



Review

# A Dynamic Interactive Approach to Music Listening: The Role of Entrainment, Attunement and Resonance

Mark Reybrouck <sup>1,2</sup>

<sup>1</sup> Faculty of Arts, University of Leuven, 3000 Leuven, Belgium; mark.reybrouck@kuleuven.be

<sup>2</sup> Department of Art History, Musicology and Theatre Studies, Institute for Psychoacoustics and Electronic Music (IPEM), Ghent University, 9000 Ghent, Belgium

**Abstract:** This paper takes a dynamic interactive stance to music listening. It revolves around the focal concept of entrainment as an operational tool for the description of fine-grained dynamics between the music as an entraining stimulus and the listener as an entrained subject. Listeners, in this view, can be “entrained” by the sounds at several levels of processing, dependent on the degree of attunement and alignment of their attention. The concept of entrainment, however, is somewhat ill-defined, with distinct conceptual labels, such as external vs. mutual, symmetrical vs. asymmetrical, metrical vs. non-metrical, within-persons and between-person, and physical vs. cognitive entrainment. The boundaries between entrainment, resonance, and synchronization are also not always very clear. There is, as such, a need for a broadened approach to entrainment, taking as a starting point the concept of oscillators that interact with each other in a continuous and ongoing way, and relying on the theoretical framework of interaction dynamics and the concept of adaptation. Entrainment, in this broadened view, is seen as an adaptive process that accommodates to the music under the influence of both the attentional direction of the listener and the configurations of the sounding stimuli.

**Keywords:** music listening; entrainment; enactment; attunement; periodicity; oscillatory dynamics; interactionism



**Citation:** Reybrouck, M. A Dynamic Interactive Approach to Music Listening: The Role of Entrainment, Attunement and Resonance. *Multimodal Technol. Interact.* **2023**, *7*, 66. <https://doi.org/10.3390/mti7070066>

Academic Editor: Cristina Portalés Ricart

Received: 16 April 2023

Revised: 16 June 2023

Accepted: 24 June 2023

Published: 28 June 2023



**Copyright:** © 2023 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Listening to music can be an overwhelming experience. It can be so immense that it blurs the distinction between the listener and the music. The result is a kind of immersion in the sounds which has been described in terms of “oceanic feelings” and other poetic portrayals. Science, however, needs operational definitions and ways of measurement which make it possible to assess and even intervene in the experience. There is, further, a tension between the objective description of the acoustic features of the sounding music and the reactions and responses by the listeners, which can be described in an objective or subjective way. Listening, moreover, is not merely a process of information pickup, but a process of dynamic attending rather than passive resonance. Listeners, then, can be “entrained” by the sounds at several levels of processing, depending on their level of attunement and alignment of their attention. There is, as it were, a continuum from passive consumption of the music in a rather unarticulated way to a highly elaborate and fine-grained real-time attentional tracking of the time-varying sonorous articulation. What is needed, therefore, is an overall synthesis and comprehensive understanding of research areas that revolve around focal concepts, such as entrainment, attunement, and resonance.

A starting point to present such a coherent picture is the concept of *entrainment*, which has the potential to bridge the gap between the physical, mathematical, biological, neural, and social studies. Roughly defined as a process in which two or more autonomic oscillators interact with each other to adjust towards and lock into a common phase and/or periodicity, it opens the scope of study not merely to the encoding and decoding of information, but

also to the process of embodied interaction and tuning-in to musical stimuli. It allows musicians to synchronize their performance and to be in time with the music, or to put in other terms, to be “entrained” to some degree [1]. This applies to playing music, but also to joint actions, such as dancing, group sports, or gaming together, and it can even be extended to listeners who tap their feet or nod their heads along with the music. Such tuning-in immediately raises issues that rely crucially on prediction and anticipation—not merely reactive behavior—and the vast spectrum of complexity that is covered by human interactions and joint and/or coordinated actions [2,3].

Much attention has been given already to performer-related studies with a central focus on rhythmic entrainment (e.g., beat tracking) and related behavioral responses (e.g., tapping to the beat), besides studies on physiological entrainment (e.g., heart rate, respiration, brain waves, etc.) [4]. Questions can be raised, however, as to the breadth and scope of these approaches, which focus rather selectively on (mostly regular) rhythmic entrainment, with the exclusion of other potentially entraining factors such as timbre and pitch, which are both also temporal phenomena—their spectral content is constituted of vibrations, each having their distinct oscillating frequencies—if considered at the highest level of temporal resolution. There is, further, no reason to restrict entrainment studies merely to the investigation of regular and driving rhythms. A broadened view, on the contrary, can be seen as an additional tool for the description of fine-grained dynamics and perceived irregularities and discrepancies in the temporal unfolding that go beyond the standardizations, accuracies, and approximations of standard music notation. It is an approach that explores the validity and generalizing power of any theory that goes beyond the limitations of restricted and defined domains, such as jazz studies, classical music theory, or ethnomusicological research [5]. It argues also that musical behavior and concepts should be investigated more explicitly in terms of entrainment, as instigated already by some founding fathers of ethnomusicology in their attempts to cope with the idiosyncrasies of typical musics of the world, and to describe the sound as accurately as possible [6–8] (and see [9] for an overview).

In what follows, this broadened approach is elaborated in depth. Section 2 takes an “enactive stance” on musical sense-making by distinguishing between the acoustic description of the eliciting stimuli and the responses by the listener. The music is described as a driving or entraining force with specific characteristics, but equally important is the way how listeners attune their attention to the sounds. Section 3 delves into the focal concept of entrainment. Besides some conceptual and definitional issues, with a major distinction between physical, physiological, and cognitive entrainment, it starts from the metaphor of master and slave, exploring how listeners can behave as follower or leader by aligning their attentional control through small adaptations, both in terms of being driven or being themselves the driver. This delicate tension is then discussed in terms of oscillatory dynamics with a major distinction between external and internal oscillators, both of which can function as timekeepers for entrainment. Section 4, finally, opens up new avenues by bringing together the concept of enactive listening and dynamic interactionism. It highlights the role of the listener as an autonomous agent who enacts the heard music by relating epistemically and experientially to the sounds. By invoking recent conceptualizations of participatory sense-making and perceptual crossing, it points out that musical understanding is not merely a cognitive process. Equally important is the role of perceptual experience and interactions’ dynamics to understand music not only in a solipsistic manner, but also in terms of social awareness.

## 2. Music Listening: Enacting the Sounds

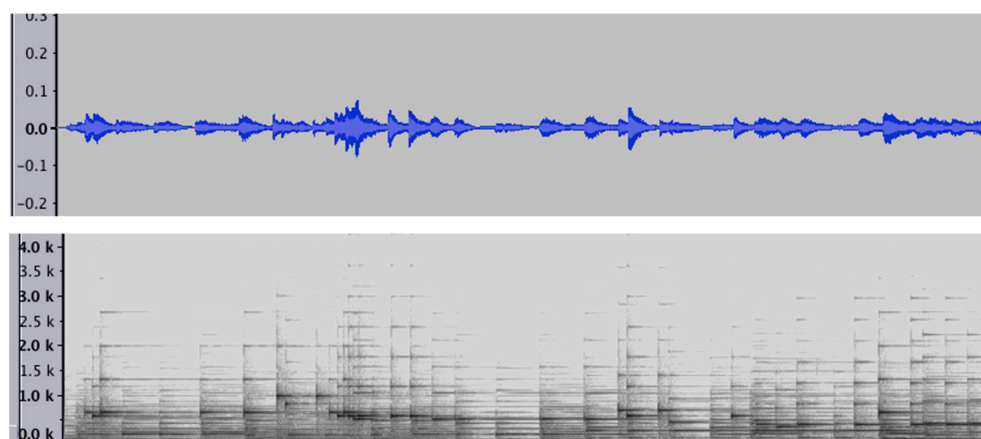
Music listening is a multifarious phenomenon. It is triggered by the acoustic features of the music, but even the most accurate physical description of the sounds cannot fully capture the subjective experience by the listener. There is, as such, no linear-causal relation between the physical attributes of the music and the feelings and meaning they evoke, given that the same musical stimuli can generate different experiences in other listeners or in the

same listeners over time and/or in different contexts [10]. Yet, it seems arguable to assume at least some non-arbitrariness between the music as a driving stimulus and the reactions and responses by the listener [11], and it is likely that the concepts of *entrainment* and *enactment* can be helpful here (see below for the terms). There are two levels of description in this regard: the acoustic description of the eliciting stimuli, and the responses by the listeners. Both levels can be measured “objectively” on the condition that the latter are measured in terms of behavioral (e.g., reaction time experiments, timing of task responses) or physiological and electrophysiological responses. The “subjective” experience, however, is more reluctant to objectification, although there are indirect measurement techniques that aim at uncovering the listener’s implicit reactions in a more objective and verifiable way—the phenomenon of physiological entrainment (see below) is a good candidate in this regard.

### 2.1. Eliciting Features of Entraining Stimuli

There is a strong tradition of describing music in terms of discrete categories or parameters with strong demarcations between them. Examples are pitch, timing, timbre, tuning, and loudness. It is arguable, however, to go beyond these category boundaries by conceiving of distinctions that exhibit more fine-grained nuances and that make it possible to describe the musical experience in more holistic terms [10,12]. Besides, it makes sense to conceive of music not only as a sounding structure, but as a “driving” or “forcing function” which invites the listener to seek what is invariant in the structure of the music. The “metrical structure” has proven to be a major candidate in this regard. It reflects the relative durations of the beats between them as contrasted with the tempo, which reflects the mere succession of absolute durations of the beats [13]. Care should be taken, however, not to rely solely on the concept of regular “beats”, given that not all music is percussive, and that most of the rhythms of music are even not strictly periodic. Most of them are complex, temporally structured sequences of acoustic events. There is, however, a widespread tendency to perceive a kind of *periodicity* (pulse or beat) and structured patterns of accentuation among these pulses (meter) in musical rhythms [14]. Such felt pulses can be defined as a kind of “endogenous periodicity”—as a series of regularly recurring, precisely equivalent psychological events that arise in response to a musical rhythm [15]. They do not necessarily coincide with every event onset of the sounding music, and they may even occur in their absence, though there is a general tendency to gravitate toward the onsets of sounding events to produce a kind of synchrony with an external rhythm in case this is periodic [14]. Figure 1 provides an example; it depicts the piano transcription by Busoni of a piece by Bach. Though perceived as music with a strong periodicity—a kind of softened beat—the graphical depiction in the waveform notation (upper pane) shows a continuous shape of the sound wave, rather than a succession of discrete slices of the temporal unfolding. The latter can be recognized more easily in the spectrogram notation (lower pane), where the onsets of each discrete event onset are much easier to discern. The graphical representations clearly show the tension between the objective acoustic structure and the structuring by the listener, which tends to discretize a continuous sonorous articulation into separate pieces that can be processed in cognitively less-demanding ways [16].

The concept of periodicity, further, has a lot of explanatory power because of its formal simplicity and similarity to perceptual experience [14]. Three major descriptive components can mainly be identified in this regard: trend (variation in tempo), cyclical or seasonal components, and irregular components [1]. The trend represents a general systematic component that changes over time (e.g., slow music, fast music, *accelerando*, *ritardando*, etc.), and that does not repeat within the captured time range; the cyclical components, on the contrary, occur repeatedly and contain information that concerns the underlying processes of rhythmic behavior.



**Figure 1.** Waveform (upper panel) and spectrogram (lower panel) of the first 31 s of Busoni's transcription of the Intermezzo of Bach's Toccata, Adagio & Fugue in C, BWV 564.

It is a small step from periodicity to the concept of *metrical structure*. Being defined as an alternation of strong and weak pulses, it is a generic concept that extends beyond the boundaries of music, sharing features between music, dance, poetry, and forms of metrical entrainment—as in, walking and marching [13]. As such, several studies have been conducted to investigate the relationship between music as an entraining stimulus and physical exercise that is repetitive in nature (e.g., walking, running, cycling, and rowing). Music, in that case, can be considered as having *ergogenic effects*, which are more pronounced when the periodicity of the musical stimuli is synchronized with that of the movement. Several musical characteristics or sound features, further, seem to explain most of the variability of the entraining stimuli. A binary emphasis in the music, stressing each beat, has an activating effect on the walking velocity and stride length; a ternary emphasis distracts from the binary structure, and has a relaxing effect on the walking velocity and stride length; and complexity of the rhythmic structure or tonal diversity diminishes the activating character of the music [17]. It can be questioned, however, whether this ergogenic effect is solely reducible to features of the sounds, as induced motor effects (found in the motor cortex and the basal ganglia of the brain) are not only dependent on rhythmic patterns in the music, but also on the evoked emotional experience, with the strongest activations in combination with feelings of joy, power, and triumph. It seems, therefore, that listening to music may incite listeners to move and walk in synchrony with the rhythm of the music, to induce involuntary movements—such as tapping the foot or swinging the head—and even to rely on active mental imagery of movement without overt movement [18].

Questions can be raised, however, by a rather narrow description of music as an entraining stimulus in terms of periodicity and metrical structure: not all music is reducible to metric rhythms; non-metric rhythms lack periodicity and regularity in their event onsets; and even metric music is not always strictly reducible to a metric grid, as exemplified in the many expressive timing mechanisms [13], tempo changes, or rubato in music performance [14], and the use of groove in jazz and popular music [9,19]. The broader concept of *rhythm*, however, can be used as an umbrella term for a wide variety of explicitly isochronous (a one-beat rhythm, such as the beating of a metronome or the ticking of a clock, with similar inter-onset intervals and durations of the events), as well as non-isochronous, time structures. An “external rhythm”, then, involves “a sequence of temporally localized onsets, defining a sequence of time intervals that are projected into the flow by some external event” [20] (p. 124). It should be noted, further, that the experience of an external rhythm as a “time structure” is not limited to the serial order of the events. The events should be perceived as temporally coherent to count as a structure as well. This holds not only for music, but for everyday events in general, most of which comprise of

structured actions and movements that display distinct and clear beginnings, recognizable rhythms, characteristic tempos, and lawful endings.

## 2.2. Attunement and Attention

The structure of the music is one thing. A “conceived” structure (by the composer, performer), however, is not the same as a “perceived” structure. Listeners, in fact, rely on listening strategies for sense-making [21,22]. There is, first, an “ecological approach” to perception, which claims that listeners do not perceive the acoustical environment in terms of phenomenological descriptions, but as *ecological events* [23,24]. These can be defined as higher-order variables or perceptual units with time-varying complex acoustic properties within temporal constraints, mostly in the range of 2–3 s (see [25–27] and [28] for an overview). Such events are continuous in their unfolding, but discrete in their labeling.

Ecological listening aligns with the natural tendency to reduce processing efforts in the search for information as an example of “cognitive economy” [28]. It is possible, however, to interfere with this “default mode of listening” by making a transition from mere information pickup to active and deliberate attention. Listeners, then, should adopt a dynamic attending approach, so as to be “tuned” adaptively to the minute changes of the music as it unfolds over time. It is a claim of recent research on attentional capture that echoes James’ early statement that “No one can possibly attend continuously to an object that does not change” [29] (p. 92) (see also [20]).

Attention, in this view, can be described in terms of “attentional entrainment” with the time structure of a distal or external event exerting control over the process of dynamic attending [20]. The key insight of the *dynamic attending theory* is that attention is not constant over time, but that it waxes and wanes with time’s passing, and that it can become coupled to the temporal structure of environmental stimuli [30,31].

There is, further, a lot of flexibility in the time structure of natural events. Most of them have some extension over time, and capturing the moment-to-moment dynamics of their unfolding entails a kind of “high-resolution” listening—referring to a temporal acuity in the range of milliseconds—that relies on *attentional tuning*.

The relation between attunement and attention has been somewhat underexplored up to now, with an exception in the domain of music affect and emotions [32], and music therapy [33]. Affect attunement, however, is not the same as attentional tuning, which aims at facilitating “attentional targeting” by optimizing the temporal positioning of attentional energy with successively finer discriminations of temporal nuances within the attentional scope. A distinction should be made, further, between the apprehension of stable rhythmic structures—as in, the search for patterns in the environment—and the adaptive responsiveness to temporal fluctuations and changes in the structure of attended events. This is commonly observed in music performances with musicians shaping the structure of rhythms and melodies to their own needs by exhibiting great temporal flexibility and expressivity [34–36]. It is arguable to assume that skilled listening should require the same sharpness and sensitivity by attuning the attentional focus to track the minute nuances in the temporal unfolding of the music.

Two major questions arise in this regard: how does an external event structure guide the observer’s attention, and how does his/her attending adapt when the event structure starts to change? A possible answer is the hypothesis of the *attending rhythms* that entrain to external events, allowing selective targeting of attention to expected points in time. They can be defined as “recurrent pulses of attentional energy that facilitate a listener’s response not only to expected information but also to temporally unexpected information” [20] (p. 121). Energy, in that view, is to be seen as a periodic attentional pulse that is targeted by an internal rhythm. It thus seems that such a dynamic attending model builds on both notions of “expectancy” and “entrainment” to describe real-time attentional tracking of time-varying events. It is a conception, held by many cognitive psychologists, which claims that perception, attention, and expectation are rhythmic processes that are subject to entrainment [1]. Or, to put in other words: human beings are inherently rhythmical with “tunable perceptual rhythms” that can entrain to time patterns in the physical world [37]

(p. 340). The claims are appealing, but a full operational description of the exact relationship between attunement and attention is still waiting for additional empirical support.

A distinction should be made, in this regard, between a “future-directed” and a more “present-centered” or “analytic style” of attending. The former occurs in the case of more coherent and periodic time structures, which makes an expectation more probable; the latter refers to stimuli that are less coherent and more complex, and thus, also less predictable [31,38]. Both styles can be associated with different kinds of experience, focusing on either teleological larger-scale organization or on immediacy and the specific qualities of the sounding stimuli [39].

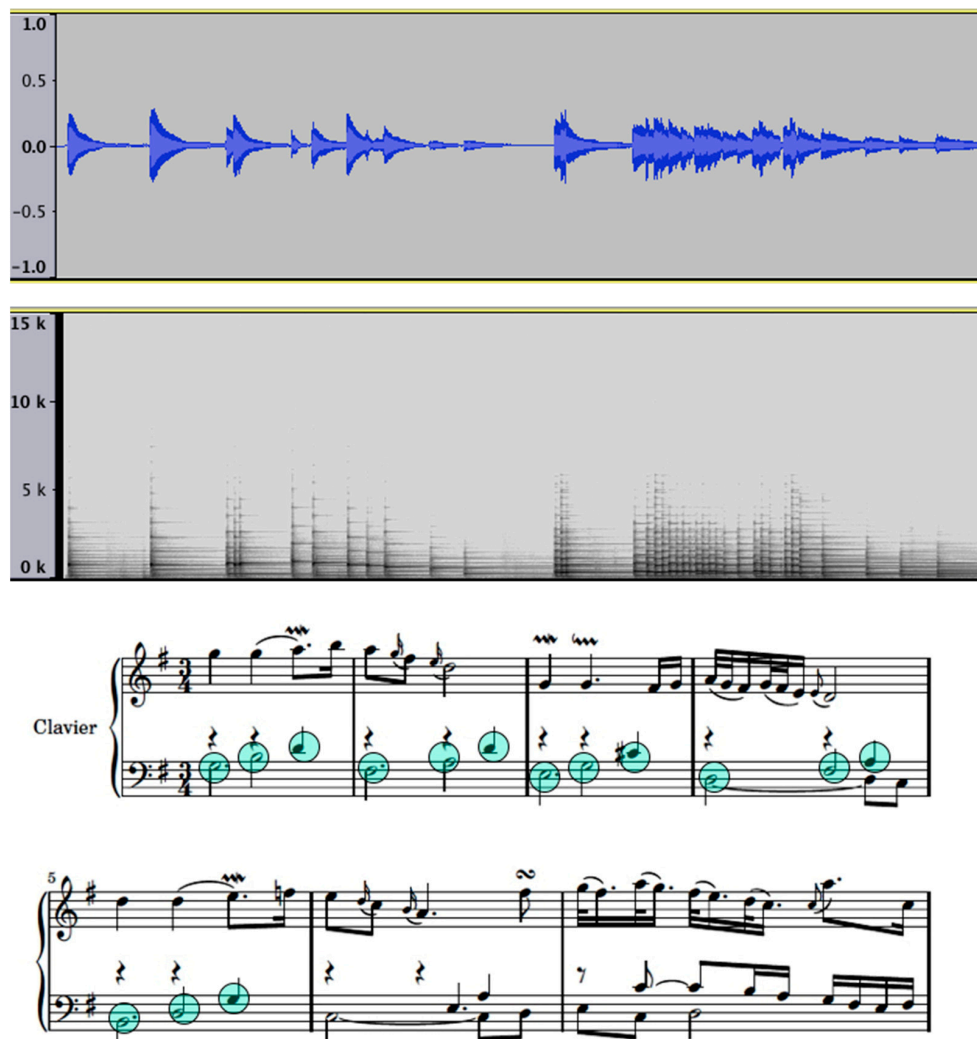
Both styles imply different levels of attention, with entrainment being described as an adaptive process to accommodate to the unexpected and the expected under the influence of both attentional direction and the configuration of the musical stimuli. Attention, then, can be understood as being partly intentional and goal-directed, and partly controlled by the characteristics of the external stimuli [1]. Both styles, further, are not opposed to each other, but alternate in a rather dynamic way, as in the case of navigating a vehicle through narrow streets. A skilled driver will not randomly focus on particular aspects of the surrounding (present-centered), but will let the environment control his/her attentional focus to some extent (future-directed), relying on a driving routine that enables him/her to cope with it in a more global way [40].

This brings us to the role of *attentional capture* and the question of whether our focus of attention is primarily controlled endogenously—by our intentions, goals, beliefs, etc.—or by environmental controlling factors. A famous quote by William James is quite enlightening here: “A faint tap per se is not an interesting sound; it may well escape being discriminated from the general rumor of the world. But when it is a signal, as that of a lover on the windowpane, it will hardly go unperceived” [41] (pp. 417–418). This “top-down” view is rivaled, however, by a “bottom-up” approach, which settles the focus of attention in the saliency of the perceiver’s environment [42,43]. It seems arguable that a correct view on attentional capture should integrate both views [40]. Figure 2 provides an illustration. What is shown is a depiction of the first measures of the aria of Bach’s Goldberg variations, both as waveform and spectrogram (upper panes) and as score notation (lower pane). As marked in the latter, there is a clear ascending pattern in the left-hand part, which is clearly identifiable as three separate notes of a broken chord, and which are clearly recognizable as discrete notes. The waveform and the spectrogram representation, on the contrary, show the continuous articulation of the spectral energy and are much more difficult to follow in real-time due to the lack of clear segmentation between the separate note values.

The encircled notes in the score representation have been manually added to show how top-down decoding or schematizing can structure the perception of the temporal unfolding. As soon as the first three notes are identified, they may instantiate, as it were, a pattern that has the ability to repeat itself. Such a pattern, then, can be described in technical terms as an *endogenous oscillator*, if abstraction is made from its position in pitch space. This means that the felt “pulse” of the music emerges as a kind of internalization of the regularity of the music as a kind of external driver, either for a limited or a local space of the temporal unfolding, or for a broader temporal span. It is a way of attentional tracking that oscillates between expectancy and expectancy violations with a dynamic interplay between an external rhythm (the music) and an attending rhythm (the internalized pulse). As such, it echoes older theories of attention that embody anticipation and surprise with dichotomies associated with the preparation and response to change or novelty (see [44–47] and [20] for an overview).

Two main mechanisms seem to be important in the dynamic attending approach: the temporal resolution of the attentional targeting (local/focal or more extended), and the coordination of the internal and external rhythms. What is crucial for good attentional coordination, however, is a narrow attentional pulse allowing to “track” the minor changes over time with attention to similarity and difference. The coordination of attention with a pulsating ambient flow—the periodicities within the time structure of an external rhythm—can tie an attender to a distal event

and shape the complex of internal oscillations via the mechanism of entrainment, allowing him/her to participate in the rhythm of such a remote event and to match up with certain time spans in that event. The notion of attentional pulse is crucial in this regard: it allows the attender to experience an event structure with regard to its expected and unexpected aspects and constitutes a form of attentional capture that reacts to the momentary changes in the external rhythm, in terms of coordination with an external rhythm, interval time properties, and synchronization quality [20]. The technical equivalents of these controlling variables are “phase”, “period”, and “focus” (see below) [48].



**Figure 2.** Waveform and spectrogram representation of the first four bars of the aria of Bach’s Goldberg variations (**upper panels**) and score notation of the first seven bars (**lower panels**), performed by Glenn Gould.

The discussion of attentional tracking, further, can be related also to the distinction between the *syntactical* and *processual* evaluation of the music [5,19], which can be reduced mainly to the distinction between discrete and continuous processing of the music [21,49]. This holds in particular for the vast majority of music that is performance-oriented, dance-derived, and partially improvised, as exemplified in jazz, polkas and blues, and many other groovy and sensual musics of the world. There are aspects of the ongoing musical progress that can be characterized in terms of groove, swing, or vital drive, and which can be subsumed under the heading of an *engendered feeling*, in the sense that the perceived rhythms conflict somewhat with the pulse, but without destroying it altogether. As Keil puts it: “something approaching complete comprehension of the processual aspect will only be possible when

we are able to determine accurately the placement of notes along the horizontal dimensions" [5] (p. 345). It is a concern that has led designers and programmers in the 1980s to design "human clocks" and "biological drummers" in order to make the music more "lively" by adding "feel" to drum machines and computer music [50]. Electronic music producers, for their part, have since engaged in several kinds of imperfections, inaccuracies, perturbations, offsets, adjustments, shifts, and feel issues—labeled with terms as snap, drag, loose, and stretched—as found in the processual discrepancies that can be found in the musics of the world [51].

### 3. Entrainment: What? What For?

As mentioned already above, the concept of entrainment has a lot of explanatory power because of its formal simplicity, similarity to perceptual experience, and computational elegance. The term, however, is used in different ways, and there are still problems that are related to its concept validity, scope, and breadth of meaning of the term. What is needed, therefore, is an operational definition, starting from the shared tendency of physical and biological systems to coordinate temporally structured events through interaction [1,52].

#### 3.1. Conceptual and Definitional Issues

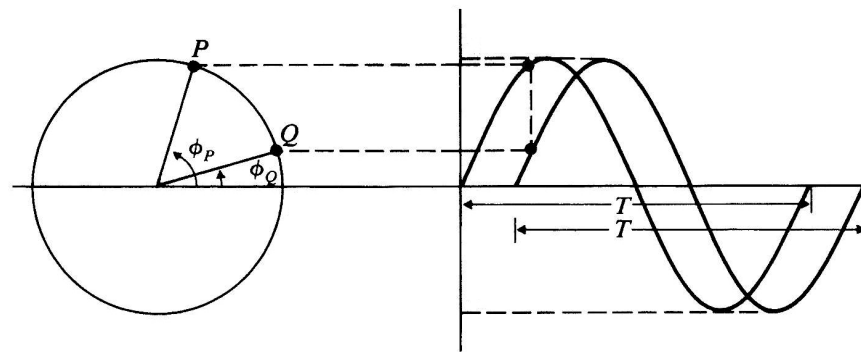
The concept of entrainment goes back to the Dutch physicist Christiaan Huygens, who discovered, in 1665, that the pendulum frequencies of two clocks that are mounted on the same wall or board become synchronized to each other. The effect—metaphorically coined as the "sympathy of the clocks"—was subsequently confirmed by many other experiments and was called *entrainment*. The concept was thereafter applied in mathematics, physical, biological, and social sciences [1,53].

Huygens' observation was a typical example of emergent behavior with the strange finding that the clocks' pendula were swinging in opposite directions (anti-phase synchronization, or "odd sympathy" in Huygens' words). His observations were also limited to the behavior of two clocks and his mathematical descriptions were related to the physics of single pendulums, without reliance on later mathematical developments in nonlinear systems and experiments on the synchronization of many clocks that swing simultaneously [54,55]. Later attempts to generalize his findings to multiple oscillators still take his findings as a starting point, but the mathematical and physical elaborations depart considerably from Huygens' original computations (see [56] for an update).

The supposed underlying mechanism is quite simple: the different amounts of energy that are transferred between moving bodies with asynchronous movement periods cause negative feedback, which drives an adjustment process until the moving bodies move in resonant frequency or synchrony. Or put differently, entrainment describes a process whereby two rhythmic processes interact with each other to mutually adjust or lock into a common phase and/or periodicity. Two conditions, at least, must be met: there must be two or more rhythmic processes or oscillators, and the oscillators must interact so that they get coupled. There is, moreover, a third condition that applies in the case of strict entrainment, namely the possibility of re-establishing synchronization after perturbations or transitions of the synchronizing process [1]. The coupling, further, can be either strong or weak, and the stronger body or oscillator locks the weaker into its frequency until both lock into a common movement period [53]. Such "frequency locking" of two oscillating bodies implies that bodies with different frequencies or movement periods when moving independently can adjust to a common period when they interact. The entrainment can occur in various temporal relationships of the movement onsets of the bodies that move together and can be described in terms of a "phase relationship", which is a commonly used measurement tool for an objective description of the dynamic behavior of the oscillators [57]. There are, as such, two aspects of entrainment that enable an operational description of the phenomenon: *frequency* and *phase-entrainment*. They describe how physically coupled oscillators are related stably in their timing (frequency) or in their spatial positions within a cycle (phase) [48,58]. Frequency or tempo entrainment refers to the periods of two



oscillators that adjust toward a consistent and systematic oscillation frequency; phase entrainment of phase-locking refers to the process where focal points in the temporal course—such as the foot striking the floor—occur at the same moment [1]. It is quite easy to represent a periodic rhythm by a cyclical motion mapped on a circular shape with one period ( $T$ ) being described as a full circle (one period is  $2\pi$ ) and the phase relationship of each moment in time in terms of the phase angle ( $\varphi$ ) and changes of it ( $\Delta\varphi$ ), with respect to the starting points of the entraining or entrained oscillations (see Figure 3). It makes it possible to describe the phase relationship between two oscillators that are frequency-locked as being either synchronous, lagging (after stimulus onset), or leading (before stimulus onset), and to describe the difference in the phase in terms of phase angle [1].



**Figure 3.** Example of phase and period relationship as exemplified in a sinusoidal depiction of a periodic stimulus.

The concept of periodicity thus seems to have a lot of descriptive power. One must assume, however, that from a broader biological point of view, temporal constancy of environmental qualities is the exception, with instantaneous shifts in response to transitions being the rule for adaptive observers. Music, even in the case of metrical music with a fixed time signature, is also mostly not totally isochronous—with equal time intervals between the onsets of the beats in each measure—and it does also not always provide a stable and sustained rhythmicity. There are mostly minor deviations, which are related to the microdynamics of the temporal unfolding, and which contribute substantially to the expressive qualities of the music. This applies also to the oscillatory dynamics of the listener if we are ready to conceive of him/her as an oscillatory system with endogenously generated rhythms that are stable over at least some lapses of time. Entrainment, in this view, has to correct for differences in cycle lengths—both external and internal—to produce specific phase relationships between the oscillating systems [59]. It can be asked, further, whether entrainment to music should be limited to the synchronization of meter and rhythm, or to the dimension of time in music at all. It has been argued, in fact, that the concept could possibly be extended in the broadest sense by including also discrepancies in pitch and timbre [9,60,61]. There is no space, however, to go into detail on this here.

Care should be taken, further, to avoid the danger of reductionism. There are, first, actually two separate, but related, aspects of entrainment: *physical entrainment*, which refers to the objective measurement of synchronization of physical objects in the world—as in Huygens' pendulum clocks; and *cognitive entrainment* or feeling of synchrony, which refers to the subjective measure of an observer or participant [62]. Besides, there is also the phenomenon of *physiological entrainment*, which holds a position in between, as it is reducible to endogenous or naturally occurring rhythms that occur within the body and which are objectively measurable (cardiovascular dynamics, respiration rate, secretion of hormones, brain waves, and many others). They are mainly categorized as the effects of music which are triggered by the music as an external driver [4]. The presence of multiple modifying and intervening factors, however, makes the coupling between the music as eliciting stimulus and the evoked effects somewhat elusive. Much is to be expected here

from the study of the inductive power of music and its major role in affective tuning [32,63], with many responses that do not even reach the level of consciousness.

Besides this first and rude classification, other distinctions complicate discussions about entrainment even more, such as the distinction between external and mutual entrainment, rhythmical/metrical and non-metrical entrainment, symmetrical and asymmetrical entrainment, within-person and between-person entrainment, self-entrainment, etc. There is no space to elaborate on each of them, but the role of an *externally imposed force* to modulate the frequency of oscillation is of major importance in this overview. It is illustrated quite clearly in the case of someone who dribbles a basketball: instead of letting the ball bounce free at its own frequency, the dribbler can externally impose the number of bounces, thus forcing the ball to bounce at a rate that is determined by the dribbler. This imposed rate can be defined as a driving, or forcing, frequency, which turns the free vibrations of the bouncing ball into forced vibrations with the dribbler acting as a forcing function that controls the motion of the ball. The ball, then, becomes a “controlled system” and the dribbler takes the role of “controller”, provoking a continuous disturbance that controls the behavior of the system [4]. It is a clear example of Lord Rayleigh’s distinction between “forced” and “maintained oscillations”, the first being forced upon a system from outside and the latter being maintained autonomously by the system itself [64]. Generalizing a little and substituting the listener for the ball, this can be applied to listening to music as an example of *asymmetrical entrainment*, where the listener cannot influence the entraining rhythm of an object or someone else—in this case, the music—and where the body is forced to adjust to externally set cyclical conditions without being able to influence the latter. There is, as such, an entraining rhythm—the “Zeitgeber”, to coin the German term—, which triggers the adjustment of an entrained rhythm [1].

It is possible, further, to define the concept of *oscillator* in technical terms as an excitable medium, which can respond strongly to an imposed, relatively weak stimulus [65], and to conceive of this as a continuous dynamical system that typically responds to a given influx of energy with an elicited reaction when the supplied energy surpasses some threshold value [66]. Care should be taken, however, as indicated already by Poincaré, not to conceive of Huygens’ and Rayleigh’s observations in terms of a linear system where changes in one variable produce predictable and linear changes in a dependent variable. The oscillators should be described, on the contrary, in terms of *interacting non-linear dynamics* [67–69]. Translated to the realm of music, this should mean that we equate the listener’s body with the controlled system and the acoustic characteristics of the music with the continuous disturbance or persisting disturbing functions that elicit possible physiological or psychological responses.

Most studies on musical entrainment have focused on external (metrical) entrainment, as in finger tapping to a metronome or dancing to recorded music. People, in that case, synchronize with an external signal, and entrain “to” the signal, mostly in a solitary or individual context. There is, however, also the possibility of *mutual entrainment*, in which case the pacing cue for synchronizing is generated by a group of subjects rather than by external cues. Such mutual entrainment is the dominant kind of entrainment in the animal world, with examples to be found in the flocking behavior of birds flying, fishes swimming in formation, crickets chirping in unison, and fireflies flashing synchronously [70]. However, even human beings manifest examples of mutual entrainment, as in the case of a rowing team, a bucket brigade, or a group of marching soldiers formation, and even more convincingly in the case of duetting pianists, string quartets, and groups of dancers [13]. The case of musicians playing together is quite interesting, as it seems to be a kind of merging of external entrainment—listening to the sounds of the others—and mutual entrainment. String players in a trio or a quartet are not only attentive to the sounds produced by themselves and by the other players, but beyond the produced sounds, there are lots of nonverbal communications which enable the high level of needed synchronization. The same also applies to dancers and some other kinds of synchronized sports, such as swimming or diving.

The case of musicians playing either alone or together is an interesting new avenue of research. In case of collective playing, there is always the need to listen to the sounds of the others, which functions as an external entrainment. There is, however, also the sound produced by the individual musicians themselves, which is self-generated and endogenous to some extent. Especially in the case of metrical music, this entails a kind of metrical entrainment that embraces three kinds of pacing movements: self-pace, mutual entrainment, and external entrainment [13]. The case is even more complicated in the case of solo playing. The only sound that is produced is self-generated, which means that there is no fixed external timekeeper. The pacing, therefore, is self-entrained. However, as soon as the performer has to synchronize with another player or with other-generated sounding music—as in the case of playing along with recorded music—the musician must be adaptive and must take also the role of follower instead of instigator of the sound. The metaphor of *master and slave* is quite appealing in this regard. It exemplifies the relation between coupled systems in terms of receiver to transmitter, follower to leader, and between being driven and driver [66]. It can even be asked whether the metaphor is also applicable to piano playing with both hands, where one hand provides the pace to proceed and where the other hand is adapting in an attempt to synchronize with the other hand. It is a refreshing idea that opens up theoretical reflections about the relation between the within-person coordination and between-persons coordination, and their common underlying dynamical principles [48].

### 3.2. Alignment and Adaptation

The metaphor of master and slave has a lot of explanatory power. An obvious application is the distinction between performer and listener. Performing is mainly considered as being active, while listening should be more passive. This is of course a bold statement which should be questioned in the sense that active listening should aim at “enacting” the sounds to make sense of the sounding music (see [22]). It means that music listening may enhance motor facilitation, to re-enact those motor actions that are required to perform the music that is heard [71], or more in general, that music may induce some kind of motor response. This is the well-known phenomenon that has been coined also as “kinaesthetic listening”, in the sense that the gestural experience of the production of sound interacts with its perception so that listeners can feel the music in their muscles and imagine what it might be like to play what they are hearing [9]. It is a claim that points in the direction of “motor resonance” as a mechanism that supports action understanding [3].

Tapping with the foot or moving the body to the beat are typical examples of overt motor responses to the music. Yet, it is possible also to feel the music at an internalized level of motor imagery [72,73]. The coupling between the objectively sounding and subjectively felt pulse as externalized in motor behavior, however, is not always very strong, as many listeners are not able of close beat-tracking when moving along with the music. It may seem that they have difficulties with the respective role of master and slave in taking the role of master as the default role of processing of the sounds. Such is an asymmetrical role in which the perceiver imposes his/her pace on the music without being receptive to the sounding input. A symmetrical approach, on the contrary, should entail an equivalence between input and output, between incoming and outgoing energy, with the perceived input triggering the motor output. It puts a lot of emphasis on the temporal alignment of the motor output with an external timekeeper so as to entrain rhythmical movements to an external timekeeper [74].

The mechanisms behind this coupling are complex and are not yet totally understood, with conflicting theoretical proposals that revolve around the notions of prediction and anticipation on the one hand, and reactive behavior and feedback on the other. A lot is to be expected here from current research on *sensorimotor synchronization* [75–82] and *predictive coding* [83–86], as well as from older theories of *central* and *peripheral theories of motor control* [87–90]. The latter revolves around the question of whether motor control is dependent upon the nervous system to coordinate movement through instructions from its

central components—memory representations or motor programs that provide the basis for organizing, initiating, and carrying out the intended actions, as exemplified in the “schema theory” of motor control [91] (see also [88])—or by information that arises from the dynamic interactions between the mover, the task, and the environment through peripheral feedback, and which is considered as movement-related specifications that can vary from one action to another. The latter, moreover, has received renewed input from the dynamical systems theory [48]. There is, however, no space to go in detail here.

Alignment to external stimuli, further, depends on the short-term attentional control with many small adaptations that represent fast, transient adjustments to novel and unexpected timing. As such, it directs attention on expectancy violations and unexpected onsets within an external rhythm rather than at long-term predictive attentional control [20], somewhat related to the above-mentioned distinction between the future-directed and analytical style of attention. Both styles, however, are complementary, keeping in mind that prediction is at least as important as tracking in the entrainment process [1]. Prediction is closely related to periodicity, but in real-life settings and also in music, there is a lot of non-periodicity that is superimposed on the periodic grid. It is the major challenge for motor programming; therefore, to let the spontaneous oscillatory properties of the body—e.g., pendular swings of the limbs and the torso—be driven to move also in the non-periodic ways that a real ecological setting requires [39]. This is exemplified quite typically in dancing along with music as a kind of emulation of the dynamics of the music where the body movements are aligned in time with the unfolding of the music. The emulation, in that case, is focused more on the continuous aspect of the movement as a whole than on the discrete markers that characterize the salient moments of the movement. What matters is the spatial and temporal alignment with the music [17].

There are further two possible problems of temporal alignment: a lack of synchronization of the phase in case of stable periodicity (lagging or leading), and minor deviations from the onsets of the externally paced events. As such, three different strategies have been identified in recent empirical research: (i) finding the beat or pulse, (ii) keeping the beat, and (iii) being in phase [17]. Finding the beat is the most problematic one. Keeping the beat reflects minor differences in musical timing that involve essentially micro-rhythmic phenomena [9]. They make it possible to define entrainment in adaptive terms as a sensorimotor process that involves a *phase-error correction mechanism* that acts as an updating mechanism toward a more stable phase synchronization [17,78,92]. Such correction, however, is constrained in the sense that the coupling strength between one tempo and another—such as between music and the entrained movement—can overcome this difference in movement period only within a specific range of differences in period—mainly less than 2.5 to 3 percent—which has been called the “entrainment basin” [93,94].

### 3.3. Oscillatory Dynamics: External and Internal Oscillators

Thinking in terms of error correction entails a normative conception of an entraining external rhythm that acts as a driver. The concept of external or exogenous rhythm, however, is somewhat ill-defined as there are also endogenous or naturally occurring rhythms that occur within the body with its multiple cycles, periods, frequencies and amplitude—examples are the cardiovascular dynamics, respiration rate, locomotion, secretion of hormones and many others, all of them at different time scales—, and which can also function as timekeepers for overt or internalized movements. It means also that not all entrainment is dependent on an external stimulus, either environmental or interpersonal. There is, e.g., the possibility of *self-entrainment*, which means that in a complex body, a gesture of one part of the body tends to entrain gestures by other parts of the body. It means that two or more of the oscillatory systems of the body become somehow synchronized in some way. Respiration and heart rate, respiration and walking rhythm, and synchronization of arm movements and leg movements in walking are obvious examples [1] (and see [95,96] for a broad overview).

There is, moreover, a distinction to be made between those endogenous rhythms or oscillations which are overt and perceptible to the instigator of the (motor) response in general, and those that are hidden from consciousness. The patterns of spontaneous human gait—a bipedal locomotor cycle—are exemplary of the former, neural oscillations, and rhythmic and repetitive patterns of neural activity in the brain exemplify the latter. The distinction is important as it questions the role of cognitive components in the entrainment process, as well as the impact of physiological entrainment as part of the responses by the entrained responder. The so-called perceptual and attentional oscillators, which are designed to automatically adapt to match external oscillators, are, in fact, highly influenced by the responder's cognitive capacities [1]. Entrainment, in that view, is a kind of interactive attending that creates a synchronous interplay between an attender and an event, with the former tending to partially share the rhythmic pattern of the eliciting event [31]. This is a conception that defines *oscillatory dynamics* as dynamical attending by actively resonating with our body and the sonic environment [20], as contrasted with passive resonance, which is not considered to be entrainment in a strict sense [1]. It means that the temporal regularity in the sensory input can be linked to temporal expectations that rely on some internal model that functions as a simple oscillator that is able to synchronize with a stimulus sequence [97,98].

The suggested frameworks seem plausible. There are, however, still many open questions, especially with regard to the so-called *internal clocks* and the allocation of *neural resources* for endogenously generated oscillating rhythms in the brain. The matter is quite complex and there are currently many disparate findings that still lack coherence and explanatory power. It has been hypothesized, in this regard, that spontaneous cortical oscillations in specific frequency ranges could explain the experience of endogenous periodicity [99]. Beta- and gamma-band oscillations have been found to be the neural correlates of responses to auditory rhythms in EEG and MEG studies on perception and cognition, with a major focus on the temporal expectancy [100,101]. The experience of pulse and meter, accordingly, should be the result of neural oscillations that resonate to rhythmic stimulation [14,102], and studies of synchronized body movement to metronomes and music have shown a natural propensity for neural or motor-based resonances, which have a limited predictive capacity [17]. There seems to be a mechanism that reflects a temporal orienting of attention via a dynamic allocation of neural resources that facilitates perception [103]. More in particular, the clocking activity of the cerebellum seems to be involved in tracking environmental inputs and preparing for alignment of motor output when required by the context as evidenced by beta power dynamics [104,105]. Beta oscillations with frequencies in the vicinity of ~20 Hz, are thought to encode prior representations of the environment, especially the time-varying structure of external events [106]. They also exhibit functional specificity for top-down processes as attentional gain [103,107,108] and predictions of causal events [109]. It has been found, moreover, that beat perception is processed through sensorimotor loops that entrain beta power at the rate of stimulation. The latter, in particular, seem to be related to motor functions and have a role in mediating top-down connections, as advocated also in current formulations of predictive timing, which requires the construction of an internal model of temporal regularities to allow the precise temporal occurrence of future events [86]. These temporal predictions modulate both the amplitude and phase of ongoing neural oscillations, thus determining the accuracy of sensory processing [100]. Generalizing a little, beta power can be considered as an index of integration and global efficiency of the brain network as a whole [110,111]. In a narrower sense, it seems to be a common denominator for predictions with evidence converging towards a shared timing mechanism that is used to pace behavior to rhythmic stimuli through action-perception loops based on predictive control [105].

The findings are challenging but not yet conclusive. They start to uncover some aspects of the relationships between some kind of external pacing and neural oscillations that resonate with these stimuli. This is the domain of *neural entrainment*, which tries to show how low-frequency neural oscillations synchronize with ongoing patterns of event

onsets through adjustments of their phase and period [98]. A distinction must be made, however, between a generalized increase in brain activity in some targeted frequency registers while entraining to an external driver and a time-varying locking of phase and frequency to this driver. It is a problem that is related to the time scales of biological rhythms—see the main contributions on chronobiology in this regard [112]—which are mostly much more extended (e.g., circadian and ultradian rhythms) than the microtimings in terms of seconds and milliseconds which are so typical for the moment-to-moment approach to music listening. This is the hallmark of the dynamic attending approach with its emphasis on the immediacy of attending, the role of expectation, and focal attending shifts [31,37] and its inscription within the entrainment hypothesis, which assumes that internal oscillations or attending rhythms generate expectancies and anticipations, and that the rhythm of external events drives these attending rhythms [20]. There are, moreover, a lot of modulating factors, which make it possible to conceive of different degrees and phases of entrainment. There are basically three of them: (i) not all oscillators will entrain, (ii) some oscillators may entrain more or less strongly and more or less quickly than others, and (iii) there is even a whole spectrum from weak to strong coupling, with respect to frequency or tempo entrainment [1].

#### 4. Enactive Listening and Dynamic Interactionism

Conceiving of listeners and music as oscillators has a lot of operational power. There is, however, a danger of reductionism by narrowing down the whole process of listening to mechanistic explanations. In what follows, an attempt is done, therefore, to broaden this approach by highlighting the role of dynamic interaction and the definition of the listener as an autonomous agent. We take as a starting point that at a primordial level all experience is passively motivated by the sensing body. Yet, passive is not synonymous with inactive. It calls forth a fundamental openness to the world and a way to actively constitute structurally coupled relationships with the environment through motivated and intentional activity [113]. Dynamic interaction, then, fits into the context of the fast-growing *enactive turn* in cognition, with a special emphasis of the role of embodiment of an autonomous agent [114,115]. There seem to be two complementary approaches in this regard: (i) a view on embodiment as a subjectively lived state, and (ii) a more objective view of our body as a living, biological organism [116–118]. Both approaches exemplify an evolution that expresses some discomfort with a rather narrow approach to cognitive science in the 1990s—mainly restricted to problem solving and representation of the world—to include also consciousness, emotion, and dynamic embodied interaction with the world so as to become more closely related with an everyday-lived human experience [119].

The enactive approach is a multifaceted approach that questions how minds relate epistemically and experientially to the world. It conceives of minds as being part of embodied biological organisms, with nervous systems that generate meaning, arising from the sensorimotor with their environment to “enact” or “bring forth” their world. It is a conception of the world and organism mutually co-determining each other in the sense that the organism’s experiential awareness of its self and the world is a central feature of its lived embodiment in the world, echoing, to some extent, the continental philosophical tradition of phenomenology [120].

##### 4.1. Interaction Dynamics and Mutual Entrainment

The enactive turn in cognition has revolutionized cognitive science. It has been continued, however, by another still more encompassing approach, namely the interactive turn in social cognition research [121,122]. This *interactionism*—often seen as anti-individualism—includes the so-called second-person approach, interaction theory, and the enactive approach (see [123] for an overview). Central in this broadened approach is the insight that interaction complements individual cognitive processes in the sense that social cognition relies on processes outside the individual while simultaneously also being done in the individual’s head. Interaction, in this approach, can be defined as “a mutually engaged coregulated cou-

pling between at least two autonomous agents where the co-regulation and the coupling mutually affect each other, and constitute a self-sustaining organization in the domain of relational dynamics." [124] (p. 572). The mutuality aspect is quite important here, to the extent that dissolving the autonomy of one agent—by reducing its role in the coupling to mere co-presence rather than being a regulator—annihilates the concept of interaction in a strict sense. It is an answer to the criticism that has been raised against the "spectatorial" view on relations with others as a detached, "third-personal" attitude that states that we are merely spectators or observers of others' behavior [125–127]. What is needed, on the contrary, is a "second-person approach" [128], as exemplified so typically in cases of joint attention, which can even be seen as a kind of behavior [40].

This *second-person approach* fits well within the area of cooperative joint action, creating a feeling of unity and common sense—a sense of "we-ness"—at the expense of a well-defined sense of self [13]. It is an emerging new area of research that is fed from mother–infant interactions in terms of entrainment [129], as well as from studies on perceptual interaction in minimalist virtual environments [130,131] and participatory sense-making [121,132]. Central in these approaches is the aim to continuously reinstate mutually responsive interactions between the participants of a communicative exchange. Such dynamic interactions can take place between performers in a musical ensemble, band, or big orchestra, which is, as it were, the default or strong version of musical interaction. It is possible, however, to conceive of weaker versions of interaction by modulating the strength of the symmetric attunement by substituting the listener for one of the performers, and by substituting a virtual performing agent or agents for the sounding music. Skilled listeners, in fact, can invest in kinaesthetic listening (see above), and listen with an "as if"-mode of listening, as if they would be performing themselves at a virtual level of imagery. Care should be taken, however, not to confuse entrainment with resonance. In physical terms, there is no real entrainment when one of the oscillators cannot influence the other—as in the case of listening—but active listeners can manifest such a degree of musical engagement that it seems as if they are really interacting with the music, and this has a lot of implications for music education and the acquisition of listening strategies.

A most promising way to deal with these interaction dynamics is to start from interpersonal or *mutual entrainment*, which is characterized by the continuous interactions and adaptive adjustments by "all" of the agents of the communicative exchange. This entails a relationship that is multidirectional, with an interplay between leading and following, and which can be conceived in terms of "entraining with" rather than "entraining to" [13] (p. 238). Entrainment, in this view, must be seen as a particular case of entrainment in social interaction, with members of the system that work out a negotiated temporal order to adjust their activity patterns to coordinate with each other. The members of the social system, in a more technically restricted view, can be seen as oscillators that entrain to or with the music in terms of conditions and implications (symmetrical or asymmetrical entrainment), such as the negotiation of the relative power between the individuals/oscillators, somewhat analogous to interaction dynamics that are conceived in terms of master and slave [1].

#### 4.2. Relational Dynamics and Participatory Sense-Making

Music can be so overwhelming that it alters the systemic operating set-points of physiological functioning, giving the impression that the music happens "within us" and not merely "to us" [4]. An obvious application of this claim is measuring the effects of music on the physiological and psychological functions of the listener (see [11,63,133]). Of particular importance, in this regard, is the research on "chills and thrills" and on "being moved" by the music [134–139]. A related area of research is to investigate the respective roles of the many *sensory modalities* in the listener's continuous engagements with music. The primary role of the acoustic medium has been clear from the beginning, as acoustic interactions are conspicuous in nature. Some of their physical and perceptual features make this channel most interesting for maintaining a sustained interaction between adaptive autonomous systems [140] (p. 28). The tactile sense, on the other hand, is also

quite significant. There is, in this regard, the phenomenon of “passive touch”, which was identified by Merleau-Ponty as the foundation of intersubjectivity [141–143]. It refers to the possibility of tactile contact with an object—such as an apple falling on our head while walking under a tree—and undergoing a tactile sensation without having the conscious experience of the presence of another subject. Such a passive experience of mere tactile impact is clearly not a sufficient condition for having the feeling of being touched by someone or something else. This is the case, however, when there is a more continuous form of actual or potential contact. It is a finding that converges on a relational hypothesis of the perception of a “you” that is constituted by the attention of another subject—or an object considered as a virtual subject—being directed to “me”. Such shared attention has been shown to have a sensorimotor signature, as evidenced from the analysis of time series of embodied interactions between individuals, with a major finding that one’s conscious perception of another person’s presence is characterized by an increase of one’s passive reception of tactile stimulations by the other [142]. Such experiences of passive touch have been found—in experimental settings—to be consistently followed by a switch to active touching of the other while the other becomes more passive. It is an interesting finding that illustrates the intersubjective experience of turn-taking and movement synchrony, as exemplified also in recent experiments on *perceptual crossing*. This experimental paradigm is a physical setup for the study of real-time and online minimal sensorimotor interactions between two persons (dyadic interactions). Conceived as a design for the study of social interaction, it provides the most basic conditions for studying the factors that are involved in recognizing others in online interactions by creating a mini-network of two minimalist devices by placing pairs of blindfolded human participants in separate rooms and asking them to interact in a common virtual one-dimensional perceptual space (see [130–132]). The aim of the experiment was to investigate whether participants were able to differentiate the perception of another intentional object from the perception of a fixed and a mobile non-intentional object in conditions of reduced sensory stimulation—only haptic information was provided. Participants were asked to move a cursor (an avatar representing their own body as an intentional object) along a line by using a computer mouse and received a tactile stimulus to the other hand when they encountered something on the line (either the avatar of the other, a static object or a displaced shadow image of the other). They were asked to click the mouse button when they thought they perceived the presence of the other participant on the one-dimensional line. The only difference between the avatar of the other and its shadow image is that the former can at the same time perceive and be perceived to be involved in live dyadic interactions. As a result, the participants were able to locate each other successfully—they were able, as it were, to “cross” the line perceptually, hence the concept of “perceptual crossing”—because the ongoing mutual interaction afforded the most stable situation with both participants being mutually engaged in the same interaction. There is, so to say, a kind of self-perpetuation of mutually responsive interactions in the sense that the agents serve as each other’s sensor interface to mutually and interactively establish a kind of coherence of the sensorimotor loops of the individual, who reinforce the interaction as a whole [144].

The perceptual crossing experiments have received a lot of resonance in different areas, and it seems arguable to apply them also to the realm of music. They clearly challenge the long-held conception that social cognition or understanding must be merely cognitive, thus underestimating the role of perceptual experience and interaction dynamics for the understanding of social awareness [128]. It is a major finding of the “relational hypothesis”, therefore, that social awareness cannot be reduced to the study of isolated individuals. This basically means that the study of the other’s intentional and attentional engagement with our self helps to constitute the awareness of the presence of the other, thus valuing the interactive above the individual approach. This holds for interpersonal interactions, but it can even be applied to interactions with oneself. By attending how the right hand touches the left hand, e.g., it is possible to experience oneself as the passive object of that action and to have experiential insight in how it feels to be the target of someone intentionally



touching us [142]. It requires little imagination to translate this to the activity of piano playing with both hands, or to listen to music in terms of being touched “as if” the music touches the listener as a virtual agent.

Elaborating more in depth on social cognition or awareness, there is also the concept of social bonding, with a distinction between one-on-one engagement in dyadic interactions and engagements with wider social groups, allowing individuals to define themselves in terms of personal relationships with individual others or in terms of collective connections to a group or social category [145]. It stresses the role of joint experience and the role of intersubjectivity and shared feelings, as evidenced in group performances when musicians are playing together. The case is more difficult, however, in case of mere listening to music (either in a live performance or from recorded music) with the major question of whether the music is to be considered as a constitutive part of the interaction between the participants of an experiential exchange. A strong version of this approach would claim that a primitive acquaintance with another in an act of joint attention plays a role in the control of the focus of attention and of behavior more generally. What matters, therefore, is not merely the presence of a person or a physical object, but the person or virtual object as a co-attender of the perceptually presented environment. It is, in other words, a question of acquaintance with aspects of mental life, commonly known as the “problem of other minds” [40].

The study of “joint action”, further, entails a comprehensive understanding of social interactions in real time, challenging the conviction that perception, action, and higher-level cognitive processes can be understood by the investigation of individual minds in isolation. Joint action, in particular, is to be regarded as any form of social interaction with two or more individuals coordinating their actions in time and space to act on the environment and/or to reach common goals. Several mechanisms come into play in this regard: a mechanism for sharing the same perceptual input and direction of attention to the same events; a close link between perception and action to form representations of the goals of others’ actions and to predict the outcomes; to predict actions based on the observation of the environment; and to integrate the observation of the actions of others with one’s own action planning [3].

This emphasis on social cognition and intersubjectivity is typical of a new emerging strand in cognitive research that challenges the “methodological individualist” or “methodological solipsist” approach to cognition and agency (see [146]). It marks a growing interest in the role of interaction between agents for understanding the nature of agency. De Jaegher’s and Di Paolo’s account of social interaction have been quite decisive in this regard. By introducing the concept of *participatory sense-making*, they argue that inter-individual interaction processes can take on an autonomous organization of their own, or stated differently: the dynamics of the interaction process as a whole can be constitutive for the individual behavior of the agents [125,132,144].

There have been several attempts to elaborate on the inter-individual interactions. One of them is the study of *interpersonal coordination*, which, according to Brown, embraces the three domains of *physical space*, *time*, and *pitch space*, intending to generate a collective coordination of emotion and motivation to drive coordinate actions by a social group [13]. The coordination in physical space is quite obvious, as exemplified by soldiers marching forward in a line or as in group dancing. Coordination in time entails the synchronization of movement and sound generation that occurs during playing music together or chorus-ing. Synchronizing with others, moreover, is highly valued as a pleasurable activity. It is a typical example of mutual entrainment and has been described both as “collective effervescence” [147] and “social attunement” [148]. Coordinative arts, such as music and dance, therefore, seem to enhance feelings of affiliation and connectedness, which can lead to a sense of cohesion and “we-intentionality” [13]. Coordination in pitch space, finally, is perhaps the most problematic domain of interpersonal coordination. Yet, humans can coordinate their melodic lines in both pitch space and time to create a diversity of choral textures, from melodic unison to the most complex forms of polyphony [149]. To quote

Brown: “[j]ust as rhythm provides ‘time slots’ for coordination, so too musical scales provide ‘pitch slots’ for coordination during chorusing.” [13] (p. 61). Tonality, in that view, is a coordinative device just like rhythm, with singing in unison, in particular, being an utmost unifying activity that creates the illusion of singing as with one voice. Making sounds together, in perfect synchrony, chorusing either in octaves or consonant or dissonant harmony, therefore, is a driving force for group-coordinative behavior [149].

## 5. Conclusions and Perspectives

Listening to music should not be reduced to mere sensory exposure and passive resonance. It entails, at best, a continuous and ongoing interaction with the sounds. Listeners, then, can be entrained by the music, that acts as an entraining stimulus that impinges on their body and their mind. There are, however, levels and degrees of entrainment, with a major distinction between strong and weak entrainment, and between a narrow and broad definition of the term. In a strict sense, entrainment should be considered as a symmetrical relation between oscillators that mutually entrain each other. This seems to be obvious in the case of musicians playing together—both being living beings—but this strict condition is more problematic in the case of mere listening, which can be conceived in traditional terms as the music acting on the listener in a kind of unidirectional way, without the listener being able to modify the sounding music. The music is then considered as an unchangeable external stimulus, taking the role of master in the “master and slave” metaphor. This unidirectional approach has been challenged, however, by new emerging paradigms in music education and recent cognitive science, which have seen the upheaval of enactive cognition. These new findings have engendered an emancipation of the listener from a passive consumer to an active agent, who interacts in a dynamic way with the music, which can be considered as a virtual agent. It is an approach that celebrates an “as-if”-epistemology that conceives of the music as if it should be an imaginary living being, and the listener as an imaginary music player. There are, of course, levels and degrees of hypostasizing, with at one extreme the listener being a music player himself/herself, and at the other extreme the music coming out of the speakers, so as to be a kind of unchangeable sounding artefact. As shown in this paper, there are lots of intermediate forms, and the concept of interaction dynamics is likely to be quite promising in this regard.

As to the operational definition of entrainment, there are still many open questions. Quite a lot of research has been done on the concept of metrical/rhythmic entrainment in reductionistic or minimalistic settings (tapping to a metronome, tapping to the beat), with a lot of major findings. Questions can be raised, however, as to the ecological validity of some of these experiments, as entrainment to the external world is not limited to entraining stimuli with strong periodicity. It is challenging, therefore, to consider the possibility of non-metrical entrainment—broadening the narrow conceptions of “beats” to the ecologically-grounded concept of “sonorous events” with some saliency—which is directed also at the minute time-varying dynamics of the sonorous articulation of sounding stimuli. This holds for music, which, in most of the cases, is not strictly metrical at the highest temporal resolution level. It can be questioned, further, whether the concept should be broadened to embrace also allocations in pitch and timbre space. A lot is to be expected in this regard from ongoing research on the role of neural oscillations—in particular, frequency bands—and their sensitivity to the spectral configuration of the entraining stimuli, but this research is just starting. Major questions are also not yet fully resolved, such as the distinction between period and phase entrainment as related to musical entrainment. It seems likely that phase entrainment, in particular, should be of primordial importance for listening that is characterized by a high level of attunement and alignment to the sounding music. The whole aspect of synchronizing with the sounding music, in fact, comes down basically to be “in time” with the music, without leading or lagging, with respect to the entraining external stimuli.

It is challenging, finally, to conceive of educational and clinical implications of the concept of entrainment. There is, first, the important aspect of attunement, which entails

a kind of dynamic attending and a heightened perceptual openness. Besides, there is the interactional aspect, both with the music and possible other music players, with already some groundbreaking research in the domain of participatory sense-making. This is all quite exciting, and it opens up new and refreshing avenues for future research, stressing the role of active engagement of the listener as an agent in a setting of dynamic interactionism, with an emphasis on continuous sound tracking rather than merely recognizing some discrete gravitational centers in the temporal unfolding.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Conflicts of Interest:** The author declares no conflict of interest.

## References

1. Clayton, M.; Sager, R.; Wil, U. In time with the music: The concept of entrainment and its significance for Ethnomusicol. *Eur. Meet. Ethnomusic.* **2005**, *11*, 1–82.
2. Sebanz, N.; Knoblich, G. Prediction in joint action: What, when, and where. *Top. Cogn. Sci.* **2009**, *1*, 353–367. [[CrossRef](#)] [[PubMed](#)]
3. Sebanz, N.; Bekkering, H.; Knoblich, G. Joint action: Bodies and minds moving together. *Trends Cogn. Sci.* **2006**, *10*, 70–76. [[CrossRef](#)]
4. Schneck, D.; Berger, D. *The Music Effect. Music Physiology and Clinical Applications*; Kingsley Publishers: Philadelphia, PA, USA; London, UK, 2010.
5. Keil, C. Motion and Feeling through Music. *J. Aesthet. Art Crit.* **1966**, *24*, 337–349. [[CrossRef](#)]
6. Blacking, J. (Ed.) *The Anthropology of the Body*; Academic Press: London, UK, 1977.
7. Lomax, A. The cross-cultural variation of rhythmic style. In *Interaction Rhythms. Periodicity in Human Behavior*; Davis, M., Ed.; Human Sciences Press: New York, NY, USA, 1982; pp. 149–174.
8. Keil, C.; Feld, S. *Music Grooves: Essays and Dialogues?* University of Chicago Press: Chicago, IL, USA, 1994.
9. Keil, C. The Theory of Participatory Discrepancies: A Progress Report. *Ethnomusicology* **1995**, *39*, 1–19. [[CrossRef](#)]
10. van der Schyff, D.; Schiavio, A.; Elliott, D. *Musical Bodies, Musical Minds. Enactive Cognitive Science and the Meaning of Human Musicality*; The MIT Press: Cambridge, MA, USA; London, UK, 2022.
11. Reybrouck, M.; Vuust, P.; Brattico, E. Neural Correlates of Music Listening: Does the Music Matter? *Brain Sci.* **2021**, *11*, 1553. [[CrossRef](#)]
12. Raffman, D. *Language, Music and Mind*; MIT Press: Cambridge, MA, USA, 1993.
13. Brown, S. *The Unification of the Arts. A Framework for Understanding What the Arts Share and Why*; Oxford University Press: Oxford, UK, 2022.
14. Large, E.; Snyder, J. Pulse and Meter as Neural Resonance. *Ann. N. Y. Acad. Sci.* **2009**, *1169*, 46–57. [[CrossRef](#)]
15. Cooper, G.; Meyer, L.B. *The Rhythmic Structure of Music*; University of Chicago Press: Chicago, IL, USA, 1960.
16. Reybrouck, M. Deixis in Musical Narrative Musical Sense-making between Discrete Particulars and Synoptic Overview. *Chin. Semiot. Stud.* **2015**, *11*, 79–90. [[CrossRef](#)]
17. Leman, M.; Buhmann, J.; Van Dyck, E. The empowering effects of being locked into the beat of the music. In *Body, Sound and Space in Music and Beyond: Multimodal Explorations*; Wöllner, C., Ed.; Routledge: London, UK, 2016; pp. 13–28.
18. Vuilleumier, P.; Trost, W. Music and emotions: From enchantment to entrainment. *Ann. N.Y. Acad. Sci.* **2015**, *1337*, 212–222. [[CrossRef](#)]
19. Keil, C. Participatory Discrepancies and the Power of Music. *Cult Anthropol.* **1987**, *2*, 275–283. [[CrossRef](#)]
20. Large, E.W.; Jones, M.R. The dynamics of attending: How people track time-varying events. *Psychol. Rev.* **1999**, *106*, 119–159. [[CrossRef](#)]
21. Reybrouck, M. Experience as cognition: Musical sense-making and the ‘in-time/outside-of-time’ dichotomy. *Int. Stud. Musicol.* **2019**, *19*, 53–80. [[CrossRef](#)]
22. Reybrouck, M. *Musical Sense-Making. Enaction, Experience, and Computation*; Routledge: Abingdon, UK; New York, NY, USA, 2021.
23. Balzano, G. Music Perception as Detection of Pitch-Time Constraints. In *Event Cognition: An Ecological Perspective*; McCabe, V., Balzano, G., Eds.; Lawrence Erlbaum: Hillsdale, NJ, USA; London, UK, 1986; pp. 217–233.
24. Lombardo, T.J. *The Reciprocity of Perceiver and Environment. The Evolution of James J. Gibson’s Ecological Psychology*; Lawrence Erlbaum: Hillsdale, NJ, USA; London, UK, 1987.
25. Wittmann, M. Moments in Time. *Front. Integr. Neurosci.* **2011**, *5*, 66. [[CrossRef](#)] [[PubMed](#)]
26. Wittmann, M.; Pöppel, E. Temporal mechanisms of the brain as fundamentals of communication—With special reference to Music Percept. and performance. *Music Sci.* **1999**, *3*, 13–28. [[CrossRef](#)]

27. Godøy, R.I. Imagined Action, Excitation, and Resonance. In *Musical Imagery*; Godøy, R.I., Jørgensen, H., Eds.; Swets & Zeitlinger: Lisse, The Netherlands, 2001; pp. 237–250.
28. Reybrouck, M. A Biosemiotic and Ecological Approach to Music Cognition: Event Perception between Auditory Listening and Cognitive Economy. *Axiomathes* **2005**, *15*, 229–266. [[CrossRef](#)]
29. James, W. *Psychology: The Briefer Course*; Harper & Row: New York, NY, USA, 1961.
30. Jones, M. *Time Will Tell: A Theory of Dynamic Attending, Online Edition*; Oxford Academic: New York, NY, USA, 2019.
31. Jones, M.R.; Bolz, M. Dynamic attending and responding to time. *Psychol. Rev.* **1989**, *96*, 459–491. [[CrossRef](#)]
32. Volgsten, U. The roots of music: Emotional expression, dialogue and affect attunement in the psychogenesis of music. *Music Sci.* **2012**, *16*, 200–216. [[CrossRef](#)]
33. Holck, U.; Geretsegger, M. Musical and emotional attunement: Unique and essential in music therapy with children on the autism spectrum. *Nord. J. Music Ther.* **2016**, *25*, 34–35.
34. Gabriellsson, A. Timing in music performance and its relations to music experience. In *Generative Processes in Music: The Psychology of Performance, Improvisation, and Composition*; Sloboda, J., Ed.; Clarendon Press: Oxford, UK, 1988; pp. 27–51.
35. Palmer, C. Mapping musical thought to musical performance. *J. Exp. Psychol. Hum. Percept. Perform.* **1989**, *15*, 331–346. [[CrossRef](#)]
36. Repp, B.H.; Iversen, J.R.; Patel, A.D. Tracking an imposed beat within a metrical grid. *Music Percept.* **2008**, *26*, 1–18. [[CrossRef](#)]
37. Jones, M.R. Time, Our Lost Dimension: Toward a New Theory of Perception, Attention, and Memory. *Psychol. Rev.* **1976**, *83*, 323–355. [[CrossRef](#)]
38. Drake, C.; Jones, M.R.; Baruch, J. The development of rhythmic attending in auditory sequences: Attunement, referent period, focal attending. *Cognition* **2000**, *77*, 251–288. [[CrossRef](#)] [[PubMed](#)]
39. Clarke, E. Timers, oscillators and entrainment. *Eur. Meet. Ethnomusic.* **2005**, *11*, 49–50.
40. Seemann, A. The Other Person in Joint Attention. A Relational Approach. *J. Conscious. Stud.* **2010**, *17*, 161–182.
41. James, W. *The Principles of Psychology, I*; Holt and Company: New York, NY, USA, 1890.
42. Theeuwes, J. Endogenous and exogenous control of visual selection. *Perception* **1994**, *23*, 429–430. [[CrossRef](#)]
43. Theeuwes, J. Top-down search strategies cannot override attentional capture. *Psychon Bull Rev.* **2004**, *11*, 65–70. [[CrossRef](#)]
44. Berlyne, D.E. Attention. In *Handbook of Perception: Historical and Philosophical Roots of Perception*; Carterette, E.C., Friedman, P.P., Eds.; Harcourt Brace Jovanovich: New York, NY, USA, 1974; Volume 1, pp. 123–147.
45. Neumann, O. Theories of attention. In *Handbook of Perception and Action. III: Attention*; Neumann, O., Sanders, A.F., Eds.; Academic Press: San Diego, CA, USA, 1996; pp. 389–446.
46. Posner, M. Orienting of attention. *Q. J. Exp. Psychol.* **1980**, *32*, 3–25. [[CrossRef](#)]
47. Sokolov, E.N. The neuronal mechanisms of the orienting reflex. In *Neuronal Mechanisms of the Orienting Reflex*; Sokolov, E.N., Vinogradova, O.S., Eds.; Wiley: New York, NY, USA, 1975; pp. 217–338.
48. Schmidt, R.; Carello, C.; Turvey, M. Phase Transitions and Critical Fluctuations in the Visual Coordination of Rhythmic Movements Between People. *J. Exp. Psychol. Hum. Percept. Perform.* **1990**, *16*, 227–247. [[CrossRef](#)] [[PubMed](#)]
49. Reybrouck, M. Music Shaped in Time: Musical Sense-making between Perceptual Immediacy and Symbolic Representation. *Rech. Sémiotiques* **2016**, *36*, 99–120. [[CrossRef](#)]
50. Stewart, M.L. The Feel Factor: Music with Soul. 1987. Available online: <https://www.mcavinchey.org/music-article/8363/details> (accessed on 16 April 2023).
51. Prögler, J. Searching for Swing: Participatory Discrepancies in the Jazz Rhythm Section. *Ethnomusicology* **1995**, *39*, 21–54. [[CrossRef](#)]
52. Clayton, M. What is Entrainment? Definition and applications in musical research. *Empir. Musicol. Rev.* **2012**, *7*, 49–56. [[CrossRef](#)]
53. Pantaleone, J. Synchronization of metronomes. *Am. J. Phys.* **2002**, *70*, 992–1000. [[CrossRef](#)]
54. Wiesenfeld, K. Huygens’s Odd Sympathy Recreated. *Societ. Polit.* **2017**, *11*, 15–22.
55. Geng, Y.; Liu, S.; Yin, Z.; Naik, A.; Prabhakar, B.; Rosenblum, M. Exploiting a Natural Network Effect for Scalable, Fine-grained Clock Synchronization. In Proceedings of the 15th USENIX Symposium on Networked Systems Design and Implementation, Renton, WA, USA, 9–11 April 2018; Usenix: Renton, WA, USA, 2018; pp. 81–94.
56. Bennett, M.; Schatz, M.; Rockwood, H.; Wiesenfeld, K. Huygens’s clocks. *Proc. R. Soc. Lond.* **2002**, *A 458*, 563–579. [[CrossRef](#)]
57. Thaut, M.; McIntosh, G.; Hoemberg, V. Neurobiological Foundations of neurologic music therapy: Rhythmic entrainment and the motor system. *Front. Psychol.* **2015**, *5*, 1185. [[CrossRef](#)]
58. Schmidt, R.; Turvey, M. Phase-entrainment dynamics of visually coupled rhythmic movements. *Biol. Cyber.* **1994**, *70*, 369–376. [[CrossRef](#)]
59. Roenneberg, T.; Hut, R.; Daan, S.; Mewes, M. Entrainment Concepts Revisited. *J. Biol. Rhythm.* **2010**, *25*, 329–339. [[CrossRef](#)]
60. Allgayer-Kaufmann, R. Identification the Entrainment Process by the Degree of Synchronization. *Eur. Meet. Ethnomusicol.* **2005**, *11*, 46–47.
61. Chang, A.; Bosnyak, D.; Trainor, L. Unpredicted Pitch Modulates Beta Oscillatory Power during Rhythmic Entrainment to a Tone Sequence. *Front Psychol.* **2018**, *7*, 327. [[CrossRef](#)]
62. Secora Pearl, J. Cognitive vs. physical entrainment. *Eur. Meet. Ethnomusic.* **2005**, *11*, 61–63.
63. Reybrouck, M.; Eerola, T. Music and its inductive power: A psychobiological and evolutionary approach to musical emotions. *Front. Psychol.* **2017**, *8*, 494. [[CrossRef](#)] [[PubMed](#)]
64. Rayleigh, J. *The Theory of Sound*; Dover Publishers: New York, NY, USA, 1945.

65. Winfree, A. *When Time Breaks Down: The Three-Dimensional Dynamics of Electrochemical Waves and Cardiac Arrhythmias*; Princeton University Press: Princeton, NJ, USA, 1987.
66. Stepp, N.; Turvey, M. On Strong Anticipation. *Cogn. Syst. Res.* **2010**, *11*, 148–164. [[CrossRef](#)] [[PubMed](#)]
67. Peña Ramirez, P.J.; Fey, R.H.B.; Nijmeijer, H. Synchronization of weakly nonlinear oscillators with Huygens' coupling. *Chaos* **2013**, *23*, 033118. [[CrossRef](#)]
68. Peña Ramirez, J.P.; Olvera, L.A.; Nijmeijer, H.; Alvarez, J. The sympathy of two pendulum clocks: Beyond Huygens' observations. *Sci. Rep.* **2016**, *6*, 1–6. [[CrossRef](#)]
69. Peña Ramirez, J.P.; Henk Nijmeijer, H. The Poincaré method: A powerful tool for analyzing synchronization of coupled oscillators. *Indag. Math.* **2016**, *27*, 1127–1146. [[CrossRef](#)]
70. Strogatz, S.H.; Stewart, I. Coupled oscillators and biological synchronization. *Sci. Am.* **1993**, *269*, 68–75. [[CrossRef](#)]
71. Gordon, C.; Cobb, P.; Balasubramaniam, R. Recruitment of the motor system during music listening: An ALE meta-analysis of fMRI data. *PLoS ONE* **2018**, *13*, e0207213. [[CrossRef](#)]
72. Prinz, W.; Chater, N. An Ideomotor Approach to Imitation. In *Perspectives on Imitation: From Neuroscience to Social Science: Mechanisms of Imitation and Imitation in Animals*; Hurley, S., Ed.; MIT Press: Cambridge, MA, USA, 2005; Volume 1, pp. 141–156.
73. Reybrouck, M. Musical Imagery between Sensory Processing and Ideomotor Simulation. In *Musical Imagery*; Godøy, R.I., Jørgensen, H., Eds.; Swets & Zeitlinger: Lisse, The Netherlands, 2001; pp. 117–136.
74. Brown, S.; Merker, B.; Wallin, N. An introduction to evolutionary musicology. In *The Origins of Music*; Wallin, N., Merker, B., Brown, S., Eds.; MIT Press: Cambridge, MA, USA, 2000; pp. 3–24.
75. Buhrmann, T.; Di Paolo, E.; Barandiaran, X. A Dynamical Systems Account of Sensorimotor Contingencies. *Front. Psychol.* **2013**, *4*, 285. [[CrossRef](#)]
76. Di Paolo, E.A.; Buhrmann, T.; Barandiaran, X.E. *Sensorimotor Life: An Enactive Proposal*; Oxford UP: New York, NY, USA, 2017.
77. Di Paolo, E.; Barandiaran, X.; Beaton, M.; Buhrmann, T. Learning to perceive in the sensorimotor approach: Piaget's theory of equilibration interpreted dynamically. *Front. Hum. Neurosci.* **2014**, *8*, 551. [[CrossRef](#)]
78. Repp, B.; Yi-Huang Su, Y. Sensorimotor synchronization: A review of recent research (2006–2012). *Psychon. Bull Rev.* **2013**, *20*, 403–452. [[CrossRef](#)]
79. O'Regan, K.; Noë, A. What is it like to see: A sensorimotor theory of perceptual experience. *Synthese* **2002**, *129*, 79–103. [[CrossRef](#)]
80. Todd, N.; Lee, S. The sensory-motor theory of rhythm and beat induction 20 years on: A new synthesis and futures perspective. *Front. Hum. Neurosci.* **2015**, *9*, 444. [[CrossRef](#)] [[PubMed](#)]
81. Torrance, S. The skill of seeing: Beyond the sensorimotor account? *Trends Cogn. Sci.* **2002**, *6*, 495–496. [[CrossRef](#)] [[PubMed](#)]
82. Silverman, D. Sensorimotor enactivism and temporal experience. *Adapt. Behav.* **2013**, *21*, 151–158. [[CrossRef](#)]
83. Koelsch, S.; Vuust, P.; Friston, K. Predictive Processes and the Peculiar Case of Music. *Trends Cogn. Sci.* **2019**, *23*, 63–77. [[CrossRef](#)]
84. Vuust, P.; Witek, M.A. Rhythmic complexity and predictive coding: A novel approach to modeling rhythm and meter perception in music. *Front. Psychol.* **2014**, *5*, 1111. [[CrossRef](#)]
85. Williams, D. Predictive Processing and the Representation Wars. *Minds Mach.* **2018**, *28*, 141–172. [[CrossRef](#)]
86. Friston, K. A theory of cortical responses. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **2005**, *360*, 815–836. [[CrossRef](#)]
87. Galantucci, B.; Fowler, C.A.; Turvey, M.T. The motor theory of speech perception reviewed. *Psychon. Bull Rev.* **2006**, *3*, 361–377. [[CrossRef](#)]
88. Sherwood, D.; Lee, T. Schema Theory: Critical Review and Implications for the Role of Cognition in a New Theory of Motor Learning. *Res. Q. Exerc. Sport* **2003**, *74*, 376–382. [[CrossRef](#)]
89. Vihman, M.M. Ontogeny of phonetic gestures: Speech production. In *Modularity and the Motor Theory of Speech Perception*; Mattingly, I., Studdert-Kennedy, M., Eds.; Lawrence Erlbaum: Hillsdale, NJ, USA, 1991; pp. 69–84.
90. Magill, R.A.; Anderson, D.I. (Eds.) *Motor Learning and Control: Concepts and Applications*; McGraw Hill: New York, NY, USA, 2018.
91. Schmidt, R.A. A schema theory of discrete motor learning. *Psychol. Rev.* **1975**, *82*, 225–260. [[CrossRef](#)]
92. Moens, B.; Leman, M. Alignment strategies for the entrainment of music and movement rhythms. *Ann. N. Y. Acad. Sci.* **2015**, *1*, 86–93. [[CrossRef](#)] [[PubMed](#)]
93. Richardson, M.J.; Marsh, K.L.; Schmidt, R.C. Effects of visual and verbal interaction on unintentional interpersonal coordination. *J. Exp. Psychol. Hum. Percept. Perform.* **2005**, *31*, 62–79. [[CrossRef](#)] [[PubMed](#)]
94. Schmidt, R.C.; Richardson, M.J. Dynamics of interpersonal coordination. In *Coordination: Neural, Behavioral and Social Dynamics*; Fuchs, A., Ed.; Springer: Berlin, Germany, 2008; pp. 281–308.
95. Klimesch, W. An algorithm for the EEG frequency architecture of consciousness and brain body coupling. *Front. Hum. Neurosci.* **2013**, *7*, 766. [[CrossRef](#)]
96. Klimesch, W. The frequency architecture of brain and brain body oscillations: An analysis. *Eur. J. Neurosci.* **2018**, *48*, 2431–2453. [[CrossRef](#)] [[PubMed](#)]
97. Canavier, C.C. Phase-resetting as a tool of information transmission. *Curr. Opin. Neurobiol.* **2015**, *31*, 206–213. [[CrossRef](#)]
98. Herrmann, B.; Henry, M.; Haegens, S.; Obleser, J. Temporal expectations and neural amplitude fluctuations in auditory cortex interactively influence perception. *NeuroImage* **2016**, *124*, 487–497. [[CrossRef](#)]
99. Palmer, C.; Krumhansl, C.L. Mental representations for musical meter. *J. Exp. Psychol. Hum. Percept. Perform.* **1990**, *16*, 728–741. [[CrossRef](#)]

100. Arnal, L.; Doelling, K.B.; Poeppel, D. Delta–Beta Coupled Oscillations Underlie Temporal Prediction Accuracy. *Cereb Cortex*. **2015**, *25*, 3077–3085. [[CrossRef](#)]
101. Fujioka, T.; Trainor, L.; Large, E.; Ross, B. Beta and Gamma Rhythms in Human Auditory Cortex during Musical Beat Processing. *Ann. N. Y. Acad. Sci.* **2009**, *1169*, 89–92. [[CrossRef](#)]
102. Large, E.W. Resonating to musical rhythm: Theory and experiment. In *The Psychology of Time*; Grondin, S., Ed.; Emerald: Bingley, UK, 2008; pp. 189–231.
103. Nobre, A.C.; van Ede, F. Anticipated moments: Temporal structure in attention. *Nat. Rev. Neurosci.* **2018**, *19*, 34–48. [[CrossRef](#)] [[PubMed](#)]
104. Andersen, L.; Dalal, S. The cerebellar clock: Predicting and timing somatosensory touch. *NeuroImage* **2021**, *238*, 118202. [[CrossRef](#)]
105. Rosso, M.; Heggli, O.; Maes, P.J.; Vuust, P.; Leman, M. Mutual beta power modulation in dyadic entrainment. *NeuroImage* **2022**, *257*, 119326. [[CrossRef](#)]
106. Betti, V.; Della Penna, S.; de Pasquale, F.; Corbetta, M. Spontaneous beta band rhythms in the predictive coding of natural stimuli. *Neuroscientist* **2021**, *27*, 184–201. [[CrossRef](#)] [[PubMed](#)]
107. Lee, J.H.; Whittington, M.A.; Kopell, N.J. Top-down beta rhythms support selective attention via interlaminar interaction: A model. *PLoS Comput. Biol.* **2013**, *9*, e1003164. [[CrossRef](#)] [[PubMed](#)]
108. van Ede, F.; De Lange, F.; Jensen, O.; Maris, E. Orienting attention to an upcoming tactile event involves a spatially and temporally specific modulation of sensorimotor alpha-and beta-band oscillations. *J. Neurosci.* **2011**, *31*, 2016–2024. [[CrossRef](#)] [[PubMed](#)]
109. van Pelt, S.; Heil, L.; Kwisthout, J.; Ondobaka, S.; van Rooij, I. Bekkering Beta- and gamma-band activity reflect predictive coding in the processing of causal events. *Soc. Cogn. Affect. Neurosci.* **2016**, *11*, 973–980. [[CrossRef](#)]
110. De Pasquale, F.; Della Penna, S.; Sporns, O.; Romani, G.L.; Corbetta, M. A dynamic core network and global efficiency in the resting human brain. *Cereb. Cortex* **2016**, *26*, 4015–4033. [[CrossRef](#)]
111. De Pasquale, F.; Corbetta, M.; Betti, V.; Della Penna, S. Cortical cores in network dynamics. *Neuroimage* **2018**, *180*, 370–382. [[CrossRef](#)]
112. Glass, L. Synchronization and rhythmic processes in physiology. *Nature* **2001**, *410*, 277–284. [[CrossRef](#)]
113. Colombetti, G. *The Feeling Body: Affective Science Meets the Enactive Mind*; MIT Press: Cambridge, MA, USA, 2014.
114. Varela, F. *Principles of Biological Autonomy*; North Holland: New York, NY, USA; Oxford, UK, 1979.
115. Varela, F.; Thompson, E.; Rosch, E. *The Embodied Mind: Cognitive Science and Human Experience*; MIT Press: Cambridge, MA, USA, 1991.
116. Thompson, E. Life and mind: From autopoiesis to neurophenomenology. A tribute to Francisco Varela. *Phenomenol. Cogn. Sci.* **2004**, *3*, 381–398. [[CrossRef](#)]
117. Thompson, E. *Mind in Life: Biology, Phenomenology, and the Sciences of Mind*; Harvard University Press: Cambridge, MA, USA, 2007.
118. Hanna, R.; Thompson, E. The mind-body-body problem. *Theor. Hist. Sci.* **2001**, *7*, 24–44. [[CrossRef](#)]
119. Torrance, S. In search of the enactive: Introduction to special issue on enactive experience. *Phenomenol. Cogn. Sci.* **2006**, *4*, 357–368. [[CrossRef](#)]
120. Thompson, E. Sensorimotor subjectivity and the enactive approach to experience. *Phenomenol. Cogn. Sci.* **2006**, *4*, 407–427. [[CrossRef](#)]
121. De Jaegher, H.; Di Paolo, E.; Gallagher, S. Can social interaction constitute social cognition? *Trends Cogn. Sci.* **2010**, *14*, 441–447. [[CrossRef](#)] [[PubMed](#)]
122. De Jaegher, H.; Di Paolo, E. Enactivism is not interactionism. *Front. Hum. Neurosci.* **2013**, *6*, 345. [[CrossRef](#)]
123. Overgaard, S.; Michael, J. The interactive turn in social cognition research: A critique. *Philos. Psychol.* **2015**, *28*, 160–183. [[CrossRef](#)]
124. Fiebich, A.; Gallagher, S. Joint Attention in Joint Action. *Philos. Psychol.* **2013**, *26*, 571–587. [[CrossRef](#)]
125. Fuchs, T.; De Jaegher, H. Enactive intersubjectivity: Participatory sense-making and mutual incorporation. *Phenomenol. Cogn. Sci.* **2009**, *8*, 465–486. [[CrossRef](#)]
126. Hutto, D.D. The limits of spectatorial folk psychology. *Mind Lang.* **2004**, *19*, 548–573. [[CrossRef](#)]
127. Ratcliffe, M. *Rethinking Commonsense Psychology: A Critique of Folk Psychology, Theory of Mind and Simulation*; Palgrave Macmillan: Basingstoke, UK, 2007.
128. Gallagher, S. Inference or interaction: Social cognition without precursors. *Philos. Explor.* **2008**, *11*, 163–174. [[CrossRef](#)]
129. Murray, L.; Trevarthen, C. Emotional regulations of interactions between two-month-olds and their mothers. In *Social Perception in Infants*; Fieldand, T.M., Fox, N.A., Eds.; Ablex: Norwood, MA, USA, 1985; pp. 177–197.
130. Auvry, M.; Lenay, C.; Stewart, J. Perceptual interactions in a minimalist virtual environment. *New Ideas Psychol.* **2009**, *27*, 32–47. [[CrossRef](#)]
131. Auvry, M.; Rohde, M. Perceptual crossing: The simplest online paradigm. *Front. Hum. Neurosci.* **2012**, *6*, 181. [[CrossRef](#)]
132. De Jaegher, H.; Di Paolo, E. Participatory sense-making. An enactive approach to social cognition. *Phenomenol. Cogn. Sci.* **2007**, *6*, 485–507. [[CrossRef](#)]
133. Reybrouck, M.; Podlipniak, P.; Welch, D. Music Listening and Homeostatic Regulation: Surviving and Flourishing in a Sonic World. *Int. J. Environ. Res. Public Health* **2022**, *19*, 278. [[CrossRef](#)]
134. Bannister, S. A survey into the experience of musically induced chills: Emotions, situations and music. *Psychol. Music* **2020**, *48*, 297–314. [[CrossRef](#)]

135. Grewe, O.; Kopiez, R.; Altenmüller, E. The Chill Parameter: Goose Bumps and Shivers as promising Measures in Emotion Research. *Music Percept.* **2009**, *27*, 61–74. [[CrossRef](#)]
136. Harrison, L.; Loui, P. Thrills, chills, frissons, and skin orgasms: Toward an integrative model of transcendent psychophysiological experiences in music. *Front Psychol.* **2014**, *5*, 790. [[CrossRef](#)]
137. Nusbaum, E. Listening Between the Notes: Aesthetic Chills in Everyday Music Listening. *Psychol. Aesthet. Creat. Arts* **2014**, *8*, 104–109. [[CrossRef](#)]
138. Vuoskoski, J.; Eerola, T. The Pleasure Evoked by Sad Music is Mediated by Feelings of Being Moved. *Front. Psychol.* **2017**, *8*, 439. [[CrossRef](#)]
139. Wassiliwizky, E.; Wagner, V.; Jacobsen, T.; Menninghaus, W. Art-elicited chills indicate states of being moved. *Psychol. Aesthet. Creat. Arts* **2015**, *9*, 405–417. [[CrossRef](#)]
140. Di Paolo, E. Behavioral Coordination, Structural Congruence and Entrainment in a Simulation of Acoustically Coupled Agents. *Adapt. Behav.* **2000**, *8*, 27–48. [[CrossRef](#)]
141. Merleau-Ponty, M. The child's relations with others. In *The Primacy of Perception*; Edie, J.M., Ed.; Northwestern University Press: Evanston, IL, USA, 1964; pp. 96–158.
142. Kojima, H.; Froese, T.; Oka, M.; Iizuka, H.; Ikegami, T. A Sensorimotor Signature of the Transition to Conscious Social Perception: Co-regulation of Active and Passive Touch. *Front. Psychol.* **2017**, *8*, 1778. [[CrossRef](#)] [[PubMed](#)]
143. Miyahara, K. Perceiving other agents: Passive experience for seeing the other body as the other's body. *Annu. Rev. Phenomenol. Assoc. Jpn.* **2015**, *31*, 23–32.
144. Torrance, S.; Froese, T. An Inter-Enactive Approach to Agency: Participatory Sense-Making, Dynamics, and Sociality. *Hum. Mentel* **2011**, *15*, 21–53.
145. Pearce, E.; Launay, J.; MacCarron, P.; Dunbar, R. Tuning in to others: Exploring relational and collective bonding in singing and non-singing groups over time. *Psychol. Music* **2017**, *45*, 496–512. [[CrossRef](#)]
146. Fodor, J.A. Methodological solipsism considered as a research strategy in cognitive psychology. *Behav. Brain Sci.* **1980**, *3*, 63–73. [[CrossRef](#)]
147. Durkheim, E. *The Elementary Forms of Religious Life*; Free Press: New York, NY, USA, 1995.
148. Boyd, B. *On the Origin of Stories: Evolution, Cognition and Fiction*; Belknap Press of Harvard University Press: Cambridge, MA, USA, 2009.
149. Jordania, J. *Who Asked the First Question? Origins of Choral Singing. Intelligence, Language and Speech*; Logos: Tbilisi, Georgia, 2006.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.