of rhythmic structure or tempo of a piece of music.

Jones’s theories of “attending rhythms” are useful in describing entrainment as an adaptive process, capable of adjusting to widely different musical contexts under the influence both of conscious direction and of the musical stimulus. As a process directed by entrainment, attention can be understood as partly intentional and goal-directed and partly controlled by the pace of an external stimulus (through an involuntary, automatic assimilation of that rate). The complementarity of these facets of the process enables the attender to respond to the environment while also having some control over his/her experience.

Another insight of Jones’ research with implications for the study of musical aesthetics, is the distinction between two different modes of attending: “future-oriented attending” and “analytic attending” (Jones & Boltz 1989, Drake, Jones & Baruch 2000). Future-oriented attending tends to occur where the stimulus has a coherent time structure, which facilitates a shift in attention to higher referent levels (i.e., longer time spans). For this reason, future-oriented attending “supports anticipatory behaviours,” like those we might expect between musicians desiring to play in time with each other (Jones & Boltz 1989: 459). Analytic attending tends to occur when the event stimuli are less coherent and more complex, such as where expectations are extremely difficult to formulate. In analytic attending, attention is switched to focus on shorter time-spans, which facilitates the comprehension of the grouping of adjacent elements rather than repeating structures (see Jones & Boltz 1989:459; Drake, Jones & Baruch 2000:256). Again, this insight may well be a useful conceptual tool in understanding the performance and reception of unmetered music, such as Indian alap: a working hypothesis might be that the temporal structure of the music directs the listener’s attention to an analytical, rather than a future-oriented mode, whereby subtleties of pitch, timbre and/or dynamics may be better appreciated.

In summary, Jones and her colleagues have shown how perceived rhythms set up expectation in the listener based upon the current context (e.g. musical cues) and schemata learned from previous musical experiences. Just as people have perceptual and attentional oscillators that are designed to automatically adapt to match external oscillators, these same perceptual and attentional oscillators may be influenced by our own cognitive capacities. According to the vision of entrainment offered by cognitive psychologists like Jones and her colleagues, entrainment is a form of interactive attending, creating a “synchronous interplay between an attender and an event in which the former comes to partially share the events’ rhythmic pattern” (Jones & Boltz 1989: 470). Entrainment appears, therefore, to be one of the fundamental processes providing an intimate connection between individuals, others, and the world around them.

---

18 According to Barnes and Jones (2000: 294), the period of an oscillator takes six times longer to adapt than does the phase.
4. Applications in music research

Attempts to characterize musical behaviour in terms of entrainment, rare as they may be, do in fact go back over 30 years: the following quotation relates to Chapple’s speculation on topics such as the relationship between possession and entrainment:

Voodoo [sic] drums, the regular and driving rhythms of revivalist ceremonies, the incessant beat of jazz or its teenage variants in rock and roll, must synchronize with the rhythms of muscular activity centred in the brain and nervous system. (1970:38).

Notwithstanding Chapple’s suggestion, given the volume of research on interactional synchrony in verbal and nonverbal communication, the rarity of such studies in music research is remarkable. In this section, we will summarize some of the important work that has treated music as a time-bound, interactive activity and experience.

4.1 Musical metre

Perhaps not surprisingly, much of the work on entrainment in music relates to metre and metrical perception. There are a number of strands of research here, but the most detailed and comprehensive application of the entrainment concept to music research comes via further application of Jones’s theory of attentional periodicity (see section 3.3) to metrical perception. For instance, Large and Kolen (1994) present "a mathematical model of entrainment appropriate for modelling the perception of metric structure." This model comprises "a network of oscillators of various native periods that entrain simultaneously to the periodic components of a rhythmic signal at different time-scales, and to the outputs of one another." (1994: 178, 190). A similar approach is described by Eck, Gasser and Port, who report that a system of coupled oscillators can be modelled in which each oscillator will entrain to one of three pulses in a musical stimulus with ternary metre (2000). In addition, they suggest building a physical robotic arm and linking it to the computer simulation, so that it can literally 'beat time' to a musical stimulus, thus demonstrating experimentally how such behaviour is afforded by the entrainment of coupled oscillators.

Other applications of entrainment to metrical analysis include those of Justin London and Robert Gjerdingen. London employs the entrainment concept in his study of "complex meter" (otherwise known as additive metre), citing Jones’ concept of the referent level (1995). Gjerdingen proposes that if rhythmic perception depends on the entrainment of oscillatory circuits in the brain, then the stimuli used in traditional psychological studies with "clearly demarcated durations, precisely located points of attack, and easily derived temporal ratios" are "among the worst imaginable" (1993:503). Gjerdingen - like the social psychologist Rebecca Warner (see section 3.2) - proposes the modelling of rhythm in terms of sine waves rather than discrete durations, in which he differs from Jones and her co-workers, who use more "traditional" musical stimuli in their experiments (i.e. they used European-derived music with highly coherent, hierarchical time structures). Gjerdingen's and Warner's position may also be supported by the approaches of Paul Fraisse, who stressed the fundamental importance of periodic motion in the psychology of time (see Fraisse 1963 and Clarke 1999). Physiological rhythms are always, however, far less symmetrical, or more 'complex', than sine waves, let alone than the simple alternations of square-toothed waveforms. A somewhat different approach to the same issue is Clynes's work on 'pulse microstructure' (see e.g. 1986).

Entrainment research in music will at some point have to grapple with these complexities, perhaps developing new "wave theories" for musical rhythm. Rudimentary wave theories of metre have been current for about half a century at least in Western musical thought, since the publication of
Zuckerkandl's *Sound and Symbol* in 1956. Entrainment research suggests that music theory may have to revisit these images, and find more systematic ways to describe musical rhythms in terms of the mathematics of wave functions.

### 4.2 Biomusicology

Interestingly, the entrainment concept is invoked in recent evolutionary accounts of music, as well as in studies of proto-musical behaviour in child development. Brown, Merker and Wallin see the ability to entrain bodily movements to an external timekeeper as a key distinction between man and other higher mammals:

> The human ability to keep time should be distinguished from the ability of most animals (including humans) to move in a metric, alternating fashion. What is special about humans is not their capacity to move rhythmically but their ability to *entrain* their movements to an external timekeeper, such as a beating drum. (Brown, Merker and Wallin 2000:12)

Merker argues elsewhere (1999-2000) that this ability, manifested in the development of "synchronous chorusing" in a sub-population of the common ancestors of humans and chimpanzees, may have marked a key phase in the emergence of the human species. If correct, this means that although all animal species exhibit entrainment in various ways, the human ability to entrain to music represents a propensity unique to our species.

In an important publication at the end of the last century, Colwyn Trevarthen sets out a view of the relation between music, gesture and communication centred on the notion of Intrinsic Motive Pulse: an example of research into music taking into account the endogenous generation of musical rhythm and the mutual entrainment of these rhythms. Elsewhere in the same collection of essays, Wittmann and Pöppel suggest that "A neurobiologically determined tempo of all the participants can easily lead to the coordination of the individual clocks of the musicians, thus leading to synchronisation of musical performance" (1999-2000:19) while Malloch discusses mother-infant interaction (another concern of Condon, incidentally) in terms of "attunement" (a synonym for entrainment). Entrainment is thus amongst the concerns of this emerging discipline.

### 4.3 Music therapy

Clinical applications of entrainment theory in music therapy may also be of interest to ethnomusicologists: although the therapeutic context is in many ways distinctive, music therapists' experience can both inform us about the nature of musical entrainment, and suggest perspectives on the relation between entrainment and socialisation. This work has been quite diverse. One therapeutic procedure, for instance, is for the therapist to play a piece of strongly metered music (using tempo in the range 50-65 bpm), with the intention of encouraging entrainment of autistic clients: the limited research data available suggests that this may be effective in modifying clients' behaviour (Orr, Myles and Carlson 1998). More sophisticated procedures include those developed by Nordoff and Robbins.

---

19 See p. 168; for a discussion of this and of the relationship between the metaphors of the wave and the cycle in music theory, see Clayton 2000:ch.2

20 Although Trevarthen uses the term 'entrainment' only intermittently (for instance, in discussing Condon's work), the idea is in fact central to the Intrinsic Motive Pulse thesis.
In time with the music 22

(summarized in Rider and Eagle 1986:231-2), and involve the therapist mimicking the spontaneous musical behaviours of autistic children in synchrony with them. In their experience, once a child realised his behaviour was being mirrored by the therapist, “there was almost universally a laugh, a smile, or some observed affective change which seemed to indicate the children were willing to enter into a more therapeutic relationship.” (230-231). One way to characterize different therapeutic approaches is in terms of different levels of attention demanded of the client: in many cases musical stimuli are deliberately kept quite simple - especially, strongly metered - so that they require relatively little attentional energy to follow and can act as a ‘carrier’ for other learning tasks. Some studies have found that simply playing ‘calm’ music aids the learning of complex tasks, with the likely explanation being that the music “entrain[s] the subjects into a relaxation stage of the learning cycle.” (236).21

Music therapists have distinguished three modes of entrainment for clinical purposes:

1. Primary entrainment: "the music, or some attribute of it, is matched directly to the physical or cognitive behaviour of the client. Once synchronized, modulation of the music causes change in those personal behaviours."

2. Secondary entrainment: "the music is synchronised with the material, skill, or concept to be learned. Many of the mnemonic uses of music utilize this process."

3. Tertiary entrainment: "the music is matched to the child’s functioning or preference level to cause a change in an unrelated behaviour…." (all from Rider and Eagle, 1986: 229).

The experience of music therapists also reinforces the message that we are dealing with non-linear systems whose behaviour is difficult to predict. An example of this comes from the use of entrainment in therapy with subjects exhibiting pathological ‘rocking’ behaviour. One study found that "retarded adolescents who body-rocked at different speeds responded differently to varying tempi of music. Generally, (a) fast rockers tended to slow down when the music’s tempo was slightly below that of their rocking tempo. (b) Slow rockers slowed down even further when the tempo of the music was slightly faster than their rocking tempo. (c) Music that matched the tempo of the rocking had a stimulative effect on the rocking of both groups. " (Stevens, summarised by Rider and Eagle 1986:232). On the basis of this and other studies Rider and Eagle conclude that "perseverative rocking behaviours of children are entrained best to musical tempi which are near, but not precisely the same as, the rocking state” (232).

Much of this work developed through therapeutic practice, without a clear sense of how or why auditory entrainment is an effective tool. Recent years have seen an increase in scientifically controlled studies of auditory entrainment however, many of them carried out by Michael Thaut and co-workers at Colorado State University. These studies have begun to demonstrate in greater detail how procedures such as "rhythmic auditory stimulation” can help those suffering from brain injuries or the effects of strokes, or diseases such as Parkinson’s or Huntingdon’s, with rehabilitation in tasks such as walking and reaching (see e.g. Hurt et al 1998, McIntosh et al 1997, Thaut et al 2002). There can be little doubt that further study of the use of entrainment in music therapy can yield further insights into the phenomenon, and in turn suggest lines of enquiry into non-clinical musical behaviours.

21 For a recent review of studies of music therapy and autism, see Wigram 2002.
5. Entrainment and ethnomusicology

5.1 Ethnomusicological studies relating to entrainment

Although entrainment studies are only now beginning to take shape in ethnomusicology, a number of earlier ethnomusicological approaches relate to our concerns. Indeed, we would argue that entrainment relates directly to most of the key concerns of ethnomusicology: the challenge is to make that connection clear and to investigate musical behaviour and concepts explicitly in terms of entrainment. This section discusses some of the more explicit connections with earlier ethnomusicological inquiry, focussing in particular on the work of Alan Lomax, John Blacking, Charles Keil and Steven Feld.

Alan Lomax was one of the earliest ethnomusicologists to take an interest in entrainment. His 1982 article "The cross-cultural variation of rhythmic style" is a concise introduction to the rhythmic aspects of his comparative approach, and this piece makes clear that he was concerned with gesture, embodiment and 'rhythmic style'; the relationship between conversation and musical interactions; and the role of rhythm in social relations.

Rhythm [plays a role] in linking people, by providing a common framework of identification. Rhythm is, after all, a prime mover in social relations. Rhythmic patterns facilitate the co-activity of groups and aid their members in coordinating energies and resources in work, nurturance, defence, social discourse, rites of passage, interchange of information, and, above all, expressive acts. The important role of rhythm in group behaviour suggests that we can view the rhythmic aspects of communication as essentially social in nature - a system that binds individuals together into effective groups and links groups into communities and polities. Each such "rhythmic style," passed on generationally, shapes many aspects of each cultural tradition… (Lomax 1982:149-150)

These fundamental concerns, however, were not pursued through study of the detailed mechanisms and effects of entrainment, but were developed as generalisations - using Lomax's Cantometrics system - of the relationship between rhythmic style as displayed in music, dance and conversation and the correlation of rhythmic style with other 'cultural' factors.22 Perhaps more compelling are Lomax's reported findings that individuals in different societies tend to move in different metrical patterns: "You can walk in a 1-1-1-1 meter - or in a 1-2-1-2 meter, or even in a 1-2-3-1-2-3. The upper body can simply go along with the legs or it can move to an independent meter or in an accompanying pattern… The combinations of the rhythmic patterns in the upper and lower body give rise to more complex meters. For example, Africans produce polyrhythms by moving arms and legs to different meters. One favourite Oriental rhythmic style consists of a steady four in the legs (and the percussion section) while the arms follow a free metered melody of a lead instrument." While the degree of generalisation may grate with some readers, it is difficult to contest that Lomax's speculations served to "open up a complex subject" by suggesting "how much light may be shed on musical rhythm by a study of its corporeal basis." (1982:161-2).

Lomax was of course, like many ethnomusicologists of his day, influenced by a tradition of comparative musicological thought that looked - however speculatively - for the roots of musical rhythm in bodily motion (see for instance Hornbostel 1928; Sachs 1953). Another particularly important figure in this

22 For instance: "The importance of women in the main productive system is related to the level of cohesiveness in the rhythmic style… The level of discipline in the child-rearing system is related to the regularity of rhythmic patterns" (1982:152-3).
area was John Blacking.23 Blacking’s 1977 essay “Towards an anthropology of the body”, for instance, talks about the importance of timing and movement, of the circadian cycle and the interrelationship of bodily rhythms. Blacking’s writings make frequent reference to the concerns of this paper - interaction, both verbal and non-verbal communication, the biological underpinning of communicative behaviour, and shared emotional or somatic states. He related these ideas to elements of Durkheim’s approach:

Co-operation and social interaction are not the consequence of rational contract or of habits learned during a long period of infant-mother dependency: they are biologically programmed and a necessary condition for the growth of distinctly human organisms. Durkheim’s reference to [society as] ‘a system of active forces’… implies powers of sensory awareness, of resonance and of communication between the individual parts of the social organism (Blacking 1977:8).

We would suggest that ‘entrainment’ would be a more appropriate term than ‘resonance’ here, as it seems that Blacking effectively describes the general parameters of entrainment, even though he never explicitly uses the term. He wrote elsewhere that the bodily basis of musicking requires “some degree of bodily resonance, as does monitoring speech” (1992:306). According to Blacking, “bodily resonance” (which he also referred to as “bodily empathy”) is the sensation or awareness of synchronising with the physical movements of others in a musical situation. Blacking described this as “the experience of ‘falling into phase’ that players shared” (1983:57). Bodily resonance is felt by the body both as an emotional connection and the physical sensation of co-ordinated motion. “Thus,” Blacking concludes, “sensuous, bodily experience was a consequence of correct musical performance… and a correct musical performance was a way of feeling” (1983:57). Blacking argued that when each performer played his/her part correctly, “the collective effort produced both new cultural forms for the ears of the performers and listeners and a richer, bodily experience for the participants” (ibid.) Blacking theorised there was a direct correlation between “bodily resonance” and increased “fellow feeling” - regard for others - through participating in social music-making, suggesting that the emotions and “somatic states” resulting from such profound aesthetic experiences were a critical force motivating a person’s identification with the social group that made the musical experience possible.

Another important strand of ethnomusicological thinking relating to entrainment is the work of Charles Keil and Steven Feld on ‘groove’, as well as Keil’s theory of ‘Participatory Discrepancies’ (PDs). Keil and Feld (1994:22-23) define “groove” in numerous ways that relate to entrainment, for instance, as the experience of “being together and tuning up to somebody else’s sense of time.” Among the many things that “groove” signifies, we suggest that “groove” could also be understood as the socio-musical process of being entrained at the preferred degree of synchronicity. The term “participatory discrepancy” is carefully chosen, as Keil demonstrates ([1987] 1994), to suggest both that musicking involves a sense of participation (referencing Levy-Bruhl and Barfield), and that participation is founded not on exact synchronisation but on appropriate degrees of being ‘out-of-time’. According to Keil, discrepancies – particularly in timing – are what create ‘groove’, or an activation of positive feel in the music. This idea is developed by Alén and Prögler in a special issue of Ethnomusicology (39/1, 1995) and also by Steven Feld in his discussions of Kaluli aesthetics: “The essence of "lift-up-over-sounding" is part relations that are in synchrony while out of phase” ([1988] 1994: 119, emphasis in original). Again, the concept of entrainment per se is not explicitly addressed in these publications, but its significance is quite clear. It seems likely that an encounter between entrainment and participatory discrepancy would be productive for both theories. The latter theory could perhaps benefit from the broader base afforded by the

23 Blacking, of course, acknowledged his own debt to Hornbostel's famous 1928 paper: see Blacking 1955.
entainment concept," while PDs may have a valuable role in directing entainment studies towards phase differences (see section 2.8).

5.2 The significance of entainment for ethnomusicology

We have seen how the entainment model has proved a powerful tool in other areas of research. It has begun to be seriously applied in musical scholarship, inspiring new perspectives on musical meter and on music’s role in human evolution, as well as practical applications in music therapy. It relates in profound ways to key issues in ethnomusicology, and resonates with the work of several ethnomusicologists of past and present generations - and yet there has been very little detailed work and virtually no empirical studies in ethnomusicology on entainment processes and their ramifications.

What, then, is the potential benefit of applying the entainment concept in ethnomusicology? What kind of studies might this idea stimulate, and what might those studies tell us?

An entainment model suggests we look at engagement with music not simply as a process of encoding and decoding information, but of embodied interaction and 'tuning-in' to musical stimuli. Musicking humans can be seen as embodying multiple oscillators (or endogenous rhythmic processes), oscillators which may be mutually entrained in a process of self-synchrony as well as entraining to external stimuli in the processes of making and engaging with musical sound. Entrainment in musicking implies a profound association between different humans at a physiological level and a shared propensity at a biological level. The implications of this view for studies of socialisation and identification are obvious, and so too is the link to questions of enculturation: someone’s ability to respond appropriately to a given musical stimulus can, since it is a learned application of a basic biological tendency, be a marker of the degree to which an individual 'belongs' in a particular social group (cf. Lomax).

The question of the degree of symmetry in entainment has obvious links to larger ethnomusicological questions. Studies in non-musical communication have shown that more dominant individuals tend to force the less powerful to adapt further in the process of mutual entainment, so that even where entainment is mutual it is not necessarily equally balanced. In some musical contexts, of course, some participants have little or no power to influence the sound production, directly or indirectly, and these might be seen as examples of asymmetrical entainment.

We need not spend too much time, perhaps, demonstrating that entainment exists in musicking - although even this would be a welcome development. The fact that musicians can synchronise their performances, and that they and others can dance or tap their feet demonstrates that they are in time with the music, i.e. entrained to some degree. It would also be regrettable if a fragmented research

24 For example, we might ask whether so-called “discrepancies” might be better described as complex rhythm patterns that are culturally defined. (Case studies 2 and 3 below offer examples of such complex rhythm patterns; see also Clynes 1986). We might also ask what are the physical and psychological parameters within which intentional play against a culturally defined rhythmic scheme is possible and potentially meaningful.

25 Of course, power relations observed in musical settings do not necessarily simply reflect those evident in other contexts. For example, there is often an inverse relation between degrees of socio-political power and the power to control a musical situation, such as when musicians come from subaltern, economically or politically marginalised groups. Our point is that a study of the exact musical processes involved in entainment has the potential to offer substantial evidence of who is entraining to whom and in which circumstances. Regardless of the specifics, one would expect studies of entainment to shed light on the power relations relevant to the musical context.
effort appeared to confirm Chapple's unfortunate prejudice that entrainment is a phenomenon only, or primarily applicable to a subset of "regular and driving rhythms." If entrainment is a factor in any interpersonal interaction and communication, we should expect that it is a factor in any variety of musicking. We feel that there are a couple of questions that may help to direct more productive research efforts: What is special about musical entrainment?, and How do processes of musical entrainment vary between musical traditions and contexts (i.e. culturally), as well as individually?

• What is special about musical entrainment?

This is perhaps a key question, and one for which we have as yet only the sketchiest of answers. It seems that certain kinds of musicking can afford particularly precise synchronisation between individuals' behaviours. This may in fact be one of the basic reasons for the development of these kinds of musical repertories, since musical metre is often (although is not invariably) more regular, with more hierarchical levels, than the rhythmic patterning of speech and other communicative behaviour. Since certain degrees of entrainment between individuals seem to be associated with positive affect, is it the case that particular patterns, periodicities, hierarchies or intensities of entrainment afford particular affects? Could positive affect be associated with a greater degree of self-synchrony as well as closer synchrony with a social group (cf. Blacking 1983, see above)? Do particular kinds of music (or for that matter, music enculturation) promote the switching of attention between temporal levels? All of the above seem possible, even probable, but serious work demonstrating not only that entrainment occurs in musicking, but also how it differs from the kinds of entrainment experienced without music, has yet to begin.

• How do processes of musical entrainment vary individually and with culture?

Something we should not lose sight of is that entrainment is a process which varies between individuals. For instance, not everyone at a performance appreciates the music in quite the same way; some may not "get it", or might appear to be entraining to a different beat (be "out of time"), and even those thoroughly enculturated listeners enjoying the music may be attending more or less to different temporal levels (see discussion of Jones, section 3.3 above). If these individual responses are averaged or summed over groups of listeners, it appears that there are also "cultural" differences - as any ethnomusicologist might guess - in the value placed on various types and degrees of entrainment. Listeners unfamiliar with a particular musical stimulus may fail to demonstrate an appropriate response to that music because they have not learned the "right," i.e., culturally appropriate, way to do so. It may be that studies of entrainment, if carefully designed, can illuminate long-standing issues of "cultural difference", of enculturation and acculturation, of what it is to be an "insider" to a musical culture, and so on, in a productive new light.

At a more general or abstract level still, a greater role for entrainment could be part of an ongoing shift towards a paradigm that sees the business of ethnomusicology as the investigation of musicking as embodied, interactive, communicative behaviour. Many ethnomusicologists will, with justification, feel that this is what they already do and were trained to do. Nonetheless, we feel that serious consideration of the concept of entrainment, and the development of methodologies which are empirical and experimental as well as ethnographic, could potentially lead to a significant shift in the focus of ethnomusicological enquiry. We also believe that this kind of approach may offer a useful contribution to discussions of musical affect, the place of emotion in music, and the concept of musical meaning.

Entrainment research within ethnomusicology relies upon the integration of musical, cognitive, and cultural theory, thereby allowing a broader description of how musical experience, while individually unique in every case, is nevertheless always social. Through exploring the phenomena of entrainment, ethnomusicologists may be able to better understand how musical sound serves as an interface that connects selves—viscerally and cognitively—to society.
Our intention in this paper has been both to set out, for an ethnomusicological readership, some of the key features of the entrainment concept and its application in music and other fields, and also to encourage our colleagues in ethnomusicology to join us in building on this base. Not only ethnomusicology, but other academic fields as well, stand to reap enormous benefits if we can develop this study further, through both individual and collaborative projects. Ethnomusicologists have an opportunity to work with scholars from other disciplines and to build interdisciplinary bridges, but we also have a chance to make a significant contribution and develop theories of wide relevance in the biological and social sciences. We firmly believe that there is a great deal that we need to explore about human thought and behaviour, which can only be done with a significant input from ethnomusicologists.

In the following sections we make some suggestions, ranging from the general to the specific, about how entrainment in musical behaviour can be investigated. We do not, of course, intend to set a limit on either topics of investigation or methods - such an attempt would be futile in any case - but to present a mixture of our own experience and our speculations. We hope that respondents to this article will have further observations and reflections, and indeed methodological suggestions, so that this publication as a whole will take on the character of a collaboration between more than just the three authors of this presentation.

6 Methods and Methodology

What are the most appropriate procedures for the study of entrainment processes? It should be clear that traditional ethnographic and music-analytical methods may generate hypotheses regarding musical entrainment, and to suggest areas of further investigation. To prove the occurrence of entrainment, however, and to describe specific entrainment processes in any detail, will require the collection and analysis of timing data derived from the musical sound, from observed motor movements and/or from other physiological processes associated with the performance. Entrainment is manifest only in actual behaviour and it is the temporal analysis of motor performances that supplies the essential data for entrainment research. As has been pointed out earlier, entrainment processes are largely constrained by non-conscious structural and procedural factors; although we must consider performers' intentions or motivations as revealed by their explicit statements they have to be evaluated against the chronometric data. The subsequent analyses of the data allow for inferences about the underlying motor performance and cognitive processes, which help us to uncover the guiding constraints and interactions between processes, either within individual performers or between performers.

One of the paradoxes here is that time analysis research runs against the grain of current ethnomusicological methods, despite the fact that entrainment theory itself resonates strongly with certain aspects of ethnomusicological thinking. Entrainment requires empirical study of something which cannot be verbalised: ethnomusicological orthodoxy stresses the primacy of informants' verbal accounts, and most ethnomusicological reports operate at a meta-discursive level. Blacking recognized this problem, writing:

> How can we measure the apparently invisible and how can we presume to say that something not recognized by the actors is real? It is as real for purposes of explanation as atoms, genes, or successive differential counts of leucocytes: that is, it may be inferred from certain kinds of observable behaviour. (Blacking 1977:17)

Blacking was not himself known for empirical scientific work - and even those ethnomusicologists who are sympathetic to the idea of such study rarely have the scientific background to actually carry it out. Therefore, if we agree that empirical studies are in principle a good thing, there is a need for practical
suggestions for how such studies can be carried out in the context of ethnographic work. The following
discussion assumes that ethnomusicologists will want to develop methodologies that have an
ethnographic element, even if they also incorporate empirical study based on evidence not collected
ethnographically, and perhaps a limited amount of controlled experimental work. Although this paper
is not a manual on ‘how to do entrainment research’, we do present below some general suggestions
regarding research methods. The following sections discuss ways of investigating entrainment, and for
convenience are divided into three sections: data collection; data analysis; and finally ethnography and
interpretation.

6.1 Data collection

We begin by considering the possible sources of evidence, or modes of data collection, available to
ethnomusicologists. They can be considered in four categories:

1. Ethnographic investigation and introspection: We should be looking to establish how the degree,
rate (and other aspects) of entrainment correlates with dance and other physical responses, and
with ethnographic reports of affect, emotional response, and so on. Discourse about music - which
is accessible only through ethnographic methods - can inform us about how entrainment feels,
and incorporates the metaphors people use for being ‘in good time’ with fellow musicians (or not).

2. Musical sound: As the case studies on chronometric analysis below demonstrate, a lot can be
inferred from recorded sound itself (not, we hasten to add, from transcriptions such as those
rendered in Western notation, which can suggest entrainment possibilities only indirectly and
unreliably). This is the main focus of the case studies below.

3. Visible physical behaviour (gesture): The work of Condon, Kendon, McNeill and others suggests
that, apart from musical sound itself, the most important evidence for entrainment in musical
behaviour might be the physical movements, or gestures, of participants, which can be studied by
means including video or film analysis. This is an area that needs careful consideration. For
instance, the problems of how to segment continuous behaviour into discrete gestures, and the
difficulty of assigning gestures carried out over a finite duration to a specific point in time, are not
trivial. If these methodological difficulties can be solved or by-passed, then this kind of study offers
huge potential. This kind of data can be gathered either in ‘real-life’ or experimental situations.

4. Physiological processes (heart rate, respiration, brain waves etc): There are almost certainly
important musical entrainment effects which cannot be observed either in the physical sound wave
or in visible motor movements. Possible areas of investigation here include electroencephalograms
(EEG, for recording brain waves), magnetoencephalograms (MEG, for recording the location of
brain activity), electrocardiograms (ECG, for recording heart rate), electromyograms (EMG, for
measuring muscular activity), body temperature, and more. In most cases, the procedures involved
would present difficulties in field-work situations due to the intrusive nature of the equipment
needed to make these types of measurements, and to most ethnomusicologists’ lack of training in
operating this kind of medical apparatus. Nonetheless advances in technology, and developments in
experimental and collaborative research methods, may in time overcome some of these difficulties.

26 We do not consider here methods such as those described by Large and Kolen for the investigation of
metre, which are essentially theoretical and involve experimental testing of computer simulations of
neurological structures: this is not to criticize such methods, simply to recognize that they lie outside the
remit of ethnomusicology.
Audio recording

Audio recordings are an essential form of data for entrainment research in music because of their excellent time resolution. With the audio equipment available today, good quality recordings are not very difficult to produce. It is advisable to generally use test or reference signals (e.g. 1kHz test tones) that allow estimation of the reliability in the reproduction and transfer processes. This does not only hold for analogue equipment, but also if digital gear is used: sometimes sound cards do not work properly with certain software (especially on PCs), or operating software may be corrupt; such malfunctioning can easily be detected by checking the reference signals.

Although recording machines do not cause too much of a problem, the number and placement of microphones requires careful consideration. The type of analyses afforded by audio recordings is very much dependent on the type of music performance recorded and the manner of recording. To be able to analyse simultaneous musical processes it is essential that the corresponding sound signals are separated or separable as much as possible. If the sounds involved have distinct spectra for instance, no other separation might be needed (for an example see the clapsticks in Appendix C), in other cases a well planned stereo separation may be necessary and sufficient. If more than two processes are to be analysed multitrack recording may be a necessity.

It has been our experience that extant field recordings can be useful for investigating entrainment phenomena that are evident in music. Our best success has been with recordings of music performances that involve small numbers of performers and/or relatively discrete, individually identifiable instruments. Examples of what can be done with such recordings follow in the three case studies below, all three dealing with recordings involving a small number of performers. Extant recordings of large groups of music participants (such as of festivals, ceremonies, and rituals) present added analytical challenges and may not be usable for some types of chronometric analysis. However, if the research focus is not on individual but in group behaviour – requiring a different type of analysis – even these recordings may offer interesting material. In some cases recording quality and the identification of individual musicians or instruments can be enhanced to a degree using digital audio processing tools (Peak and Pro-Tools are widely available tools for digital post-processing, but as far as complex filtering is required to isolate or enhance certain acoustic events from a recording, CoolEdit (only for PCs) is very flexible and easy to handle).

Video recording

While audio recording is of course indispensable to this research, for many (perhaps most) purposes some other form of data will also be required on motor movement patterns or other physiological processes. The most accessible method, after audio recording, is undoubtedly video recording. Although video may be valuable, it is important to bear in mind the relative strengths and weaknesses of audio and video data, the most obvious of which lie in the domain of sampling rates. Audio has a much higher definition than video (sampling rates up to 48kHz on standard digital audio equipment are some 2,000 times higher than standard film or video frame rates). Another factor is that proper set-up for video recording is far more critical than even for audio recording. With audio, it may be possible to get around the fact that microphones were poorly placed or levels not set optimally. But with video, if the camera is not properly set up, then the footage will certainly be unusable for movement analysis.

Some of the basic points regarding the collection of video data are the following:
• for movement analysis the camera shot needs to be static (no zooming, panning or tracking).

• footage from a single camera can be used to analyse movement in two dimensions only (perpendicular to the direction the camera is pointing); ideally a reference frame of known dimensions should be placed in shot to allow calibration (although in practice a rough idea of the dimensions of the space is often sufficient).

• three-dimensional movement analysis requires at least two cameras: these need to be calibrated using a specially constructed three-dimensional model, and synchronised. Depending on the performance situation, more than two cameras may be necessary. (Note that biomechanics laboratories tend to use at least five or six synchronised cameras).

• a variety of camera speeds is available (e.g. the PAL and SECAM standards use 25Hz, NTSC uses 30Hz, but specialist high-speed cameras are also available). 25-30Hz may prove to be adequate for many research purposes, although this remains to be seen; using high-speed cameras in any case produces far more data, which takes proportionately longer to analyse, and therefore sets practical limitations.

There are other ways to gather movement timing data, and over the last century or more each generation has found ways to adapt extant technologies. For instance, some researchers in recent years have used the MIDI – musical instrument digital interface – system (see Clarke 1995, Arom 1976, Busse 2002).

Modern motion capture systems, such as those used in the entertainment industries, can offer possibilities not afforded by traditional video: by recording movement data directly rather than storing video images, they are able to use high-speed cameras and can handle data with timing resolution much higher than conventional video. Although promising, these systems are very expensive, intrusive (they require reflective markers to be fixed to subjects), and at present would be impossible to employ in most fieldwork situations.

For other physiological data standard medical equipment is required, which again can be difficult or impossible in many field-work settings. Some motion analysis systems (including high-end video recording systems) do allow other data such as EMG to be recorded in parallel with video or motion data.

6.2 Data analysis

Attempts at analysis of rhythm in ethnomusicological research go back to the early days of comparative musicology, and have been facilitated by important technological developments including Seeger’s melograph as well as modern audio editing software. In spite of Seeger’s promotion of this technology however, desire for this type of detail in musical analysis has waned and the methods he promoted have rarely taken centre stage in recent ethnomusicology. Recent efforts at detailed rhythm analysis in ethnomusicology based on audio timing data include the work of Prögler and AlØn (of jazz rhythm section and Cuban dance: see Alén 1995, Prögler 1995), Widdess and Clayton (separately on Indian

27 IRCAM, for example, has had a fruitful history collaborating with ethnomusicologists to develop and adapt technologies for experimental applications for use in the field.

28 Such as those produced by Motion Analysis Corp. and VICON.
music: see e.g. Widdess 1994, Clayton 2000) and Will (on Australian Aboriginal music; Will, 1998; Will, in press).

Film and video have also (of course) been widely used by ethnomusicologists to assist in musical transcription and analysis (see the work of Kubik (1965) with African xylophone players, Baily (1985) with Afghan lute players, and others), as well as for documentation. It has also been used in behavioural observation: the most detailed exposition of this method is in Qureshi’s work on South Asian qawwali (1987, 1995[1986]).

Even in most of these cases, the analysis of timing data is relatively crude, and does not employ the scope offered by statistical analysis software. This section considers the latter possibility, beginning with methods for extracting timing data from audio and video recordings.

Audio analysis

Given good quality recordings, extraction and processing of audio data in computer programs does not seem to present serious difficulties. Recordings need to be uploaded to a computer (in the case of analogue recordings, most present day computers have the necessary analog-digital converters built in). Depending on the variables of interest (timbre, pitch, rhythm, or other large scale periodicities) the digitized audio files can be re-sampled (e.g. from 44kHz down to 11kHz) to reduce the data volume. For rhythm or periodicity analyses, the re-sampled sound files are then loaded into an editor program, e.g. Praat or Sigmund that permits one to label or mark sound events of interest in a defined and consistent way. These markings, together with the respective time information, are then saved as text (ASCII) files that form the raw data of the measurements and can then be displayed, processed, and analyzed with a statistics program. The case studies below include practical examples of these procedures.

Video analysis

Data extraction and analysis from video can be split into two basic approaches:

a. **Behavioural observation** involves marking down event onsets (e.g. when someone starts or stops singing) and the frames at which they occur. In Condon’s day, this work had to be done manually, which was a notoriously time-consuming business. Nowadays, software exists for the automatic logging of behavioural observations, and the linking of these observations to video time code (e.g.

29 A related and important contribution from jazz studies was published as a special volume of *Music Perception* (2002). A historical summary of methods of chronometric investigations related to jazz studies is included (Collier & Collier), as well as examples of measuring and analytical techniques, including the use of MIDI (Busse). Topics of the contributions cover swing and ensemble timing (Friberg & Sundström), as well as expressive “microtiming” (Iyer 387).

30 Praat [http://www.fon.hum.uva.nl/praat/](http://www.fon.hum.uva.nl/praat/); Sigmund: contact U.Will (will.51@osu.edu)

Observer Video-Pro).\textsuperscript{32} Such software itself can be used for statistical analysis, and/or allow export of timing data for further statistical analysis in other programs.

b. **Movement tracking** involves marking particular parts of the subject’s body (e.g. joints) manually as they appear on a video frame displayed on a computer screen. Once this time-consuming job has been completed for a sequence of movement, software can analyse the movement and generate data on variables such as the velocity and direction of movement. This is fairly simply done for two dimensional analysis.\textsuperscript{33} It can also be done for three dimensional analysis, provided the recordings have been properly prepared.\textsuperscript{34} As noted above, three dimensional movement tracking can also be done automatically using modern motion capture systems. In some cases these systems integrate automatic motion tracking with video recording.

The tradition of sound-film analysis from Condon and Birdwhistell onwards has looked for a correlation between the articulations of speech and physical gestures, and has suggested that periodicities in the region of 0.1 sec (frequencies c.10 Hz) can be observed. Using video shot at 30 frames per second, the limit for observation of periodic behaviours is maximally 15 Hz (0.06 sec); for 25 frames per second (PAL) this is 12.6 Hz (0.08 sec).\textsuperscript{35} This might suggest the level at which we may look for periodicities in visible behaviour. If periodic behaviours with much faster frequencies are significant, they will have to be investigated using high-speed cameras or motion capture systems.

In these analyses we should be aware of the relationship between analogical and digital processes: gestures can often be described either in terms of an event occurring at a time point (e.g. a foot tapping, a stick striking a drum head) or as a continuous gesture (a melodic phrase or ornament, the movement of arm and hand preparing to strike the drum head). Since time points are much easier to deal with mathematically, it is often expedient to treat musical behaviours in such discrete, 'digital' terms, but one should be aware that this potentially oversimplifies the analysis.

**Analysis of time series data**

Although synchronization does not in itself prove entrainment, entrainment can only be suspected if two oscillators interact and their behaviour becomes coordinated or synchronized. The first step of analysis, therefore, is to look for synchronization. Generally, caution is demanded, as synchronization is a complex dynamic process, not a fixed state. A single observation is not sufficient to identify synchronization: what are required are time series data, i.e. a series of observation data or measurements of the respective (system) behaviours in time. These data can then be analysed in order to detect the presence of synchronization. In some cases, this may be an easy task. However, the detection of synchronization of irregular oscillators and noisy systems is not so easy. Simple visual inspection of the data may often not be sufficient and the mere estimation of phase and frequency of complex time series may be complicated in real world data. The more noisy the data are – and real

---

\textsuperscript{32} Observer Video-Pro [http://www.exetersoftware.com/cat/observer.html](http://www.exetersoftware.com/cat/observer.html)

\textsuperscript{33} Using programs such as SiliconCoach: [http://www.siliconcoach.com/](http://www.siliconcoach.com/)

\textsuperscript{34} Using systems such as Peak Motus, [http://www.peakperform.com/](http://www.peakperform.com/); or APAS, [http://www.sportscience.org/Main/adw-04.html](http://www.sportscience.org/Main/adw-04.html)

\textsuperscript{35} Software is available for splitting each video frame into its two constituent two "fields" (e.g. SiliconCoach), which can effectively double the timing resolution of video recordings (with a commensurate loss of picture resolution).
systems are always noisy, not only in ethnomusicology – the more difficult it is to distinguish between synchronous and asynchronous states or to detect transitions to synchronization. Some basic approaches in the analysis of chronometric data are demonstrated in the case studies that follow.

However, the mere demonstration of a covariation of two variables is not sufficient to prove the presence of entrainment (some studies have shown that correlated signals may in fact originate from unsynchronised oscillators, e.g. Tass et al. 1998). In order to identify entrainment one needs to examine perturbations or transitions of the synchronization process; only if synchronization is re-established after these disturbances - as in Huygens’ clocks - does it seem justified to describe the interaction between the oscillators as entrainment. We mentioned above the complex oscillations of real life processes. Studies have clearly shown that, being coupled, irregular or complex oscillators can also undergo entrainment (e.g. Rössler 1976). In these cases, however, it might be necessary to use a different set of variables to describe the systems, for instance to replace frequency with short time mean frequency (i.e. frequencies averaged over some tenth or hundreds of milliseconds), because entrainment of these complex systems might follow a temporary mean value of these variables, not any instantaneous values. However, frequency related measures may not be the most adequate to analyse entrainment, because they concern events occurring in individual components of the entrainment process. The phase relation between the components involved is obviously better suited to describe observed patterns of interaction as well as transitions between them. The calculation of one such measure, a point estimate of the relative phase, is described in appendix C, and applications are demonstrated in case study 2 and 3.

6.3 Ethnography and interpretation

As we suggested above, it is our belief that empirical and/or experimental research can flow out of and serve to illuminate the interpretation of ethnographic research. Serious consideration of entrainment in ethnomusicological research, including the development of methodologies that are empirical and experimental as well as ethnographic, could potentially lead to a significant shift in the focus of ethnomusicological enquiry. The particular ways in which ethnographic and empirical methods can be integrated will vary substantially between research contexts, so it is difficult to offer detailed suggestions here. We do believe however that each should feed into the other: in other words, (a) ethnography ought to direct empirical research, as ethnographic research will offer intuitions as to which phenomena seem to be particularly important in particular performance contexts; and (b) empirical research can generate questions for ethnography; e.g. if the numerical data tells us that an interesting entrainment effect occurs at a certain point in a performance, we may show greater interest in that portion in future field work. Furthermore, if numerical analysis allows us to identify and analyse entrainment processes, in many cases it will be the ethnographic knowledge that permits the identification of intervening variables and order parameters\(^{36}\) and an ethnomusicologically relevant interpretation of the analysis (for an example see case study 3).

---

\(^{36}\) These are parameters that determine a system’s behaviour at transitions from one state into another, e.g. at transitions from synchrony to asynchrony or from an entrained to an non-entrained state and vice versa.
7. Case studies of chronometric analysis of rhythm performances

In the following sections we are going to present some case studies that demonstrate a couple of entrainment phenomena and illustrate some basic procedures of entrainment research. We begin in case study 1 with an analysis of rhythmic motor behaviour by single performers and ask, what time series analysis can tell us about the individual components of an entrainment process. Although the focus is more on general aspects of rhythmic performances than on entrainment itself, this analysis will give us a chance to present basic methodological approaches and elementary concepts in temporal analysis, a field largely ignored by ethnomusicology but essential for entrainment research. In the second case study we will then analyse the phenomenon of self-entrainment, the coordination of simultaneous motor activities in individual performers. In that study we will also touch upon questions of experimental and non-experimental data and the importance of comparative studies. Finally, in the third case study, we are going to analyse a process of synchronization and entrainment between two performers. Although the analyses of these case studies are done here with virtually no ethnographic input, it will become evident, at least for cases two and three, that ethnographic data are an indispensable and integral part the interpretation of the analyses. These analyses are but a first, though essential, step in understanding how entrainment works in a socio-cultural context like music making. Before presenting the case studies we would like to introduce some basic concepts and procedures of time series analysis that, we feel, might be helpful in following the subsequent analyses.

As discussed above (section 6.2), audio recordings can be saved on a computer, resampled, and loaded into an editor program such as Praat, where relevant events can be labelled (Figure 1 illustrates an audio file in such a program, with time labels for some of the bell strokes). The raw data sets which these labels generate are time-ordered series of measurements – sound events in performance time – in which the chronological sequence is characteristic for each data set. This is important, because time series analyses can inform us about the dynamics of the underlying processes, something a randomly sampled data set would not afford. Often the primary interest is not to know when a certain event occurred but rather how is it related to preceding or successive events. In these cases we need to calculate the time differences between successive events, which gives us the period length – if we are dealing with periodic events - or the duration between events. Fig.2 displays such a series of durations for the bell pattern of fig.1. This performance consists of two types of strokes, a short and a long one, displayed as two different groups of dots in fig.2.

Fig.1: Sound wave of a rhythm played on an Ewe iron bell, gankogui (Kpeel-su audio cassette, White Cliffs Media, 1992, ISBN#:0-941677-42-7) The maxima of the first four bell sounds are labeled by vertical markers. These markers and the respective time information form the basic data file for the chronometric analysis.
In time series analysis, as in most other analyses, it is assumed that the data consist of a systematic pattern combined with random noise, which usually makes the pattern difficult to identify. Time series data can be described and analysed in terms of basic, identifiable classes of components: trend, cyclical or seasonal, and irregular components. The trend component, for example, represents a general systematic component that changes over time and does not repeat (or at least does not repeat within the time range captured by the data). Cyclical components are those variations in the data that occur repeatedly and contain information concerning the underlying processes of the rhythmic behaviour.

For the case considered here, the trend has an obvious musical interpretation; it is the variation in tempo during the performance. In the performance of the above Kpegisu bell pattern the tempo accelerates from the beginning towards the 15 sec time mark, slows down slightly towards the 35 sec mark, accelerates again and finally slows down after the 70 sec mark (see fig.3).

![Fig.2: Plot of the bell stroke durations vs their time of occurrence in a Kpegisu song performance. Performance time (in sec) is displayed on the horizontal axis, and the duration of the bell strokes (length between two subsequent strokes) is displayed on the vertical axis. Obviously the performance consists of two different types of strokes, short (ca. 0.25 sec; lower series of dots) and long ones (ca. 0.425 sec; upper series of dots).](image)

![Fig.3: Plot of the long bell stroke durations vs time of occurrence from fig.2. The continuous curve represents the trend component and indicates the tempo changes during the performance.](image)
interfere with the subsequent analysis that assumes stationary processes (for more information on time series analysis see e.g. Box and Jenkins, 1976). This is the case in the following study in which we are analysing the time structure of rhythm performances and its cognitive implications in order to better understand what is involved when a musical behaviour becomes entrained.

**Case study 1: Production of one rhythm by one performer**

In section 2.6 (self-entrainment) we pointed out the importance of an understanding of the internal timer organisation for understanding the temporal organization of the various movement components in music making. We also indicated that the timer organization could put important constraints on the way in which rhythmic actions can interact and entrain. Take for example a musician performing two simultaneous rhythms: The possible ways in which these rhythms can entrain, the modes they can ‘lock in’ with each other, depend very much on whether the two rhythms are controlled by a central clock – a master timer – or not. Now, in order to analyse entrainment and to detect the constraints at work ethnomusicologists do not have to wait for psychologists or neuroscientists to tell them what the timer organization underlying the observed behaviour actually is. All the necessary information can be obtained from the data at hand, the audio tracks of the documented performances. In the following case study we are going to indicate some possibilities of how this information can be extracted, how certain assumptions about timer organization relevant for entrainment studies can be tested, and how such an analysis helps us to identify possible constraints for the entrainment of multiple rhythmic processes. Although this case study does not demonstrate entrainment as such, it illustrates some basic concepts and procedures essential to entrainment research.

The basic ideas to begin with are as follows. It is one of the basic tenets of cognitive sciences that cognitive processes consume time, and more complex processes consume more time than simple ones. This principle is for example exploited in reaction time experiments. These experiments study the effect of specific factors on cognitive tasks by analysing the influence they have on the timing of the task responses. Complex, temporally structured behaviour such as speech or musical performance requires two different operations, (1) the preparation of the movement components and their correct temporal arrangement (establishment of the serial order of the behaviour), and (2) the timing of the units of the behavioural stream.

Models of motor behaviour such as that of Sternberg et al. (1978), assume that abstract representations of the movements are assembled and their serial order established prior to the response onset. These motor programs are consulted prior to and during the response execution to activate the peripheral neuro-muscular effector system. Under these models, the complexity of the cognitive processing that leads to the generation of the motor programs is reflected only in the response onset time and the complexity of the motor program is reflected in the time needed to consult the motor program during the response execution: the more complex the motor program, the longer the response time. For these models, cognitive processes that prepare the motor programs for execution only exert an influence on the timing of the onset of the response units, but not on the time course of the response itself.

However, it has recently been shown that different hierarchical levels of language processing do in fact influence the time structure in written language production (writing and typing). These findings suggest a processing architecture in which the peripheral motor system essentially connects at several hierarchical levels with central processing units (Will et al. 2002, in press). Such an architecture can account for the fact that, besides influences of the peripheral motor system, central processes like lexical access are also reflected in the time structure of serially organized motor behaviour like writing and typing.
In contrast to speaking, writing and some other, implicitly timed motor behaviour, musical behaviour is generally explicitly timed, i.e. constrained by the tempo requirements of the performance. One interesting question in this case is, whether under these temporal constraints hierarchical representations can be reflected in the temporal structures of rhythmic performances or whether musical timing can be adequately accounted for by only the sequential organization of temporal behaviour (for a more detailed explanation of these two different timer models, see Appendix B). We are trying to get an answer by analysing two examples of rhythmic performances, a West-African bell stroke pattern and a clap stick pattern from North-Eastern Australia. We shall start our analysis by testing an existing model for the analysis of temporal properties of rhythmic performances.

This model, which seems to be of interest for analyses of musical performances, was originally proposed by Wing and Kristofferson (1973a,b) for the timing of inter-response or inter-onset intervals in periodic tapping. It is based on the idea that the temporal variation or inaccuracy in periodic finger tapping can be explained by two components, the variability of a (hypothetical) central timekeeper and the temporal noise in the executing motor system. The hypothesized underlying process is described as a first order moving average process, with the expected timekeeper interval as constant and the temporal noise as random errors: this means that the tapping interval at any point in the sequence equals the timekeeper interval plus a noise component that is proportional to the mean of some previous noise components. The hypothesis can be tested with the autocorrelation function. (This function describes how a variable (a sequence of values) correlates with a delayed or shifted versions of the same variable. See Appendix A). For a first order moving average process the autocorrelation function for lags > 1 is zero, and in the Wing model the autocorrelation at lag 1 is negative. This makes it interesting for music performances, because it means that, as a stationary process, it is self-correcting: if one interval is larger than the expected value the next one will be shorter, and so forth. Models similar to the Wing model have been developed for speech timing (Kozhevnikow and Chistovitsch 1965) and for bio-rhythms (Ten Hoopen and Reuver 1967). However, the Wing model does not assume that higher cognitive processes are reflected in the time structure of motor performances and we want to test, whether this model is suited to adequately describe rhythmic process in performed music.

Case A: Kpegisu

The first item analysed here is one of the Kpegisu (Ewe war drum) songs, ‘Agbeme nuawo ken li’ as performed by Godwin Agbeli, who sings and plays the bell pattern (the Kpegisu audio cassette was published by White Cliffs Media, 1992, ISBN#:0-941677-42-7). For background information on this music see: Locke, 1992.

The Kpegisu bell pattern, which consists of seven strokes, has already been introduced above. In order to keep this presentation as simple as possible, we are going to analyse only the sequence of long strokes (5 per pattern), but the analysis can be extended to cover the complete bell pattern without changing the main results of our study. Prior to the autocorrelation analysis we ‘detrended’ the data, that is, we eliminated that component of the data variability that is due to tempo changes. The detrended data are plotted in fig.4. In comparing this graph with the previous one, a reduction in the range of variability (distance from the highest to the lowest data value) from about 0.13 sec to 0.085 sec is obvious, and the dotted line in the lower graph can be thought of as the flattened, straightened out tempo curve of the upper graph. The data variability that remains, then, cannot be explained by performance tempo variations, and we are now trying to identify the underlying process that produces these remaining variations in the data.
The apparent random variation of the beat durations is in fact highly structured. This structure is disclosed by marking all durations sequentially in their order of appearance within the pattern (see. fig.4; stroke numbers 3 and 7 are missing because they indicate the short strokes not considered here). Note for example the clear separation of strokes number 5 and 6.

![Plot of the residual values for the long bell stroke durations following the elimination of the trend.](image)

**Fig. 4:** Plot of the residual values for the long bell stroke durations following the elimination of the trend.

In order to specify the underlying process for these data, we calculate the autocorrelation function. As predicted by the Wing model, the plot of the autocorrelation function (fig.5) shows a significant negative value at lag 1 (in musical terms: if one beat was too short the next beat makes up for it in being longer, and vice versa). However the values for lags>1 are not zero and for lag 4 and 5 and multiples thereof we obtain significant negative and positive values, respectively. Hence, the underlying process cannot be considered to be a first order moving average process as predicted by the Wing model. The autocorrelation function is periodic with a period length of 5 lags. As these sequential dependencies demonstrate, the sequence of long bell strokes is obviously organized into higher-level units the size of which is expressed in the periodicity of the autocorrelation function. Possibly this is an indication of a hierarchical timer organization in which a high-level timer controls the duration of the rhythmic groups and subordinate timers control the durations within the groups.

One possible realization of such an organization could be as follows: The highest level timer initiates the execution of the first stroke and subordinate timers activate execution of subsequent strokes. At the end of the pattern, in the present analysis after five strokes, the highest level timer becomes active again and the whole cycle repeats. Now, if we examine the variance of the sums of each five successive bell strokes, the model makes the following predictions: the variance should be smaller if the successive strokes belong to the same group (to the same cycle of timer activation, starting with the highest level one) than if they are taken from two adjacent groups ( the variance of the sums of beats 1+2+3+4+5 and 3+4+5+1+2 should be different if in the second case 3+4+5 and 1+2 belong to different groups). This follows from the fact that in the latter case timing variations from two groups contribute to and increases the variance of the summed responses (see Appendix B). However, our tests for equality of variance, performed on the various sums did not reveal any significant differences (all p> 0.27). There seems to be no evidence for a hierarchical timekeeper organization; the results are, however compatible with a multiple sequential timer model.
In time with the music

Fig.5: Autocorrelation plot of the detrended long bell strokes of a Kpegisu song (strokes 1, 2, 4, 5, and 6).

Obviously, the timing structure of the performance of the Kpegisu bell pattern is more complex than suggested by the Wing and Kristofferson model (1973a,b). On the one hand, the autocorrelation function clearly indicates an organization of the bell strokes into higher-level units, and it seems that it is a complex pattern of seven bell strokes (5 long and 2 short strokes), or rather a cognitive template of this bell stroke pattern, that is ‘programmed’, not just a sequence of individual strokes: The complete pattern has a time structure, that recurs with every repeat of the pattern. Interestingly, the timing structure shows no particular representation for the beginning of the group, which will have to be marked by additional, dynamic features. This finding may be an important factor in explaining the flexibility in performance and alignment of this pattern with other, simultaneously performed rhythmic patterns. On the other hand, there is no indication of a hierarchical time keeper organization and the execution of the strokes seems to be controlled by a sequential time keeper. However, our analysis does not support the assumption of a fastest pulse as the basis for the performance of bell patterns. This assumption is made by some scholars of African Music (e.g. Kubik 1998; Koetting 1970) to explain certain features of African rhythms. If this were the case, we should have been able to identify a hierarchical arrangement of a pulse generator plus additional generators for the short and long bell strokes. Furthermore, the mean stroke duration should have been the same for all long bell strokes as they all span two units of the ‘fastest pulse generator’. However, our results show (see fig.3) that stroke 5 and 6 differ significantly from the other long strokes, a result incompatible with the assumption of a common ‘fastest pulse generator’ for the performance of this pattern.

Case B: Gama AT

The second example is a song from the CD “Dyirbal Song Poetry” published together with a book of the same name by Dixon and Koch (1996). The musical tradition is that of the Dyirbal speaking people south of Cairns in North Queensland (Dixon, 1972). They have two categories of songs, ‘corroboree’ and ‘love songs’, and the example analysed here is from one genre of the former, Gama songs. Dyirbal songs are performed by individual singers, mostly accompanied by clap sticks or boomerangs, played by
the singers, and in some cases by an additional ‘lap drum’ – a skin stretched across thighs, played by a woman (Dixon & Koch, 1996).

As the tempo variation for this performance is considerably smaller than the variability in beat duration, we did not eliminate any trend from the time series. At first glance, the plot of the beat durations (fig.6) looks like a series of random variations around a mean of about 0.3 sec, but in fact, it is highly structured. Marking all durations sequentially by two arbitrary labels (dots and circles in fig.6) and plotting them correspondingly discloses this structure: on the average, every second beat seems to have a slightly longer duration. This grouping into two beats is confirmed by an analysis of variance, showing that the difference between the two types of beat duration is highly significant ($F(1,100)=39.298$, $p<0.0001$) with a mean duration for the two types of 291 msec and 310 msec, respectively.

![Fig.6: Time-duration plot of clap stick beating in Gama AT.](image)

The calculated autocorrelation function shows a significant negative value at lag 1, as predicted by the Wing model, but as for the Kpegisu bell, values for lags>1 are not zero. The autocorrelation function is again periodic, but now with a period length of 2 lags. The significantly positive values at lag two and multiples thereof is obviously an indication of the higher-order grouping of the clap stick beats into units of two. Again, the test for a hierarchical timer arrangement was negative (see previous case): The variance for summed clap durations within and between the groups did not show any significant differences ($F(1,48)=0.735$, $p=0.289$), and therefore suggests a sequential timer model.

**Discussion**

The timing structure in both the African and the Australian example is clearly more complex than suggested by the Wing and Kristofferson model (1973a,b). A comparison of the variance of within- and between-pattern durations showed that the higher-level units of organization of the clap stick beats, as revealed by the autocorrelation function, are not an indication of a hierarchical timer organization. In both examples the results are, however, compatible with a sequential timer model. What, then, is the significance of the periodic structure of the autocorrelation function?

Following the interpretation of comparable results from a grouped tapping experiment by Vorberg and Hambuch (1977), we suggest that the higher-level organization indicated by the autocorrelation function reflects the establishment of serial order of the pattern to be performed. It is a more central, cognitive aspect of the behaviour and relates to the mental representation of the bell pattern, the pattern template. This means, what performers recall or activate in performance is not a series of individual strokes but complete patterns with distinct temporal fine structures. On the other hand, a ‘peripheral’ sequential timekeeper seems to control the duration of the sequence of strokes during motor execution and its organization is reflected in the variance of the stroke timing.
In both the Kpegisu bell pattern and the Gama AT clap stick pattern, the time structure of the rhythmic performance appears to contain information about at least two aspects of the motor behaviour, the control of timing and control of serial order of the rhythmic pattern. The time structure therefore informs us not only about the execution of motor programs but also about some aspects of the cognitive processes that lead to these motor programs, and musical behaviour shares this feature with other timed behaviour like writing and typing (Will et al. 2002, in press). It is these two aspects of the temporal structure of musical rhythmic activities that we are dealing with when we try to understand the phenomenon of synchronization and entrainment between two or more of such activities. Because of the time constraints of most musical behaviour, it would be most interesting and challenging to explore how these two aspects are involved in and contribute to the observed phenomena.

Case study 2: Simultaneous production of two rhythmic activities by one performer

In this case study we are going to examine two examples of simultaneous rhythmic processes in one body. The question in these examples is: How can we detect the presence or absence of synchronization between the processes and what does the identification of synchronization tell us about entrainment? We are going to show this by a further examination of the two musical examples from case study 1. The analysis of the individual examples, being somewhat representative for a range of ethnomusicological material, shows that the mere identification of synchronization may hint at, but is no proof for entrainment. Interestingly, however, the cross-cultural comparison of the two examples suggests that socio-cultural factors may influence entrainment, thereby requiring a reformulation of certain entrainment phenomena that hitherto have been interpreted exclusively on the basis of results from laboratory experiments. This control function, the re-evaluation within a variety of different contexts, is obviously one of the important contributions ethnomusicology can make to entrainment research.

On the central issue of coordination of limbs and body in performance of complex actions, dynamic system theory in motor control provides a tool for analysing the way that individual body parts cooperate to form patterns in space and time (e.g. Kugler and Turvey 1987). Cooperation is thought to result from the coherence of the movements of the parts under certain energetic constraints, without any explicit or conscious control, i.e. without cognitive intervention (Kelso 1995). Among the best studied cases are those concerning rhythmic limb movements in humans and other animals – several of these studies have already been mentioned in previous sections – and it seems that quasi automatic self-entrainment is dominantly found in homologous limb movements, for instance in coordination of movements of the right and left arm. Kinsbourne’s functional distance principle may provide a possible explanation for this. It posits that two concurrent (motor) activities interfere with each other to the extent that they are based on similar neuronal activity patterns in highly connected cerebral regions. Now, simultaneous movements in two corresponding limbs requires, amongst others, the activation of homologous structures on both sides of the brain, some of them directly connected via the corpus callosum, which might contribute to the strong coupling observed in this type of movement coordination. However, strength and automaticity of self-entrainment for movements in non-homologous body parts are less clear. Although numerous experimental studies have found an influence of vocal activities like speaking on arm or finger movements, the review of Kinsbourne and Hiscock (1983) clearly indicates that whether the two tasks interfere depends very much on two factors: the difficulties and the nature of the motor tasks. It seems that the more difficult and the more similar the tasks, the stronger the interference. Speaking is generally reported to affect (to slow down) the tempo of arm or finger movements. However, the study of Inhoff and Bisiacchi (1990) demonstrates that speaking significantly constrains the timing of concurrent unimanual tapping, i.e. it reduces the tapping variability.
As both manual motor activity and speaking are controlled by the left hemisphere (Kimura 1993; Mattingly and Studdert-Kennedy 1991), several studies have been performed to test whether non-verbal vocalization, supposedly controlled by the right hemisphere, also interferes with manual movements. Results are not unanimous: Lomas and Kimura (1976) as well as Johnson and Kozma (1977) did not find any influence of nonverbal vocalization on concurrent movements with either hand, whereas an effect was reported by Hicks (1975) and Hicks et al. (1978). An interesting question arising from these studies concerns the type of interactions between manual movements and vocalization that occurs if vocalization implies the involvement of both hemispheres, as is the case in song performance.

In the following examples our analyses focus on coordination, and the degree of synchronization between singing and manual performance of a rhythm by individual musicians from two different cultures. The two cases presented here are a continuation of the analyses from case study 1. In addition to playing a rhythm instrument, both performers were also simultaneously singing in both performances. The data for the manual activity (playing the rhythm instrument) have already been presented (see case study 1). Singing, the vocal motor activity, is analysed here in terms of timing of speech segments, i.e. the onset of consonants and vowels. As both instrumental rhythm and singing were part of the same recordings, they share a common time base and the chronometric data of both activities can be compared directly. The segment onsets were marked with a computer program that allows for simultaneous display of the original sound wave as well as the spectral representation of the sung voice (see fig.7), and the markings, together with the associated time information was combined with the data for the manual rhythm and saved as an ASCII file for further treatment.

**Fig.7:** Example of sound wave (upper tier), spectral display (middle tier, with curves showing intensity (upper curve) and pitch (lower curve)), with segment transcription, segment labels and segment onset markings (vertical lines) in the lower tier. The example is from a section of a spoken song text.

**Case A: Kpegsisu (Ghana/Africa)**

For every segment onset a special program function calculated the distance to the closest bell stroke. These measurements are plotted against the time of the bell strokes and split by segment type (c = consonants, v = vowels) in fig.8. Obviously the segments are aligned with the bell strokes, graphically represented by their arrangement in parallel to the zero line (the bell stroke reference), with c and v segments having different distances to the bell strokes. The vowel onsets tend to be placed on or close to the bell strokes (mean distance 0.009 sec), whereas there are two populations of c segments, a large,
In time with the music

‘early’ group with a mean distance of about -0.076 and smaller, ‘late’ group with a mean of 0.146 sec. This seems to indicate that the vowel onsets are synchronized with the bell strokes.

Fig.8: Plot of segment distances to the nearest bell strokes in a Kpegisu song.

The absolute differences, however, do not inform us about the synchronization pattern. To find this, we need to express the difference between the two variables (i.e. timing of bell strokes and c/v onsets) with respect to the time cycle or period of one of the variables. This measure is called the relative phase, and can be calculated, for example, by dividing 360 degrees by the time between two beats and multiplying the result by the difference between the first of the two beats and the c/v onsets. If this calculation is done for all bell stroke pairs and their respective c/v onsets (see Appendix C) we can plot the data and obtain the distribution pattern for the vowel and consonant phases as shown in fig.9. The relative phase for the vowels (v phase) has a major peak close to 0 degrees. The actual peak is about 8 degrees off, and this difference is part of the ‘groove’ of the music. Another characteristic is the minor peak close to 180 degrees, a synchronization of vowels in antiphase to the bell stroke corresponding to the off-beat feel of this music.

Fig.9: Relative phase of bell strokes and vowel (v phase) and consonants (c phase) in a Kpegisu song

The relative phase for the consonants (c phase) shows a flatter distribution and a broader peak at about -60 degree. The consonant onsets are neither in phase nor antiphase with the bell strokes, but there appears to exist a timing relationship with the vowel onsets. Indeed there is a highly significant, strong correlation (R=0.967, p<0.0001) between the c segments and the corresponding v segments. This correlation strongly suggests that the stream of sung language can be characterized in the same way as the speech chain in spoken language, as a continuous vowel flow probably represented by the alternate movement of jaw openings and closings (Rhardisse and Abry 1995) on which consonant gestures are
superimposed (Fujimura 1995; Öhmann 1996). It seems that in this music the timing of the vowel flow is such that vowel onsets are synchronized with the timing of the bell strokes.

Case B: Gama AT (Queensland/Australia)

Again, following the labelling of segment onsets, a special program function calculated the distance between segment onsets and the closest clap stick beat. These measurements are plotted against the time of the clap stick beat and split by segment type (c=consonants, v=vowels) in fig.10. In contrast to the Kpegisu song, there seems to be no obvious alignment between c or v segments and clap sticks.

![Fig.10: Plot of segment distances to the nearest clap stick beats in Gama AT.](image)

This is confirmed if we examine the relative phase plots for the clap stick beats and c/v onsets in fig.11. The relative phase for both consonants and vowels is rather flat, with the vowels showing a moderate peak at +90 degrees and hence a slight preference to occur after the beat.

![Fig.11: Relative phase of bell strokes and consonants (c phase) and vowels (v phase) in a Gama song](image)

Despite the lack of synchronization between clap sticks and c/v segments there is again a high correlation between v and c segments (with a correlation coefficient of $R= 0.845$ and $p<0.0001$). It seems therefore that the temporal relationship between consonant and vowel segments is independent from whether or not one of these segment types is synchronized with another activity. This supports our
suggestion following from the previous analysis that segmental time structure in song language can be understood in analogy to that of spoken language.

Discussion

Although entrainment has not been proven in either of the two cases - that can only be done by analysing the dynamics of the interaction at a point of perturbation or rate change — their comparison demonstrates that there is no automatic synchronization or self-entrainment between simultaneous manual and vocal activities in one body. It is possible, as in the coordination of speaking and arm or hand movements, that synchronization is more likely to occur with the more difficult tasks, as in the present examples the Kpegisu pattern is evidently more demanding than the Gama clap stick pattern. This hypothesis could easily be tested by analysing a series of comparable songs in which the rhythmic accompaniments cover a broad range from simple to complex forms. However, the complexity of the accompanying rhythm does not seem to be the only factor that influences synchronization. Although all Dyirbal Gama songs have essentially the same clap stick pattern, they actually show a variety of clap stick-singing interaction pattern that range from non-synchronization, like in Gama AT, to significant synchronization (Will, forthcoming) and the degree of synchronization seems to be a ‘trade mark’ of individual performers, just as it seems to be the case for melodies in this culture (Dixon and Koch, 1996). This suggests that the degree of coherence in synchronization or entrainment is also affected by musical and cultural factors whose influence needs to be determined and analysed. For example, in a recent study Tom Beardslee (unpublished) was able to show how the introduction of electronic rhythm machines and sequencers affected the synchronization between singers and rhythm section in West African High Life music.

In the present case study we were able to quantitatively describe the different degrees of synchronization between the two concurrent activities of beating/clapping and singing. With the given examples it was not possible, however, to show how the two activities influence each other. However, there are other musical genres that would allow for such an analysis. For instance, in North Australian music there are genres that consist of sections with clap stick beating only, sections with clap sticks and concurrent singing, and sections with singing only. In these genres it should in principle be possible to analyse whether clap stick beating affects singing or vice versa and, finally, to describe by what processes the synchronization is established. For Central Australian music, one of us found that when songs are accompanied by beating, there is a noticeable entrainment effect on the singing with beat and syllable alignment and regularization of the sung text rhythms (Will, in press). Nevertheless, the experiments of Inhoff and Bisiacchi (1990) suggest that, at least in principle, an inverse entrainment – that of the beating accompaniment by singing - might also be possible.

Case Study 3: Coordination of rhythms between two performers

While the previous case study concerned processes within individual musicians, we will now take a look at synchronization and entrainment between two performers. This case study will also demonstrate that analyses of recorded performances are not limited to the mere identification of the degree of synchronization. As already indicated in section 6, in order to be able to analyse the process of synchronization in non-experimental studies, we have to look at non-stationary sections (‘perturbations’) in the performance, for example tempo or pattern changes. Only by analysing the behaviour of the performers at these transitory sections are we able to identify the presence (or absence) of entrainment, that is the re-establishment of synchronization, and to describe the dynamics of the underlying processes. The example for this case study was chosen because it contains several non-stationary sections – an initial phase of coordination and three pattern changes – within the short
In time with the music 46

time span of about 35 seconds. It also demonstrates that in music we are rarely, if at all, dealing with strictly periodic oscillations, but exact periodicity has been shown to be not a prerequisite for entrainment (e.g. Pikovsky et al. 2001). In addition we want to show how to describe the interaction between the two rhythm performances in terms of their relative phase, an important measure in the behavioural analysis of complex systems. For the kind of data we are dealing with here, we can construct a point estimate of the relative phase by calculating the latency (time of occurrence) of one series of clap stick beats with respect to the period (duration between two successive beats) of the other (see Appendix C). Such a description is also accessible for musicologists, because 'relative phase' can immediately be translated into 'musical' terms: if the reference period (the pulse) is represented as a quarter note, a phase difference of 180 degree corresponds to a binary subdivision of the reference period (e.g. 8\textsuperscript{th} note), a 90 degree phase difference means the timing of the players differs by a 16\textsuperscript{th} note, etc.

The example for this case study is taken from the first track of the Djambidj recording, a clan song series from Arnhem Land, published by Clunies Ross and Wild (1982, 1984). The performance includes two singers who also play clap sticks, and a didgeridoo player. Djambidj songs have been described to consist generally of three parts distinguished clearly by the accompaniment: part one is accompanied by clap sticks, part two by clap sticks and didgeridoo and part three is unaccompanied (Clunies Ross and Wild 1982). The clap stick beating of the first Djambidj song consists of several sections (see fig.12). Following some initial beats at 'half speed', there is a long section of more or less regular clap stick beats with average durations of 0.26 sec (see also fig.13). At the 22 sec time mark the performers change the pattern, then they return to the first pattern at the 28 sec mark and finally introduce a concluding pattern at about 33 sec. We are going to start this analysis by looking first at the clap stick performances in the first part up to the time mark of 22.6 sec. The different clap sticks of the two players were identified on the basis of their different sound spectra.

Fig.12: Clap stick beat durations of a Djambidj song played by two performers (circles and squares).

Fig.13: Temporal alignment between the clap stick beats of the two players (1 and 2) in the first 10 seconds of the Djambidj song from fig.12.
One of the singers starts with the beating and after his second beat the other player begins. If we compare the timing of the two players (fig.14) we see that the initial difference of 0.17 sec is rapidly reduced and within a couple of clap stick beats the performers achieve synchronization of their clap stick beats, i.e. entrainment takes place. The regression of the data (continuous line in fig.14) reflects this entrainment effect with an initial exponential reduction in timing differences and a subsequent plateau with values of about 10 milliseconds, which is reached at the time mark of 9.4 sec. As can be seen from both, fig.14 and fig.15 (relative phase plot), the initial entrainment proceeds in several steps. Following a quick reduction of the timing differences at the start of the performance, there is a short stabilization of the phase, a ‘locking in’ with a mean phase of about 40 degree for about nine beats (see fig.15). After the 8 second mark there is a change in phase and a further reduction of the differences to a mean phase difference of 9.6 degree and the performance between the 9th to the 22nd second settles down to a quasi-stationary section. The mean phase is also called a phase attractor: it seems to ‘attract’ the behaviour of the system in this part of the performance. Synchronization during this section is characterized by a peculiar pattern of differences in clap stick timing between the musicians: The phase shifts randomly, and there is only a weak, hardly significant cross correlation (r = 0.25, p = 0.053) between the musicians, indicating that the timing of the beats of one player has only a very weak influence on the timing of the other. Residual analysis suggests that it is actually the timing difference between the two players that is crucial here: these differences drift randomly within a band of ± 30 milliseconds and only when they exceed this value does there seems to be a directed correction. Furthermore, the autocorrelation function for the beat durations of the first player – the one who starts the song – shows a significant negative correlation at lag 1 and corresponds to the function as predicted by the Wing model, that means this player controls his beat sequence through self-correction. In contrast, the autocorrelation function for the second player has no significant values for any lag, suggesting that external factors influence the timing of his beatings.

In the following section, starting at 22.5 sec (first mark in fig.15) with the change to a different beat pattern, we find a break down of the previously achieved synchronization. The range of phase variation increases notably, even reaching a value of 71 degrees for the penultimate beat of this section. Although the phase is not stable there seems to be a certain patterning, a self-similarity of the phase changes with a period of three. As can be seen from fig.12, the players perform a three beat pattern with one long and two unequal short beats. The phase change pattern seems to indicate that the musicians do not synchronize individual beats. It seems that the reference point for their synchronization, for their ‘playing together’ is the starting point of the pattern, as there is minimum discrepancy at the beginning of each pattern (two of these points even have nearly zero phase), but increased phase differences within each pattern. The type of synchronization in this section can best be described as intermittent
phase locking where repeated beat patterns tend to be synchronized at their onset but not within the pattern.

At 28.3 sec (second mark in fig.15) there is another pattern change, the return to the first pattern. For the first beat of this pattern there is a phase difference of 60 degrees, but within two beats phase synchronization is re-established and phase variation is limited to the same narrow band as in the first section. The final pattern change, the introduction of the concluding formula at about 34 sec is not marked by a noticeable phase change and therefore will not be considered separately here. With the analysis of several such pattern changes it would be possible to arrive at a more precise description of the coupling force. In the present case the relaxation time, the time it takes a system to return to a previous state following a perturbation (an indicator of the coupling force), does not seem to be longer than one or two clap stick beats (a detailed study of Djambidj beating accompaniment is in preparation).

The mean phase for the whole last section (i.e. from 28.6 sec) is about 7.6 degrees and slightly smaller than in the first section (where it was 9.6 degrees). Interestingly the two sections with continuous phase locking seem to have a phase attractor that is not zero but has a marked positive value. With our two calculated phase means we might guess that the phase attractor probably lies somewhere between 7 and 10 degrees.

Discussion

Our analysis has shown that the synchronization between the clap stick beats of the two musicians is not a complete or absolute one, but one marked by fluctuations within specific limits. One obvious explanation for this ‘synchronization bandwidth’ of approximately 30 milliseconds comes from psychophysics: It has been shown that this time is necessary to evaluate the temporal-spatial order of acoustic stimuli (for review see Pöppel 1997). If the time difference between two musicians is considerably less than 30 milliseconds then, although they can still hear that the sounds they produce are not absolutely synchronous, they cannot tell which of the two is first and therefore cannot take directed action to improve synchrony. Only when the difference between them is larger than this interval is it possible for a musician to take explicit action in order to reduce the difference. Interestingly, the mean absolute difference between bell strokes and the ‘synchronized’ vowel onsets in the Kpegisu song of case study 2 was also of this magnitude, i.e. slightly larger than 30 milliseconds (see fig.8). This interval is a reflection of some basic neuronal mechanisms, as is suggested by the fact that attentive
brain states in which incoming stimuli are evaluated, produce EEG patterns with a strong 30 Hz component indicating heightened, synchronized cortical activity in this frequency range.

However, our results are not meant to suggest that synchronization cannot be better than about 30 milliseconds. It is well known that skilled performers can overcome this limitation and improve their synchronization considerably. The point here is that this improvement can only be done with recourse to other means than an evaluation of the time structure of the produced auditory events, for example by using anticipatory action. Anticipation studies seem to suggest that the superior reactions of highly skilled performers in some cultures are due largely to learning. Although players are generally not conscious of the cues they are responding to in anticipation, careful analyses of performance contexts and detailed experimentation might unveil at least some of the cues guiding the performers.

The finding that the attractor, the mean relative phase between the musicians, is located between 7 and 10 degrees, is interesting in connection with modelling of coupled oscillators in music psychology (e.g. Large and Jones 1999). There, it is simply assumed that the phase attractor is located at phase zero. However, the instabilities of that location due to psychophysical constraints (see above), make it possible that the actual location of the phase attractor in non-laboratory, real world situations can easily be influenced by socio-cultural and/or performance conditions. We are going to argue that the position of the phase attractor in this case reflects the specific socio-cultural condition of the entrainment process under consideration, an entrainment between a ‘song owner’ and an ‘accompanying’ performance participant.

Ethnographical background knowledge is essential in interpreting the analysis and understanding the socio-cultural significance of the identified entrainment process. Generally, the ethnomusicologist who performs the analysis is also the one who did the recording and can, therefore, identify the musicians. This is not the case for the present example and despite the excellent booklet accompanying the record we cannot tell from the recording, which of the musicians is playing which pair of clap sticks. However, our analysis seems to suggest that those clap sticks that show a negative autocorrelation at lag 1 are played by the ‘song leader’: it is he who starts first, his performance is strongly controlled by his internal pulse, and his beat durations show the smaller variance of the two. The performance of the other musician, which exhibits a larger variance in his timing and less reliance on an internal pulse – as indicated by the autocorrelation function - seems to indicate that we are dealing here with a case of asymmetric entrainment: while the first player ‘sets the pace’, it is the performance of the second that establishes and assures the synchronization between the musicians. The processes that are indicated by our analysis, led to a degree of synchronization that satisfies the culturally defined criteria for ‘playing together’ of the two musicians with different roles in the performance. Different sets of criteria in other cultures might necessitate the recruitment of additional or different processes in order to establish the respective degrees of synchronization.

Musicians’ brains obviously cannot be thought of as just a memory store for motor routines - blind, inanimate warehouses which produce stereotyped responses when triggered by conscious commands or external stimuli. Analyses of entrainment processes produce abundant evidence to show that there is considerable leeway for - if sub-conscious – intelligence and creativity that mark entrainment as an important adaptive capability of living systems to be found whenever coordination and cooperation within a group, not the achievements of individuals, are of importance.
A. The autocorrelation function

Autocorrelation describes the correlation of a signal or a variable with shifted versions of itself. Take for instance the variable represented by the series of connected dots in figure 16. If we shift all values by one data point we obtain the sequence of open triangles in the upper graph. This shift by one data point is called lag 1. Comparing dots and triangles at the same points on the abscissa, we see that the values of the two series change in opposite directions: when point values increase, triangle values decrease, and vice versa. This is the reason why the correlation coefficient for these two data series is negative.

Fig.16: Plot of a variable (dots) and two lagged versions of the same variable, shifted by one data point (lag 1, triangles, upper graph) and by two data points (lag 2, x, lower graph). Numbers on the right show the correlation coefficients for the correlation between the variable and lag 1 (upper number) and lag 2 (lower number).

Shifting the lag 1 series by another data point we obtain the lag 2 sequence (lower graph). In this case the data values of the original and the lag 2 series vary in the same direction, if one increases, so does the other, and the correlation between them is positive.

The autocorrelation function produces a series of numbers that are the correlation coefficients for the variable and a specified number of lagged versions thereof. This function allows us to detect whether values of a variable are dependent on previous values of the same variable, for example whether the duration of beats or tones is influenced by the duration of preceding ones. As demonstrated by case study 1, it also permits us to identify (e.g. temporal) patterns hidden in the data series: these would show up as repetitive pattern in the autocorrelation function.
B. Sequential and hierarchical timer models

Timed complex behaviour seems to require the activation of multiple timers. Even as simple a movement as a drum stroke, for example, needs time-coordinated activation of a large set of different muscles. One of the interesting questions is whether timing of musical behaviour, e.g. the production of a rhythmic pattern, can adequately be accounted for by only a sequential organization of temporal behaviour or whether it requires a hierarchical organization. The basic idea behind these different models is as follows.

According to the proposal of Wing and Kristofferson (1973a,b) the timing structure of motor behaviour can be considered to consist of two components: a) temporal structures produced by timers (e.g. neural oscillators) and b) temporal structures produced during the execution process (implementation of timer ‘instructions’ in the muscular system involved). This second component, the ‘executional’ part, is assumed to be the same in both sequential and hierarchical models. However, the difference between the two models lies in the arrangement of the timers.

We will examine these arrangements with an example of a pattern of four movements or strokes of unequal duration. In the sequential model (fig.17) each timer starts a motor action and initiates the activation of a subsequent timer. The timer for the last motor action in the series (T4) re-activates the first timer (T1) and then the cycle starts all over again.

![Fig.17: Sequential timer model for four timers activated in a cyclical fashion. The model has two components of variance, variance due to the imprecision of the timer and variance produced at the level of the motor delay (e.g. imprecision caused by the activation of the muscular system).](image)

In the hierarchical model (fig.18) the first timer (T1) starts a motor action and initiates the activation of the subsequent timer (T2). The second timer starts a motor action and initiates the activation of timer 3 and 4. In contrast to the sequential model, the subsequent cycle is not initiated by the last timer in the sequence (T4) but by the first timer (T1). This ‘high-level’ timer T1 controls the duration of the whole cycle. The model shown in fig.18 is but one possible arrangement; with four timers 24 different arrangements are possible.
Vorberg and Hambuch (1977) have shown that – at least in principle - a sequential timer organisation can be distinguished from a hierarchical organisation on the basis of the time structure of the resulting motor actions. In our case studies we make use of this to distinguish between sequential and hierarchical arrangements, a distinction that seems to be relevant for understanding and interpreting musical entrainment processes. The distinction can be made, as demonstrated by Vorberg and Hambuch (1977), by comparing the variance of the inter-response intervals summed up within a cycle and across cycles. The summed interval within a cycle is $I_p = I_{p}(t_1,t_2) + I_{p}(t_2,t_3) + I_{p}(t_3,t_4) + I_{p}(t_4,t_1)$ and a sum across cycles is e.g. $I_s = I_{s}(t_3,t_4) + I_{s}(t_4,t_1) + I_{s}(t_1,t_2) + I_{s}(t_2,t_3)$, where $I_{n}(t_m,t_{m+1})$ is the interval of the $n$-th cycle, between the actions initiated by timer $T_m$ and the subsequent timer. With a sequential timer arrangement the variance of the sum of the response intervals should not depend on how they are chosen, the variance of $I_p$ and $I_s$ should be the same. Under the hierarchical assumption the variance of $I_p$ should be smaller than the variance for $I_s$, because in the variances from two cycles add up in the latter case.
C. The relative phase of two event series

Figure 19 shows how the relative phase between two periodic events series can be calculated. The figure shows waveform and spectrogram of an excerpt from a recording of singing with clap stick beating. There are two pair of clap sticks (played by different performers) that can easily be discerned by their spectral components: one pair shows a lower and higher frequency component (arrows to the right of the last pair of beats) that are missing in the other pair.

![Sound wave and spectrogram](image)

Fig.19: Sound wave (upper display) and spectrogram (lower display) section from a recorded performance of singing with clap stick accompaniment (two pairs). For explanation see text.

To calculate the relative phase we need the period of one clap stick pair, for example clap stick one, which is $t_{32} - t_{31}$ in the above example. Then we need to calculate the latency of the second clap stick pair, which is $l = T_{32} - t_{31}$. The relative phase can then be calculated as $\phi = l \times 360/(t_{32} - t_{31})$. This procedure is repeated for the next period $t_{33} - t_{32}$ and all subsequent ones.
9. References


Waterman, C.A (1995) [Response to Participatory Discrepancies Special Issue], *Ethnomusicology* 39/1, 92-94.


**Notes on the authors**

**Martin Clayton** studied at the School of Oriental and African Studies (SOAS) in London, where he completed a BA in Music and Hindi (1988) and a PhD in Ethnomusicology (1993). He has worked in the Music Department of the Open University since 1995, and has also taught at several other institutions including the universities of London, Cambridge, and Chicago. He served as editor of the *British Journal of Ethnomusicology* from 1998-2001 (with Suzel Reily), and is currently Chair of the British Forum for Ethnomusicology.

Email: m.r.l.clayton@open.ac.uk

Web pages:  [www.open.ac.uk/arts/music/mclayton.htm](http://www.open.ac.uk/arts/music/mclayton.htm)

**Rebecca Sager** holds a MM (1993) and PhD (2002) in Ethnomusicology from the University of Texas at Austin, receiving research funding from the Wenner-Gren Foundation for Anthropological Research. She has taught in the Department of Music at Istanbul Bilgi University and currently lives at 1017 Cherokee Dr., Tallahassee, FL 32301, U.S.A.

phone: +1 (850) 219-9966

Email: rsager@music.org

**Udo Will** studied music, sociology, and neuroscience. He holds PhDs in both musicology and neurobiology and is professor of cognitive ethnomusicology at The Ohio State University. He currently leads research projects on cognitive aspects of music performances in oral cultures, on rhythm and melody processing by the human brain, and on a comparison of the cognitive architecture of music and language production.

Email: will.51@osu.edu

Web pages: [www.arts/ohio-state.edu/Music/Ethnomus/Ethno.html](http://www.arts/ohio-state.edu/Music/Ethnomus/Ethno.html)