What is Entrainment?
Definition and applications in musical research

MARTIN CLAYTON
Durham University, Department of Music

ABSTRACT: Entrainment theory describes the process of interaction between independent rhythmical processes. This paper defines entrainment in this general sense, then briefly explores its significance for human behaviour, and for music-making in particular. The final section outlines a research method suitable for studies of entrainment in inter-personal coordination, and with reference to published studies suggests that the study of musical entrainment can be a source of rich insight also for the study of human social interactions and their meanings.

Submitted: 2012 January 6; accepted 2012 July 13

KEYWORDS: dynamic attending, neuronal oscillators, recursiveness, interaction

The term ‘entrainment’ refers to the process by which independent rhythmical systems interact with each other. ‘Independent rhythmical systems’ can be of many types: what they have in common is some form of oscillatory activity (usually periodic or quasi-periodic in nature); they must be independent in the sense of ‘self-sustaining’, i.e. able to be sustained whether or not they are entrained to other rhythmical systems (thus sympathetic vibration, as when a violin’s soundboard vibrates at the same frequency as one of its strings, is not an example of entrainment). In order for interaction to take place some form of coupling must exist between the rhythmical systems, and this too can take many forms. This process of interaction may result in those systems synchronising, in the most common sense of aligning in both phase and period, but in fact entrainment can lead to a wide variety of behaviours.

The classic example of entrainment is that of pendulum clocks, which were observed by the Dutch physicist Christiaan Huygens to synchronise when placed on or suspended from a common support (see Pikovsky et al. 2001, 357ff). Numerous other mechanical instances of this phenomenon have been identified, but the phenomenon also extends to the biological world, where examples include those of synchronising fireflies, and the resetting of body clocks by sunlight (circadian entrainment). Even these few examples reveal some of the many variables encountered in the study of entrainment. Huygens’s original example was of two clocks mutually influencing each other (symmetrical entrainment); the fireflies are many, with each individual potentially both influencing and being influenced by several others; circadian entrainment features one rhythmic process – the cycle of day and night – influencing the internal body clocks of many individual organisms, while those individuals are not able to influence the time of the sun’s rising and setting (asymmetrical entrainment). Entrainment between rhythmic processes may thus be one to one, one to many, or many to many, and may be symmetrical or asymmetrical. This paper will focus on examples of entrainment that are particularly relevant to human behaviour, and more narrowly still to musical behaviour. Even within this relatively narrow brief, entrainment comes in many forms, in terms of the number of rhythms involved, their timescales (periods), (a)symmetry, and so forth.

Entrainment is not a single phenomenon that occurs only in human musical behaviour: it is an abstraction describing a process common to many different phenomena occurring at different scales of space, and in both biological and mechanical systems. As such it has been studied with the aid not only of observation of natural phenomena, but also with the help of mathematical models developed within the broader framework of dynamical systems theory. It is possible to study musical entrainment productively with reference to only a tiny fraction of this substantial research field, but nonetheless a certain part of this wider body of theory is indispensable. One essential abstraction is the notion of phase in periodic or quasi-periodic processes. Many observable rhythmical processes are by their nature continuous, but the more or less arbitrary definition of a reference point (such as the audible tick or a clock or the moment at which an insect begins to emit light) allows us to study the phase relationships between them. If we are concerned with entrainment between musicians, we might identify the moment an individual strikes a drum-head or taps a foot as the focal point of a quasi-periodic movement, and based on this choice we may study the relative phase of a particular pairing. If
two such events occur at precisely the same time then the rhythms are in phase (relative phase 0°), if one occurs precisely midway between the other they are in anti phase (relative phase 180°), and so on. If two rhythms are entrained, then, they do not necessarily fall precisely into phase with one another. Rather, the evidence for entrainment will be (a) a stabilisation of the relative phase relationship, and (b) the reassertion of this stability following a perturbation. In other words, when two rhythms are entrained the relationship between them (be it in-phase, anti-phase, or somewhere in between) will stabilise, and will be sufficiently robust to reassert itself; in the example of the clocks the test of entrainment is not simply that they tick in synchrony (which could happen by chance), but that if we upset that synchrony, for instance by momentarily stopping one of the pendula, they will resynchronise.

**ENTRAINMENT IN HUMAN BEHAVIOUR**

An example of human entrainment has already been noted above – the resetting of internal body clocks by sunlight. In fact, many instances of entrainment can be observed both within and between human individuals. Examples on a very small scale include entrainment between different neuronal oscillators, or between ‘pacemaker’ cells found in the heart. Still within the individual but on a somewhat larger scale, many common actions involve the synchronisation of movements between different body parts – for example walking. At the inter-individual level, many forms of manual labour clearly involve the synchronisation of actions between individuals – the cooperative pounding of roots or grains, for example, or a cooperative sawing motion. High-performance athletes may benefit from types of entrainment not experienced by most individuals: for instance, rowers and cross-country skiers tend to entrain their breathing to their limb movements (Steinacker, Both et al., 1993; Fabre, Perrey et al., 2007). In short, entrainment is displayed in many different forms of human behaviour, with periods ranging from a few milliseconds to a day (and possibly longer). That observed in inter-personal entrainment tends to fall within a somewhat narrower temporal range.

Looking at this from a somewhat different perspective, entrainment affects our behaviour in many ways: if our body clock falls out of synchrony with the cycle of day and night we suffer from jet-lag, which can result in unpredictable patterns of tiredness and wakefulness, irritability and so on. We will suffer in very different ways if our heart pacemaker cells do not function properly. The broader implications of entrainment for human behaviour have been discussed across a range of disciplines including social psychology and even history. Historian William McNeill, for instance, begins his 1995 book with reflections on the visceral effects of military drills and what he calls ‘muscular bonding’, extending his discussion to topics including dance, religious ceremonies and social cohesion in small communities. In the process he tackles many questions which have interested social scientists at least since Durkheim, who argued that:

“Probably because a collective emotion cannot be expressed collectively without some order that permits harmony and unison of movement… gestures and cries tend to fall into rhythm and regularity, and from there into songs and dances.” (Durkheim 1995 [1912]: 218)

One of the important aspects of human entrainment to have been studied to date is inter-limb coordination: a long tradition of tapping coordination studies has shed considerable light on this phenomenon. The basic dynamics of two independent but coupled body parts (published studies focus most frequently, but not exclusively, on index fingers) can be modelled with the help of a simple mathematical equation known as the Haken-Kelso-Bunz (HKB) model, which simply assumes the existence of two dynamical systems and some form of coupling between them (it is not therefore dependent on any specific biomechanical information): this model predicts two stable relationships, in-phase and anti-phase, with the anti-phase relationship becoming progressively less stable as the frequency of oscillation increases (Kelso, 1995: 54ff). This is exactly what empirical studies of fingertapping show, whether the two index fingers belong to one individual or two.

A related group of studies explore unintended entrainment between two individuals performing actions such as swinging pendula from the wrist or rocking in rocking chairs (Schmidt & O’Brien, 1997; Richardson et al., 2005; Richardson et al., 2007). These experiments clearly show that when individuals interact socially, for example in conversation, the rhythms of their actions tend to become entrained. A key factor here is mutual attention in social interaction, because simply being in the same room is not a sufficient condition for entrainment to occur.

Indeed, according to Mari Riess Jones’ Dynamic Attending Theory, attention is a key factor in human social entrainment. Jones proposes that attentional behaviour is quasi-periodic, and that our attentional rhythms may become entrained to regularities in our environment, which may include rhythmical behaviours of other individuals (Jones & Boltz, 1989; Large & Jones, 1999). The entrainment of attentional rhythms can be understood as underpinning a number of human social behaviours, including speech and music: if I can entrain to your behaviour and you to mine then we are able to coordinate our actions. Numerous examples of human behaviours – notably musical ensemble –
suggest that some mechanism of this kind must exist, while for others – such as turn-taking in conversation and other aspects of linguistic behaviour – it may also be a convincing explanation.

A particular phenomenon which seems to be distinctive of entrainment in and between humans, which is related to the periodicity of attention, is our ability to coordinate actions with an external periodic auditory stimulus. The notion that this is exclusive to humans has been challenged by Patel: exploring his hypothesis that beat perception is linked to vocal learning (2006), studies appear to show that a cockatoo is able to synchronise its movements with a musical stimulus, if intermittently (Patel et al., 2009). Whether or not this proves to be a widespread ability in other vocal-learning species, it certainly seems to be fundamental to many behaviours that we think of as essentially human, such as the use of language and music. Dynamic attending, especially when employing auditory information, permits a wide range of temporally coordinated behaviours in humans. Timing coordination can be observed in many different species, but the range of applications of inter-personal entrainment seems to be uniquely broad in humans, precisely because we can attend rhythmically to a range of senses including hearing.

ENTRAINMENT IN MUSIC – DISTINCTIVE FEATURES

As noted above, entrainment theory is not domain specific, but is rather an abstraction that can be used to make sense of many different phenomena. In studying entrainment in music too we can find many different phenomena sharing some common features. In fact, I argue that it is important to distinguish between different manifestations of the phenomenon (see Clayton, in press), and that it is convenient to do so at three different levels, viz:

1. Intra-individual entrainment takes place within a particular human being. An important phenomenon at this level is the entrainment of networks of neuronal oscillators, which appear to be responsible for metrical perception (Large & Kolen, 1994; Large, 2000, 2010; London, 2004). Another aspect of intra-individual entrainment, as noted above, is the coordination between individual body parts (e.g. the limbs of a drummer).

2. Inter-individual/ intra-group entrainment concerns co-ordination between the actions of individuals in a group, which is essential for ensemble playing in any musical tradition. This is largely facilitated by the entrainment of attentional rhythms to auditory information, although other sense modalities – vision in particular – often also play a part.

3. Inter-group entrainment concerns the coordination between different groups. This is less widely recognised, being a rare phenomenon in Western art music, but in fact it is a widespread phenomenon (see Lucas et al., 2011).

These different levels of musical entrainment are interdependent, most obviously in the sense that each builds on the previous level: intra-personal entrainment allows us to perceive metrical structures in musical stimuli and to coordinate our actions to those structures; without this it would not be possible for individuals to play in time with each other; and without musicians playing together in groups we could not have multiple groups interacting with each other. The interdependence may not all be in one direction, however. Clayton (2007) demonstrates how hierarchical timing relationships can come into being without prior planning or explicit recognition, as an emergent behaviour of a group of people performing quasi-periodic actions: this raises the possibility of metrical patterns emerging directly from joint action, rather than necessarily coming into existence first at the neuronal level and then being expressed behaviourally.

As noted above, entrainment does not necessarily result in synchronisation in phase between rhythms of matching periods. Musical behaviour offers many other manifestations of entrainment, some of which may even be unique to human music-making.

A. Different rhythms can entrain not only in unison, but also in hierarchical or polyrhythmic relationships. Examples of coordination between musical parts which are relatively slow and those which are relatively fast (in relationships such as 2:1, 4:2:1, or 6:3:1) are so common in music as to be trivial. Less common but still very widespread are polyrhythmic relationships between parts (3:2, 4:3); Clayton (2007) introduces an example of a 3:2 relationship which emerges unintentionally from musical interaction.

B. Hierarchical relationships can not only be observed behaviourally, as in the case of parts which move at different speeds but are mutually coordinated. They also account for metrical percepts: computer models which aim to illustrate this process in a simplified form, show how hierarchical percepts can emerge spontaneously in response to relatively simple stimuli (Large 2010).

C. Just as common as musical parts in hierarchical temporal relationships are parts with matching periods which are synchronised out of phase – for example, a snare drum that falls in an anti-phase relationship with a bass drum. While so many commonly-cited examples of entrainment seem to involve in-phase relationships, it is important to keep in mind that in music, entrainment can involve a wide range of phase relationships.
D. Entrainment can be symmetrical (as in Huygens’s clocks) or asymmetrical (as in circadian rhythms). In the case of music it can be either: symmetrical in an ensemble made up of peers, asymmetrical when people play or dance along with pre-recorded music they cannot influence. It can also be relatively symmetrical: in most musical ensembles any individual can influence any other, but in practice some people are more likely to have influence than others (e.g. conductors, section leaders, soloists, senior musicians). Music may then be a particularly good forum for investigating the interdependence between timing coordination and social power relationships.

In summary, musical entrainment is recursive: individuals perceive and generate hierarchical temporal structures; they coordinate their actions within groups; and groups of people coordinate to form larger groups. It is also diverse: it can involve matching periods as well as hierarchical and polyrhythmic relationships, it is out of phase as often as it is in phase; and it can fall almost anywhere on the symmetrical-asymmetrical continuum. Musical entrainment is observed with periods in a range of roughly 100-2000 msec (corresponding to frequencies of 0.5-10 Hz), from the fastest beat to a typical measure (metrical and hypermetrical structures can however be considerably longer than 2000 msec, see e.g. Clayton, 2000, p. 87).

Given the diverse examples of entrainment in musical performance, it follows that a wide range of methods may be applied in studying these phenomena. The particular focus of the work summarised below has been the role of entrainment in interpersonal and inter-group interactions, as manifested both in the sounds produced and in patterns of movement. The analyses are therefore sensitive both to musical knowledge and to personal and social factors – in other words they take account of the fact that entrainment is being observed in real-life, meaningful human activities, and assume that both the features of the music that people aim to produce, and the range of meanings that the activity holds for them, will be significant in relation to the entrainment dynamics per se.

**INVESTIGATING MUSICAL ENTRAINMENT: LESSONS FROM CASE STUDIES**

This final section briefly introduces an approach to the study of musical entrainment in natural settings, which employs a ‘stroboscopic’ method designed for the investigation of entrainment between quasi-periodic rhythms (Pikovsky et al., 2001, Clayton et al., 2005). This method is straightforward in principle, involving the following steps:

a) Identification of quasi-periodic rhythms and extraction of time series data

b) Calculation of relative phase relationships from pairs of time series

c) Investigation of entrainment using this relative phase data, employing visual inspection and statistical measures

The case studies referred to here all concentrate on either inter-individual or inter-group entrainment. The rhythmic processes at stake can be either sound-producing or silent movements: timings can be taken from the onsets of particular sounds (in an audio file), or identifiable points in a periodic motion such as the moment a drumstick strikes a drum head, or the highest or lowest point in a foot-tapping or head-nodding movement (from video recordings – motion capture technology would be another source of this data). Whichever rhythm is at stake, and whatever the data source, series of time data points need to be derived.

The next phase in this process involves the calculation of the relative phase of each point in one series with respect to the other series. If we are looking at the location of bass onsets with respect to a ride cymbal part in a jazz ensemble, for example, the two cymbal onsets closest to a given bass onset define a period, and the location of the bass onset with respect to that period is calculated. The relationship of the bass part to the cymbal part is thus expressed in a single series of phase angles.

When these phase angles are plotted against time, it is possible to get an immediate impression of the relationship between the two rhythms. If they are unentrained, that is uncoupled, then the relative phase plot will generally proceed as a gradual drift, showing up on the chart as a series of diagonal lines. If they are entrained then the relative phase will be stable, and the plot will form a (more or less) horizontal line.

Another way of plotting the same data is to simply plot the distribution of phase angles on a circle, in effect removing the time dimension. The grouping of data points will indicate the phase relationship to which the rhythms tend. If the distribution is unimodal, calculation of a mean vector will enable a quantification of this mean phase angle \[ \mu \] and the degree of entrainment \( r \), on a scale from 0 to 1. Multimodal distributions may indicate hierarchical or polyrhythmic relationships. Three stages of this process are illustrated in Figure 1: (a) a time series plot, (b) a plot of relative phase against time (in seconds) and (c) a phase distribution plot with mean vector.
Fig. 1. Investigating the relative phase of two musicians’ silent hand gestures (data selected from the study reported in Clayton, 2007, for clarity of illustration). a. Time series data. b. Relative phase of the harmonium player’s finger taps calculated with respect to the singer’s hand taps. c. Distribution plot of the data from chart b with mean vector ($r = 0.961$, $\mu = -8^\circ$). (Chart c and mean vector calculation produced in Oriana).

Three studies cited here apply this method in different ways (further studies using similar methods can be found in Clayton et al., 2005). Clayton (2007), from which the illustrative data in Figure 1 are derived, studies the relationships between musicians in an Indian classical ensemble. This paper considers two different aspects of an ostensibly unmetred performance: a) entrainment between the silent hand gestures of soloist and harmonium accompanist, and b) that between different players of the drone lute tanpura, who are not supposed to coordinate with each other or with the music they hear. Doffman (2008) uses similar methods to study the timing relationships between the members of jazz trios, and correlates these findings with interview material from the same musicians talking about...
timing relationships and their socio-musical significance. Lucas et al. (2011) applies the same methods to the study of relationships between different groups in the Afro-Brazilian Congado ritual, again seeking to make sense of the results in terms of participants’ musical and ritual relationships. All of these studies employ audiovisual recordings of real-life performances in natural settings. Of these three studies Clayton uses data derived from video observation, Doffman’s timing data were derived through a process of onset detection using audio files, while Lucas et al. derived timing data by tapping along to audio files (which were also contextualised with the help of video recordings). The different sources of the data affect the timing resolution to some extent, but otherwise are treated in very similar ways.

In the case of Clayton (2007), this approach demonstrated that in the case of the silent hand gestures, the two musicians were clearly coordinated – but loosely so, with \( r = 0.67 \) a much lower index of the strength of coupling than those reported in other studies. (As Figure 1 illustrates however, at times the coupling between the two is nonetheless much stronger than this, with \( r = 0.961 \) for this ten second extract). The study of the relationships between tanpura players reveal a mixed picture: sometimes they show no phase stabilisation, but in one pairing the two players show a (clearly unintentional) stabilisation, which occurs in a 3:2 relationship (the periods of the two plucking patterns are roughly 3 seconds and 2 seconds). In this case the stabilisation occurs between the soloist Veena Sahasrabuddhe and her student sitting behind her, and occurs when the student fixes her visual attention on the soloist’s back or shoulder: visual information seems to be important, and to this extent the entrainment must be asymmetrical as the soloist cannot see her accompanist. Less surprisingly perhaps, a study of the relationship between the soloist’s hand tapping and her own tanpura playing showed entrainment in a 3:1 relationship. As noted above, these findings illustrate not only that interpersonal musical entrainment can occur unintentionally, but that it can do so in hierarchical or polyrhythmic relationships, and can result in temporal hierarchies at least as complex as those intentionally reproduced as musical metre.

While the previous study was concerned largely with exploring the possibility of unintentional entrainment, Doffman (2008) looks at the musical and social coordinates of timing relationships between jazz players. Not surprisingly, the pairs of musicians (drummer, bassist, and guitar or piano soloist) are tightly entrained, with \( r \) typically >0.9. More interestingly, small nuances and shifts between relatively tight and loose coordination, or between a particular musician being slightly ahead or slightly behind another, can be intensely meaningful for these musicians, and these phenomena are tightly interwoven with musicians’ estimates of their own and others’ capabilities and characteristics as musicians, and with their understanding of the ideals to which jazz performers should aspire. Doffman studies the dynamic shifts between tighter and looser coordination, and concludes that timing relationships aimed for in jazz groove cannot be reduced to a single ideal (e.g. in phase relationship and degree of entrainment); rather, the ideal relationship is inherently dynamic and playing jazz involves meaningful variations within the permissible range of looseness and out-of-phasesensness.

Lucas et al. (2011) employ the same ‘stroboscopic’ method to investigate entrainment between groups in a form of ritual processional music. Here ethnographic study strongly suggests that the groups are invested in the notion that playing in time together is an index of ritual unity; concomitantly it is important not to fall into time with groups from other communities, which would indicate a breaking down of necessary ritual bonds and barriers. Study of several occasions on which different groups play in close proximity allowed a number of different factors to be distinguished. Different groups belonging to the same community entrain, and fall into synchrony, relatively easily – at least when their tempi are fairly close together and they are in close proximity. When the groups belong to different communities this is not so: sometimes they manage to retain their mutual independence, using strategies such as exaggerating tempo differences and looking away from each other; at other times one or both groups will simply stop playing to avert the possibility of falling into time with the other. On one occasion, however, two groups performed a mutual greeting ceremony while attempting to avoid playing in time, despite the fact that they were playing similar rhythms at similar tempi. The result was that they actually fell into a tightly entrained relationship for over two and a half minutes (\( r = 0.988 \)), but they managed to do so out of phase by 223°, which meant that although they were tightly entrained they did not perceive the relationship as such. Again, the complex interrelationship between entrainment dynamics, intentions and meanings is apparent.

These three studies, as varied as they are, all address a particular type of musical entrainment phenomenon – they are concerned with studying the interactions between people while making music in real-life situations. The various findings demonstrate not only that it is possible to shed light on entrainment dynamics in natural musical performances, but also that it is possible to relate these findings to information about the intentions, experiences and discourses of the people involved. This is a particular approach to a range of entrainment phenomena in music, which is interdisciplinary and committed to rigour in both quantitative and qualitative research methods and to a principled investigation of their interrelationship (see Clayton, in press). There can be little doubt that much more can be learned about human music-making, indeed about human interactions in general, through studies of this nature.
REFERENCES


