JNϕphi

Rhythm in Music, Encoded in Neural Networks, and in the Mind

Peter van Domburg

Abstract

Rhythm is ubiquitous in nature and has fascinated scholars from times immemorial. Rhythmic activity also underlies many forms of communicative interaction both in biology and in artificial computational systems. A rapidly growing issue, both in technology and philosophy, is whether this kind of communicative interaction from the most sophisticated applications of artificial intelligence (AI) is comparable to the interaction of human beings and their minds. A now historic debate on this quickly suffers from exceeding the limits that must be imposed on the use of terms from different reference domains, like the concept of intentionality and the emergence of conscious representations in a mental world. In this paper rhythm in music, with its characteristic roots in a culture, is explored as a representation of encoded information with particular Gestalt character, but meanwhile, in the composition of modulated frequencies, also comparable to the oscillatory activity in neural networks. Rhythm in music is a complex phenomenon and the carrier or "medium" of meaningful representations, while it can ultimately be traced back to modulated oscillations in sound waves, the auditory system and related sensorimotor and information supporting networks in the brain. The phenomenon of rhythm in music is explored, in such a way that it becomes clear why it can serve as an illustrative representation for the comparison of "intelligence" in the living brain and that in AI.

Key Words: Rhythm in music, neural networks, artificial intelligence, embodied cognition, encoded information, Gestalt, intentionality

Nothing but the beat?

Music is a uniquely human achievement and so is rhythm in music (RiM). To hear, enjoy and go along with musical rhythm is part of our "musicality" which is a typical human faculty, although some animals can learn it to some degree (Honing, 2012; Merchant *et al.*, 2018). Rhythm is ubiquitous in nature and has fascinated scholars from times immemorial. The way in which rhythmic activity can be modulated also forms the basis for the transfer of information in many systems, both natural and artificial. The well-known statement of the

Corresponding author: Peter van Domburg, MD, PhD
Address: Kapelweg 2, 6267 BW Cadier en Keer, The Netherlands
e-mail ✉ pvdomburg@icloud.com

communication theorist Marshall McLuhan applies to these cases: the medium [rhythm] is the message. Today, researchers agree that at the root of RiM lies the phenomenon of *beat induction*: the psychological extraction and allocation of a regular pulse or "tactus" (the beat) in an auditory sequence (particularly music) that permits synchronous tempo-flexible, responding to this pulse, even during intervals, irregularities or the absence of sound (Honing, 2012; Vanden Bosch der Nederlanden *et al.*, 2019; Greenfield *et al.*, 2021). In his overview Henkjan Honing noted: "the term *beat induction* is preferred here over *beat perception* to emphasize that a beat does not always need to be physically present in order to be perceived". Because beat induction is based on timed interaction with an outside source it is often lumped with "synchronization", with reference to the widespread appearance of synchronization of rhythms in nature. But, beat induction involves some "awareness" that is not required for many cases of synchronization in nature. This awareness refers to the engagement of supervising structures and may explain the multifaceted way in which music determines different cultures: synchronization than only applies to people who grew up or are familiar with that culture.

Extensive interdisciplinary research over the past few decades has yielded far-reaching insights into the evolutionary, biological and neuroscientific foundations of the phenomenon of beat induction and the meaning of RiM. Following a recent multidisciplinary overview, we will pragmatically continue to use the concept of 'beat perception and synchronization' (BPS) as an umbrella term, despite the fact that beat *induction* has been replaced by beat *perception* again (Greenfield *et al.*, 2021). BPS concerns a single non-ambiguous clearly structured part of our everyday experience with RiM that lends itself well to scientific research, because of its well-structured character. This necessarily excludes significant other components of timing, such as tempo (pace), meter, accent shifts and, most of all, the creative and fundamental coupling to pitch, timbre and melody. These latter components give a composition a contextual *meaning*, with a focus on subjective appreciation and recognition, rather than on interaction. These "mindful" representational or encoded aspects of RiM are much more difficult to access and somewhat underexposed in the neuroscientific literature. BPS generally is a driving force in music, but only paves the way to perceive, recognize and appreciate those more encoded elements in the appreciation of rhythm in a musical composition.

The concept of BPS implies that *rhythmicity* of serial signals triggers sensori-motor coupling in the central nervous system (CNS), both on a computational, psychological and an interactive behavioral level (Lenc *et al.*, 2021). The way in which subcortical modules in the CNS, like the basal ganglia and the cerebellum, play a leading part in this phase of perception is well established and recently reviewed (Damm *et al.*, 2020; Kasdan *et al.*, 2022). This automized phase of rhythm perception, apparently directed to synchronous interaction, also suits to substantiate the concept of "biomusicality" (Honing,

23

2018). Yet, at this phase of perception there is much more than just reflex-activity, at least in humans. As we know from our own experience, BPS often is accompanied by an as pleasant experienced "willingness" to move: the *groove* (Janata *et al.*, 2012). This comes about through the involvement of reward networks in the brain, facilitating the engagement of motor networks originally destined for interaction, even when more complex rhythms don't seem to serve this goal (Matthews *et al.*, 2020). Groove is already a sophisticated concept, that leads us into that inaccessible domain of an individual mental world, based on phenomena like *timing* (Ross and Balasubramaniam, 2022), *prediction and reward* (Matthews *et al.*, 2020), or *entrainment* (Lakatos *et al.*, 2019; Damm *et al.*, 2020). In summary, two features are central to BPS. First, the involvement of motor networks, through a direct but dynamic interaction between the sensory/auditory and motor system in the human brain. Second, BPS works as an operational component of music: it facilitates the emergence of more complex sensations and feelings, of which the groove is the most elementary, but not the only, variant. The emergence of these more complex *meaningful* sensations and their mental representations are the subject of this survey. Forced to skip many details from the extensive research into BPS, I have tried to present the findings in a preliminary representational form (figure 1), only to serve as a representative reference in our further explorations.
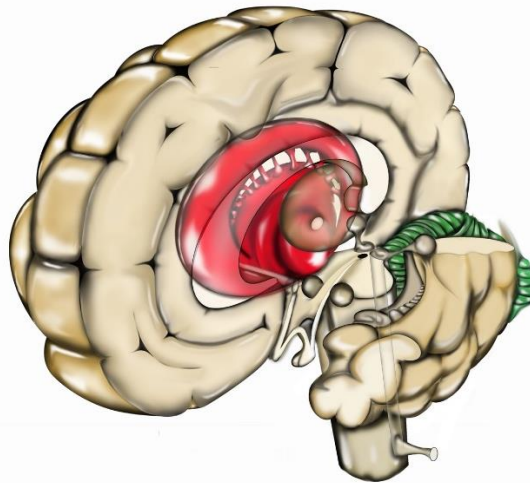
Figure 1. Stylized representation of modules involved in BPS, to be used as a reference

Encultured BPS and embodied cognition

As a lover of roots and folk music I try to broaden my horizons every now and then by getting acquainted with more exotic styles from distant cultures. To my surprise it is often hardly possible to experience anything like BPS when listening to this music, even though rhythm evidently plays a prominent role in it. Maybe it is because I am not trained as a musician? But should I be, if BPS is

determined predominantly biologically? Why don't I experience that attunement to that particular rhythm, while I immediately recognize it as "rhythm", as that familiar phenomenon in nature and in our own body (Couzin, 2018). An isochronous rhythmic signal easily attracts our attention, even shortly after birth, often followed by some rhythmic motor behavior, although without a clear adaptation to a change of pace in this stage (Ravignani and Madison, 2017). The ability to detect and react to an isochronous rhythm appears to be innate and a link between rhythm and the motor system has been shown in 7-months old's, long before they know it might be *music* what they hear (Trainor and Marsh-Rollo, 2019). The extraction of a regular pulse is learned automatically in the first 2-4 years, but keeping the meter has to be learned, similar to learning to use grammar in language. Our attention appears to be automatically drawn by the way our educators respond to such modulations, as will later be explained. It is similar to the way we intuitively appear to understand that someone else's use of unknown words reflects her intention to communicate, even before we now wat language is (Brandt *et al.*, 2019). In this way modulations in rhythmic sequences, determined by a particular socio-cultural context, are easily adopted in this stage of development, if they fit with the biological properties of our innate oscillatory network activity. A substantiation for this is provided by the uniformity of rhythm preferences in lullabies and simple children's songs across cultures (Mehr *et al.*, 2019). Many of these neural, behavioral and anatomical traits underlying rhythm perception and production are also shared with a broad range of species and can also be compared to the way vocalization patterns are achieved in other vocalizing animals during evolution (Wilson and Cook, 2016; Patel, 2021). In this way a kind of fingerprint in our neural networks is formed at this early phase of development and the basis is laid for the application of more complex modulations in the context of a particular culture. What sets the human species apart is their ability to acquire even more complex, particularly also abstract (like "aesthetically") encoded, rhythms of musical compositions, language and other formalized ways of communication (Pulvermüller, 2013; Stolk *et al.*, 2015; Dehaene *et al.*, 2015). This "rhythm syntax" is less extensively discussed in the literature, and yet it is precisely here where *meaning* presents itself as an elusive mental phenomenon and human singularity, that nevertheless has a unifying character beyond our bio musicality (Dehaene *et al.*, 2022).

When an isochronous rhythmic sound attracts the attention in newborns this will soon diminish, similar to the way we adapt to the sound of a clock or metronome. This is caused by the phenomenon of *adaptation* in isolated neurons in their response to a persistent unchanging stimulus (Wark et al., 2007). Although neural networks adapt quickly to an unchanging stimulus, they have a kind of memory for the timing of a rhythmic sequence, through "consolidation" (Buzsáki and Draguhn, 2004). That starts with the way we keep on

track in BPS, even during intervals or sound changes. Children memorize rhythmic sequences automatically from an early age and as adults we all know those unwanted "earworms" we experience when a catchy piece of music occupies our mind. One way to combat adaptation to an isochronous rhythm is to insert slight modulations, that attract our attention again. This offers the opportunity to incorporate *meaningful elements* out of the interactive socio-cultural context, as is also done in the acquisition of language. In the earliest phase of life, the consolidation takes place in the developing sensorimotor system, which now also becomes a *procedural memory* function. It will continue to play some automated leading role in the emergence of BPS in a particular socio-cultural context (Graybiel, 2008). This priming of sensorimotor networks, particularly those between auditory and motor cortex, might serve as a prototype for the way "culture" takes shape in our brain and can therefore be considered an epitome of *embodied cognition* (Pulvermüller, 2013; Ross and Balasubramaniam, 2022). A next phase of coding and consolidation, in later developing cortical association networks but co-determined by this priming phase, is required for our adult preferences and familiarity with a more complex rhythm syntax, as will be discussed below. This acquisition of new patterns or relevant changes in automated patterns becomes more difficult as we grow older, while the sensorimotor synchronization of BPS remains relatively well preserved in healthy aging and neurocognitive disorders (von Schnehen *et al.*, 2022; Sauvé *et al.*, 2022). That's why we have to be involved much more time in a strange culture to experience the interactive qualities of its rhythmic patterns, despite knowing it is the rhythm of their music what we hear now.

RiM disembodied: the meaning of rhythm

In a philosophical survey devoted to rhythm in its representative encoded form, Peter Cheyne calls this the "encoded rhythm", as the aesthetic counterpart of the embodied rhythm, as elaborated above (Cheyne, 2019). He tries to place these "encoded" aspects of rhythm, represented in a standard performance or musical notation, alongside "embodied rhythm", that unfolds in time. Encoded rhythm only can be represented in a disembodied way, contrary to the particular experience itself that only exists in the passing moment of conscious awareness. In order to meet the demands of aesthetics the composition of RiM has to be *objectified* in some external materialized form, like a musical score, an exemplary recoding or model performance. It seems paradoxical that the representation of our most intimate subjective experience, which we call "aesthetic", only exists in an objectified disembodied state. The aesthetic design represents our understanding of the way embodied cognition can take shape and not the embodied cognition itself. I have to disembody my idea of a piece of art and physically reshape it, before it can be used as an aesthetic cue for *your*

imagination. A musical score is essentially not intended, indeed, to be a straight reversal of the subjective experience. It leaves much room for our intuition and expressivity, as emphasized by Peter Cheyne. That is what makes it "my" music, in the passing moment of conscious experience of a performer or listener. At this very moment nature and nurture meet, creating a particular ambiguity, which we experience as a tension that only can exist in the living organic reality. In his survey "*Rhythm, preceding its abstraction*" of the same volume, Deniz Peters, professor for artistic research in music, addresses this paradox based on his own experience in performing music with other musicians: "Thought on rhythm usually begins where the phenomenon of rhythm itself ends - at the point where it turns into a representation, at the fringes of its experience" (Peters, 2019). He evaluates the *experience* of rhythm in its "living origin, the flow" whose identity (its "Gestalt", as will be shown) is as stable as its form is shapeable. He points to the emotional satisfaction he derives from the malleability and ambiguity of an encoded rhythmic syntax, which serves as a cue for improvising in harmonic interaction with his fellow musicians. This experience goes far beyond that of BPS, adding a creative dimension that makes it tangible how you can "capture an intricacy that is always on the move" (Cheyne, 2019). A passive listener might likewise be taken along on the waves of such an unsettled flow, directed by his or her imagination, and an ambiguity in which chance and natural law go hand in hand. This particular experience of RiM can only come about in that unique conscious awareness. In this respect we are all musicians, as Henkjan Honing says.

In an attempt to overcome the dilemma of mind-body dualism aesthetically designed "encoded rhythm" is often also considered as a form of "embodied cognition", but I think that is precisely not the case. Paradoxically it is out of the body, *not* in some esoteric mental space, but in the physical form of formalized notations or unspoken (cultural) agreements, as metaphors, symbols, concepts, images or some particular performance of music. Our fleeting conscious experience of encoded rhythm only can be objectified in the domain of aesthetics, art, music psychology or philosophy. It only is embodied at its inaccessible awareness, in conjunction with the psychomotor response of BPS. It emerges from ongoing neural network activity, triggered by that external cue or possibly reconstructed from the inside as a symbolic representation (Stolk *et al.*, 2015; Dehaene *et al.*, 2022). Even for an experienced musician like Deniz Peters there is no room for "thought on rhythm" or some external representation during improvisation with his fellow players. He is triggered by it (in recordings, musical scores or the sound of fellow players), as an incentive for his ongoing neural network activity which then goes its own way. Musical scores, recordings, the gestures of a concert conductor and the sound of fellow players can all be suitable as an incentive stimulus for our evolutionary specialized senses that will guide these representations into neural network activity.

From aesthetic coding of RiM to "neural resonance"

The focus of a treatise on music generally is its *aesthetic* appreciation, in which BPS and groove play a subordinate role. This aesthetic dimension places different demands on the rhythm than BPS, such as a more complex encoded arrangement and the integration with other compositional elements of sound. Rhythmic events must creatively be arranged and grouped in complex patterns that we have learned to understand, recognize and enjoy as a characteristic aspect of music beyond the incentive to move (Fitch, 2013; Pouw *et al.*, 2021). In this way a *rhythm syntax* emerges, that now constitutes the aesthetic character of music as a human singularity. The information it provides is generally represented in abstract concepts such as meter, tempo, adagio, rubato and various musical notations. Yet this representation only serves as an aesthetic model that must be freely expressed in a performance or during recordings, in which the composer and musician are focused on the interaction with a (potential) listener, or when they themselves might be the listener in a jam session. Unlike its formalized pattern in an abstract encoded form, the execution of a rhythm syntax requires a flexible ongoing tuning to the listener to achieve the intended aesthetically defined goal in a particular musical context. That context is co-determined by the listener, with its inaccessible subjective experiences of recollection, pleasure, admiration, relaxation, inspiration or reminiscence. The sensorimotor coupling of BPS here makes way for a subtle sensing of the experience of the listener, for whom the shape of a calm, cheerful, sad or jubilant rhythm might be required, or even must be impressed with the unique skills of the musician. The paradox of RiM is that it is fixed in a composition, while it can be applied differently each time.

In terms of neurophysiology a composer or musician (the *sender*) actually aims to align the naturally occurring ongoing oscillatory activity in the cortical association networks of a listener (the receiver), to an aesthetically encoded model that already has proven its value (Müller *et al.*, 2022). The experience of RiM reflects the attunement of ongoing oscillatory activity in cortical association networks with the frequency spectrum of the rhythmic syntax provided, similar to the entrainment of sensorimotor networks in BPS. Because of the relaxing effect it has on our mind this attunement must require a *minimal* amount of energy for an *optimal* experience. The activity in our cortical association networks that lead to the awareness, recognition and appreciation of RiM, can therefore be interpreted as "resonance", even if we do not yet have the means to record this exactly (Ravignani and Madison, 2017). Of course, the concept of *resonance* is somewhat misleading. Although it originally stems from the domain of acoustics it generally refers to the physical interaction of objects with equal vibration frequency in a stable context (Glass, 2001). As a conceptual analogy it is useful however for several reasons (Large et al., 2023).

Resonance refers to the natural coexistence of oscillatory activity in two different objects that have particular properties in common. It also can be the consequence of the way these objects have been *artificially shaped*, of which the resonance of musical instruments is prototypical. Moreover, the patterns coincidentally caused by resonance can emphatically appeal to our sense of aesthetics, of which fascinating examples are easily found on internet. Nature also provides many examples of unexpected "creative" pattern formation based on automated interaction in animals and living systems, which lend themselves to comparison with rhythmic syntax and the aesthetic appreciation it can evoke (Couzin, 2018). Artists and particularly also musicians often are inspired by such patterns, apparently appealing to their imagination. Although RiM is not adopted by our CNS through physical resonance as such, internal oscillatory activity in neural networks of the perceiver somehow appears to be tuned to an externally composed frequency spectrum at that very moment.

It takes a lot of intermediate steps before RiM is experienced as such, like the transitions in a complex auditory system and the algorithm of prediction, which we will return to later. But in the end, there is a form of attunement of complex frequency modulated patterns, by which some external source and internal neural network activity come to interact, which also fits with the way we experience it. What matters is that the computational nature of a rhythmic syntax, ultimately based on the modulation of frequency patterns, somehow corresponds to the ongoing computational nature of frequency modulation in neural networks. For the time being and with some reservations, this opens the way for research into the way an external abstract (aesthetically) coded rhythmic pattern correlates with the embodied pattern of internal neural network activity, as a model for our interactive communication. This form of communication is so obvious that we might forget the fact that it can only be expressed for the time being in concepts derived from the domain of psychology or philosophy. The concept of neural resonance is also derived from another, physical, domain, but it opens the way to a falsifiable model.

Encoded neural network activity also does not explain the way in which mental images or experiences emerge in our consciousness (Roy *et al.*, 2018; Brette, 2019). This brings us into the debate about "emergence", which still lacks a sufficient explanation, despite many attempts (Feinberg and Mallett, 2020). Yet, although the *awareness* of RiM is crucial, because it in turn determines the way in which some artificial external syntax or any piece of art must be composed in order to serve its purpose, it is legitimate to sidestep the debate about its emergence. The purpose here is only to explain the way an image is *represented in neural network activity* and *not* its mental content nor the epistemology of a particular conscious awareness. Composers, performing musicians or DJs are in search for the best possible external aesthetic representation in scores, concepts, recordings, metaphors, symbols, images and, of course, in their own intentions

29

and activities. For music and other arts such external representations are called "aesthetic" if they succeed in that purpose. In a multitude of symbolic resources that human beings have at their disposal, RiM stands out as a congeneric incentive for the modulation of neural network activity. The advantage of RiM as a carrier of information is that the symbolically represented information itself is also based in a pattern of rhythmic activity, which is a categorical similarity to oscillatory activity in neural networks. With RiM we can avoid the common pitfall of expressing the *meaning* of an outcome or any finding from research in terms from another conceptual domain (Bennet and Hacker, 2003).

From neural resonance to shared intentionality /communication

For neural resonance to take place with de neural networks of a receiver, the performing musician or composer (the sender) somehow has to "know" that he/she is on the right track, during a performance or (potentially) while composing. Neural resonance only occurs if the necessary conditions are met with in both the rhythmic syntax of a representation, the sender and in the oscillatory activity of cortical association networks of the recipient, who only subsequently can confirm its experience. The inaccessibility of the content of another's mind was already a reason for Clemens von Brentano to introduce the concept of "intentionality" in the late 19th century and became adopted in the domain of communication theory (Jacob, 2023). It means that we have the ability to deduce the *intentions* of others from their behavior, as a reflection of their thinking. This faculty also has fascinated neuroscientists, particularly also after the identification of so called "mirror neurons", characterized by the expression of electrophysical activity that corresponds to that of targeted actions of an observed conspecific (Bonini *et al.*, 2022). A similar feature has also been attributed to a population of "spindle neurons", particularly well developed in the fronto-insular and anterior cingulate cortex of humans, the disruption of which is associated of the loss of a "theory-of-mind" ability in frontotemporal dementia and autism (Allman *et al.*, 2011). A further discussion of this is beyond the scope of this paper, but briefly *intentionality* now can be conceptualized as a correlate of disrupted activity in certain network modules that also underly our ability to understand other person's intentions. It is considered a further step in evolution underlying the sophisticated communication abilities of the human species through *shared intentionality* (Tomasello, 2008; Pouw *et al.*, 2021). Regardless of the existence of a particular class of neurons, it is tempting to regard the concept of shared intentionality as the phenomenological correlate of neural resonance. Tailoring and understanding rhythm syntax by a sender and a receiver in music might be based on this shared intentionality.

30

An example from our clinical experience is provided by people suffering from Parkinson's disease, in which the understanding of rhythm syntax is preserved, while the BPS component of RiM now is disrupted: a substantiation based on *double dissociation*. People suffering from Parkinson's disease generally are *not* restricted in their understanding of the *intentionality* of others or the understanding and appreciation of an aesthetically encoded rhythm syntax in music (Bellinger *et al.*, 2017). But attuning to an isochronous beat, which is the condition for the development of BPS, is disrupted due to impairments in their sensorimotor networks. Therefore, ideally, cueing paradigms for the improvement of gait and posture ideally should be based on dynamic (non-linear) interaction instead of cueing based on a steady beat. To be carried away in "the flow" of RiM requires more than just an internal perception of time. Rhythm syntax might give RiM the image of some "art object", that takes over our imagination or that of the composer and performer. For this aspect of rhythm other non-motor modules in the CNS are required that nevertheless must be intertwined with those of BPS. This intertwining enhances the appreciation, as Peter Cheyne tried to show in his approach to the interweaving of thinking, feeling and doing in music (Cheyne, 2019). We are still far from understanding exactly the underlying networks and computations for this interrelatedness, even now that the neural basis for BPS has been established. Previous research into the underlying networks in *language* have posed a similar challenge, about which a great deal of knowledge now exists that can serve as a guide. Our faculty of language is believed to rely on a similar interdependence of innate and acquired traits (Berwick *et al.*, 2013). Insights that have emerged from this human singularity of language also came to shed new light on the possible overlap between music and language, with regard to their role in communication, cognitive, emotional and social development and the generative systems in the CNS that underlie it (Peretz *et al.*, 2018, Schön and Morillon, 2019).

The compositionality of rhythm syntax and language

With his name making concept of an innate faculty of language Noam Chomsky challenged the concept of language as a purely *acquired* (cultural) ability in the 1960s (Hauser *et al.*, 2002; Chomsky, 2017). It has been the subject of ongoing debate and much interdisciplinary research (Petkov and Marslen-Wilson, 2018). One reason for Chomsky and his followers to argue that the principles underlying the generation of language are biologically determined is that children acquire the underlying rules semi-automatically through regular exposure, resulting in a large amount of implicit knowledge (Brandt *et al.*, 2019). In retrospect, Chomsky unintentionally brought language already closer to music, with the assumption of a biologically determined grammar faculty, in which some innate universal

generative computations are engaged (Hauser *et al.*, 2002). Both language and music seem deceptively trivial in the way they create a virtually limitless variety of complex hierarchically structured sequences with an abstract content, representing objects, thoughts, feelings, experiences and other phenomena out of a limited number of (linguistic or musical) elements. The computational algorithm employed in these patterns is based on the principle of *recursion*: the generation of hierarchically build tree structures of increasing complexity out of discrete elements (like phonemes, words or word groupings), into an expressive composition (the syntax) that is understandable to others sharing similar knowledge (Zuidema *et al.*, 2018). Marc Hauser and Jeffrey Watumull argued that this algorithm of recursion might underly a more universal generative faculty, realizing not only the building of syntax in language and music, but also other human singularities like mathematics and morality (Hauser and Watumull, 2017). This model gained additional significance when it could be shown that lesions in the so-called area of Broca lead to language disorders with *agrammatism* (Friederici, 2023). The extent to which it determines a universal grammar, also central to the composition of syntax in rhythm and music, is still up for debate. Its value lies mainly in the theorem of the computational mechanism in neural networks that also underlies the composition of more complex symbolic representations (Dehaene *et al.,* 2022). As a human singularity it explains the inability of non-human creatures to generate and understand encoded patterns like those in music and language, whereas they are able to communicate based on innate context related auditory signs.

32

Debates about a biologically based algorithm, like that of a universal grammar, usually lead to the question of how this could have developed as an evolutionary adaptation. Some of the underlying computational networks, required for the discrimination and production of ordered sound sequences with a meaning destined for communication, are innate and shared with other animals and might have the same evolutionary roots (Lenc *et al.*, 2021). Vocal learning in songbirds and some other animals seems to be a suitable prototype (Patel, 2021). Such animals learn to use rather complex structured meaningful vocalizations, for instance to attract or warn their peers. But it is hardly thinkable to provide an external abstract (disembodied) representation of this, that can be applied to a fellow member. It is even less likely that such an abstract representation could be tailored in use to a specific individual in a certain mood or situation. Like the difference in the meaning of my expression "I love you", if directed to my wife, my grandma, my dog or even my smartphone. Likewise, a musician has to map a rhythmic design in his or her performance in a way that not only leads to the experience of BPS (our bio-musicality), but also that extra meaningful dimension, that allows a message to be conveyed that can be tailored to individual listeners in a particular socio-cultural context. With that music has served a successful

sophisticated *communicative* purpose, which in turn has contributed to the further development of underlying networks and cognitive development based on the interaction in social networks (Pinker, 2010; Hodges, 2019). It is ontologically replicated in the way we acquire language and music in our individual development, as a culturally determined, but also vital, faculty (Brandt *et al.*, 2019).
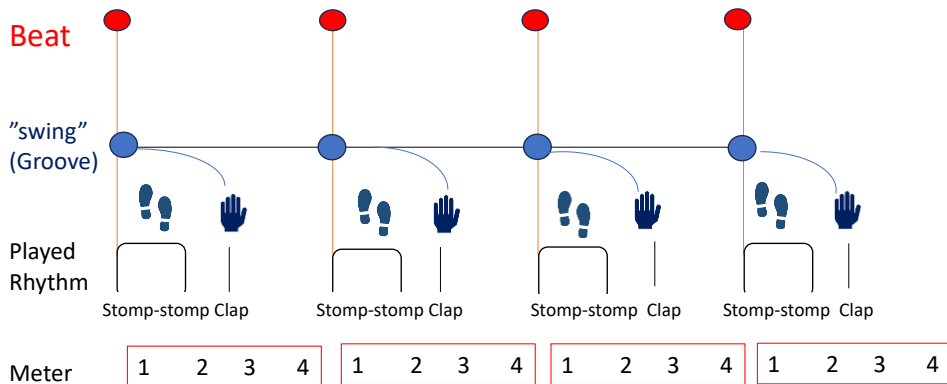

Figure 2. Syntax tree of the well-known introduction rhythm of "We Will Rock You" from Queen. The first "stomp" (footstep) with the feet of the musicians, where the melody will later start, is the start of the beat. The "clap" with the hands, that fills the gap between two beats, creates a feeling of continuity and "swing" in a meter pattern of 4/4.

33

While language is viewed as a formal communicative system of ideas and propositions, music focuses on sharing emotions, comfort, appreciations, beliefs and cravings, carried on the driving force of BPS. As with language this faculty reaches maturity at a critical early stage of development, at a time when the boundary between the (particularly also rhythmic) properties of both the faculties of language and music is still vague (Brandt *et al.*, 2019; Mehr *et al.*, 2019). For its full development, both healthy early development and a social context are required, in which biology and culture go hand in hand (Clayton, 2019). The question remains not only which additional hardware allows for the creation of a grammatical structure and meaningful syntax but particularly also how this can be mapped to both the sensory-motor (phonological) and a formal conceptual (semantic) interface. Particularly RiM can serve as a model for this unexplained way of mapping, because it is made up of a modulated frequency spectrum, process-wise comparable to the natural occurring oscillatory patterns in cortical association networks. The elements in the syntax of encoded rhythm in music are related in *time,* whereas those in a linguistic syntax are determined by their structural relations, for which time is not the primary determinant. It lends itself not only to exploring the way in which encoded rhythm can be

creatively integrated in the "living flow" of the musical experience, but particularly also for the way its external, disembodied, representation can be "understood". With the emergence of sophisticated imaging and physiology recording techniques, it has been shown that certain cortical association networks in the CNS, characterized by a high degree of plasticity, are involved in the computational process underlying both the generation and understanding of syntax in our ability of intelligible communication (Elimari and Lafargue, 2020).

Deep learning and the tuning of neural networks

For most of us amateur musicians it takes a lot of effort to become familiar with an unexpected arrangement or style of music, either when playing our listening to music. Some "thought on rhythm" is required now, before the unfolding in time can lead to the embodied experience of understanding, enjoinment or inspiration. It requires effort and adjustment to prime the networks involved in the abstract representation of a rhythm syntax, before they can "capture that intricacy that is always on the move" (section 3). The rhythm syntax of any composition in any culture might be analyzed, learned, compared, reproduced or even used as a basis for a new design for composers. It has been shown recently that even an artificial neural network architecture can be trained for this kind of "meta-learning for compositionality", even if we do not exactly know the acquired computations underlying it (Lake and Baroni, 2023). But with this development, history seems to repeat itself: Soon after Paul Broca identified the brain region involved in the composition of a grammatical syntax in1861, researchers like Carl Wernicke, Ludwig Lichtheim and also Sigmund Freud, pointed out that this doesn't explain the *understanding* of that syntax. This sparked a now historical debate about the additional modules and network needed for the understanding of language (Eling and Whitaker, 2022). Current AI applications are not so much limited in their compositional abilities as in processing ambiguous symbols or information that mainly derives its meaning from interaction with someone (or something) else. For the effective exchange of information in a communicative process, bidirectional and often multi-level, attunement between a sender and receiver is required. This brings us into a domain where theoretical biology and philosophy meet.

"Being in-the-world" instead of "Knowing the world"

When animals are born, they already have some "knowledge" (instinct) of the world around them. This will be supplemented with examples of their predecessors if they have the resources, like a nervous system, that is receptive to this. The term "Umwelt" (experienced outer world) was coined for this by the biologist Jakob von Uexküll (1864-1944)

and became adopted in theoretical biology, to represents the integrated whole of the sensory cues relevant for survival of a species in its unique habitat. This "Umwelt" is contrasted with an "Innenwelt" (inner or organic world) determined by the physio-chemical properties and needs of the organism itself. In the beginning of evolution this "Innenwelt" coincided with the outer living world. In eukaryotic organisms this physico-chemical information became encoded in the DNA of genes, leading to an innate encoded Innenwelt. Gradually the growing complexity, mobility and interaction of organisms led to the emergence of a sensory-nervous system, in which newly acquired information from the habitat could be incorporated, in addition to the genetically encoded Innenwelt. The nervous system now became the primary task of comparing the Innenwelt and Umwelt and adjusting an appropriate action accordingly. While in animal's instinct innate "knowledge" dominates over acquired features of their Umwelt, in the human species this relationship is reversed: an artificial culture has increasingly determined our human Umwelt. Newborn children have to learn a lot of things if they want to survive, for which they have an extremely extended postnatal period of complete dependence and ignorance. This became an important theme of Uexküll's contemporary Adolf Portmann (1897-1982) who explored the idea that as human beings we are born in a state in which even the awareness (an instinct) of a living environment still needs to take shape. Whereas animals can rely on instinctive behavior that meets the demands of nature itself, our living environment is ever more determined by the artefacts of a culture. Even our most basic categorical concepts, like "time" and "space", have to be acquired before we can put them into practice. A significant part of our brain consists of modules intended for the analysis and integration of incoming sensory information, allowing behavior to become guided by accurate inferences about the external world (Sheya and Smith, 2019). Our Umwelt is not just a knowledge of an unchanging world, but an ever changing "being-in-the-world" and its socio-cultural context, in the expression of Maurice Merleau-Ponty.

In an early phase of development children prefer to interact synchronous to auditory cues. It is much easier to let the "embodied cognition" of primordial BPS do its job, because it is founded in a biological predisposition, realized in the innate achievement of auditory association networks. Gradually our auditory system becomes tuned to the culture, that will define our Umwelt experience, and its "auditory scene" (Trainor, 2018). The development of our musicality in evolution went hand in hand with the natural tuning of a multi-level auditory system, similar to the way language and its benefits exerted a selection pressure on the development of appropriate modules in the CNS. We have gradually come to understand and respond to an increasingly complex auditory scene, instead of reacting to it instinctively. "Understanding rhythm as being generated in such situations offers a route to understanding the links

between social situation, cultural context and rhythmic structure, an area which has suffered from being limited to simplistic homology theories" (Kotz *et al.*, 2018).

Time and space

To attune to a rhythm syntax, we need to have some knowledge (perhaps incorporated in a recursive tree structure) of its unfolding in *time*. The question now is: "how does the brain encode temporal sequences of items, such that this knowledge can be used to retrieve a sequence from memory, recognize it, anticipate on forthcoming items, and generalize this knowledge to novel sequence with a similar structure?" (Dehaene *et al.*, 2015). Briefly stated: how do we *know about the timing*, if not by the interaction with an ongoing sequence of events and their duration? We must realize that even our awareness of "time" can only exist in an externally encoded form, like *units* on a clock or some time-scale or in relation to the duration of events in our living environments (Eichenbaum, 2017; Buzsáki and Tingley, 2018). We are all raised with the conceptual framework of classical physics, based on the categories of space and time. These categories are assigned to the natural course of things, to gain control and arrange events from the past into the future. The image of a universe in which place and time are the categorical constitutive elements took root in a long history from Euclides, Galilei and Kant to the design of classical Newtonian physics. But after all, it was culture and not nature that provided us with these concepts. Space and time even appeared to be interrelated, as was shown in the spacetime model of general relativity in the 20th century. Nature only provides us with a pulsatory course of events, to which we might synchronize (Drayton and Furman, 2018). Natural events, indeed, don't take place in this kind of a "theater", where they unfold on a timeline. They just take place…. in an undefined, yet pulsatory course. This pulsatory course lies at the root of *synchronization*: a closely related natural phenomenon, that has been the topic of much research in the physical, neurological, psychological and social domain (Couzin, 2018; Greenfield *et al.*, 2021). Our sense of time is based on oscillatory activity in neural networks, that we share with everyone and therefore can provide an intuition of commonality (Schirmer *et al.*, 2016). Time and space can only be represented in the encoded form of *distance and duration,* that we internalize as tacit knowledge (Buzsáki and Tingley, 2018). A particular brain module, the hippocampus and its related memory network, underlies the generation and experience of sequence patterns, both in time and spatial relationships (Billig *et al.*, 2022). A well-known example is that of the structural changes in the hippocampi of London taxi-drivers that evolves through the acquired spatial knowledge of a complex environment, but without some representation on a map (Maguire et al., 2000). Apparently, a rhythm

36

syntax might be mapped in the same way, which only comes to expression in its use.

Many animal species are familiar with time intervals between sensory and motor events and use these temporal representations in simple computations. Yet we do not assume that they have any sense of "duration" or the meaning of such intervals on a time *scale*. Their neural networks simply synchronize with ongoing sings and events. It is a matter of interactive timing by neural network activity, which is a common phenomenon in nature. Natural processes generally have a pulsatory course indeed. In a solid review on this subject, Leon Glass states that "these rhythms arise from stochastic, nonlinear biological mechanisms interacting with a fluctuating environment" (Glass, 2001). "Physiological rhythms are rarely strictly periodic but rather fluctuate irregularly over time.... there is continual interaction between the environment and internal control mechanisms." They wax and wane, inspiration and expiration, acceleration and deceleration, arsis and thesis. Even at a basic level an acoustic sequence first has to synchronize in some way with the dynamics of our auditory system and remodeled into neural oscillations and network activity, before we may experience it is the beat that allows us to bob our heads or clap our hands in time to the music (Rosenblum and Pikovsky, 2003). Because the processing of information about RiM is based on ongoing oscillatory activity, there is an analogy between the ontology of encoded disembodied RiM and that of the temporal flow of information in the brain: "the music of thought" (Müller *et al.*, 2022). We have to grasp the abstract nested tree structures of some encoded rhythmic pattern, devised by the composer or a fellow musician, and integrate it with the already present feeling of a beat. In the unfolding of this encoded rhythm a tension can be created by going slightly out of step with the anticipated items, such that the resolution of this aberration will lead to a feeling of reward in the listener. This always involves the flexible use of time intervals, while maintaining an overview that can be anticipated and adopted. This is the moment when culture takes over control from nature.

Our brain predicts, it does not command

So far, our model of information processing in neural networks focuses on the oscillations of interacting neurons, with their on-off signals. Extensive research has shown how this information carried by endogenous low-frequency neuronal oscillations in turn modulate the excitability of task-relevant neuronal populations (Canolty and Knight, 2010; Cebolla and Cheron, 2019). In psychological terms this leads to a *preparedness*, which can be objectified and measured as an electrophysiological signal: the "readiness potential" (Schurger *et al.*, 2021). It indicates that a selection has been made out of the complexity of an individual repertoire of potential behavior patterns, based on the

predicted best chance of success. This selected behavior pattern in turn appears also to predict the occurrence of the *next sensory event* out of a sequence, in which the cerebellum plays a leading role (Morillon & Baillet, 2017; Damm et al., 2020). This again seems to be a description in anthropomorphic terms, but it is all about RiM, ultimately also determined by rhythmic sequences, the format of which corresponds to that of the activity in neural networks. Moreover, it is not a matter of "storing information" about the state of affairs, but the brain constantly compares incoming signals with desired modes and capabilities of the body itself (Clark, 2016). The auditory system is also an outstanding example because it can "generate evoked responses to an absent but *expected* stimulus" (Wacogne *et al.*, 2011, cited in Clark, 2016). Prediction errors lead to "surprise" which represents a tension in this "predictive system" of our CNS, that can turn into a feeling of reward if the signal can still comply with the expected pattern (Petter, 2018; Cheung *et al.,* 2019). Reward, again, is an anthropomorphic concept representing the pleasant feeling of *reinforcement,* but it is also reminiscent of the rapidly increasing amplitude of a resonating sound from a speaker in the proximity of a microphone. In our brain this feeling of reward is represented by the release of, particularly dopaminergic, neurotransmitters by the organism: the basic requirement for the activation of purposeful motor programs, as we have learned from patients with Parkinson's disease. We also are familiar with this, when we go through the pleasant inclination to move by the *groove* of music (Morillon and Baillet, 2017). Prediction errors, within the boundaries of a culture-determined alignment, can cause a tension that affects our mood and intention to move (Fitch, 2016). Based on research with syncopation and polyrhythm as prototypes of rhythmic complexity, a framework of the (both learned and innate) limitations within which this occurs, has been outlined (Matthews *et al.*, 2020). This field of dynamic tension between a particularly external encoded and internal embodied rhythm in music can be creatively explored and shaped in de musical score (like the meter and changes of tempo) to be freely explored again in the performance (particularly in jazz) and the participation, like dance. With the encoded rhythm composers and performers seek to push the natural boundaries of timing and prediction, creating an expanded domain we call "aesthetic".

RiM as an auditory object: The ventral "Gestalt" stream

In parallel with the development of the first connectionist (network) model needed to explain both the composition and understanding of a language and its disorders, at the end of the 19th century, Christian von Ehrenfels (in the footsteps of Ernst Mach) introduced the concept of "Gestalt". This concept of Gestalt is based on the experience that a *melody in music* can be expressed in infinite ways and yet be understood as that same melody. It must exist as some abstraction,

beyond any concrete representation: "As soon as one is committed to the idea that something other than the sum of the tones makes up the melody, one has in effect accepted what we call the tonal Gestalt" (Robin and Ierna, 2022).[2] Rhythm syntax takes on the character of a particular Gestalt that must take shape in time and not in space: it is "formative" rather than a "form". Compare it to "craving", which we all know from experience as a "formative" drive so easily triggered by a lot of symbolic representations in the external world, without a final realization preventing it from taking over us again. It can also be compared to the way our "beliefs" and the emergence of symbolic representations in turn have contributed to a further development of association networks in our brain, in the course of evolution (Seitz and Angel, 2020). It was not without reason that the concept of "Gestalt" was based on the model of music and not language, with its less ambiguous syntax. There is something magical about the way music takes hold of us, perhaps because, more than with language, our entire body seems to be involved. Particularly with music, it is clear that although our CNS is needed for the composition and perception, this CNS is still at the service of our entire organic being. The operation of triggers with Gestalt features can apparently not simply be reduced to that of a structured information transfer.

Additional research has now provided sufficient evidence for the existence of a "ventral stream", in which modules, particularly also in the associative auditory cortex, play a role in the understanding of language but also in the processing of other complex auditory information (Sridharan *et al.*, 2007; Weiller *et al.*, 2022). The processing of rhythm syntax in music is similar but not reducible to the processing of syntax in *linguistic* trees, for which both a "dorsal constructive" stream and a "ventral receptive" stream have been conceptualized (Bornkessel-Schlesewsky and Schlesewsky, 2013; Hickok, 2022). The Gestalt features of RiM are much more pliable than the propositional features of an expression in language (with exceptions in poetry or lyrics). The incoming data in music are highly variable, surrounded by noise, context dependent and deliberately kept ambiguous, while this has to be avoided with language. The signals come from different and interacting sources such as musical instruments, voices, pitch changes or accents and even silent gaps in a particular pattern. It also has to be performed in a changing context, with different instruments and different interacting performers. For example, Beethoven's "Ode to Joy", is part of his symphony No. 9. referring to a poem of his contemporary Friedreich Schiller in which he expressed his inspiration of a brotherhood in the making. Through this Gestalt feature it later became the driving melody, with its appealing cadence, to symbolize the European Union. There is an "idea" behind this, different from the representation in the musical

---

[2] The concept of "Gestalt" would later also play a role in Brentano's concept of intentionality, referring way we understand an inaccessible mental image by the way it is expressed.

notation or some performance. Some leaps of thought are needed to understand what is meant by the wording on the EU website: "In the universal language of music, this anthem expresses the European ideals of freedom, peace and solidarity". Unfortunately, this goal only exists in our world only as an idea, a Gestalt of which we know all too well what it refers to.

Our ability to understand representations with Gestalt features, without being aware that we do so, is in fact the fruit of millions of years of evolution. In our individual life, a lifelong tuning of neural network activity must take place, from the cochlear membrane to cortical association networks, in which nature and nurture go hand in hand. Nowadays many of the underlying oscillations and their modulations can in principle be recorded. With advanced imaging it can also be shown which parts of the brain are most metabolically active at that moment. But what we hear as the *message* remains elusive. Although the concept of "Gestalt" is suitable for this, we still don't know what we are talking about, unless it is represented in some symbolic disembodied way. What sets RiM apart is that both the message and the way this is transmitted in neural networks is determined by a frequency spectrum. Recent research also has shown *which* networks are involved in the formation and understanding of abstract symbolic representations, like words in language, for which the term "semantics" applies (DeHaene *et al.*, 2022). The meaning of musical elements, however, is much more ambiguous than the meaning of words. The understanding of music seems more like the way animals learn to "understand" a physical or vocal expression of a conspecific in a particular context, for which the concept of "biosemiotics" applies (Kleisner, 2008). The difference with our human way of understanding is that the message also can be represented in an encoded disembodied format, with the features of "Gestalt" that we share in a likewise elusive "common spirit". Because we do not know exactly what this is about, it cannot be ruled out that something similar could finally be attributed to AI. This is actually already suggested when a form of intentionality is attributed to AI (Zhu and Harrell, 2009). Whatever this kind of "knowledge" might be, the processing of a Gestalt carrying information in humans is based on the biological roots we have in common, from a long evolutionary perspective: our *shared* intentionality. So, the final question is if AI is able not only to create an output with Gestalt-like features but also to *understand* my representations with Gestalt features (like a melody and RiM), as a condition for shared intentionality or neural resonance.

The attunement of RiM to our networks that moves us

Both BPS and rhythmic syntax and their associated networks are intertwined and any subdivision is artificial. The same applies to the various modules that we distinguish in the brain, like the basal ganglia, Broca's area, a hippocampal memory system and a thalamo-

cortical attentional network, defined in conceptual domains whose features we have defined ourselves, based on phenomena in experimental conditions or symptoms in disease (Bennet and Hacker, 2003). The structures in my anatomical illustrations are only conceptual "nodes" in a larger network, called the "connectome" (Wilkins, 2019). A subdivision has proven pragmatically useful, in communication and in our work, aimed at targeted control of symptoms or recovery from localized injuries. We are forced to classify CNS diseases and syndromes based on *similarities*, while in natural circumstances the CNS leads to *diversity while maintaining identity*. As a neurologist I have to look at the CNS as an isolated organ, although I know that it is built up in close connection to the organism as a whole and its living environment, such as can currently still be found in squids (Burkhardt *et al.*, 2023). The CNS works in the service of the organism that has to interact with a multifaceted and ever-changing living environment, both in health and disease.[3] It has to identify and bypass unpredictable obstacles or emerging needs of the organism itself, which only *than acquire the meaning of a goal* to be achieved or resolved in a particular context (Damiano and Stano, 2021; Draguhn and Sauer, 2023). As we have seen the underlying incentive in the choice of the steps to be taken is based on *prediction*, which applies to both natural and artificial neural networks (Clark, 2016; Lenc et al., 2018). But in the natural situation the challenge has yet to be *identified* as the purpose for that moment, before prediction can take place. Each time a new unique set point is created based on the needs of the organism itself and the demands of a living environment with natural, social and cultural components. A mix of components from different domains composed in a "message" that is able to *attract the attention of a CNS*. The Gestalt properties of its composite content have been created through similar situations in the evolution of both the CNS and its lifeworld, as has been shown for language and belief (Seitz and Angel, 2020). This might be the ultimate determinant for the difference between neural networks in a living CNS and artificial neural networks: not their *computations*, based on prediction, but their *receptivity and the pragmatic way its attention is attracted* in attunement to an Umwelt. A living CNS seems to be needed for attending and understanding Gestalt representing messages, as they are common in language, music and religion. Gestalt-like messages that also reshaped our mental and living worlds in evolution, by building abstract external disembodied representations of neural network activity, like a Pythagorean cosmology or God.

Life as such, is not a "goal" to which our attention can be focused. Only when life, its autopoiesis, is threatened does that threat take on the meaning of a goal. Both our genes and our nervous system become

---

[3] Research into the way this function is maintained in disorders of the CNS, rather than just its limitations, was also central to the work of the neurologist Kurt Goldstein (1878-1965).

intentional not by their "selfishness" but by setting possibilities and goals in the service of life. Life just takes its course, and to overcome unpredictable obstacles and meet with upcoming needs a CNS came into being as an intentional but particularly also *attentional* instrument. The question is not whether AI can be *intentional* (which is a matter of semantics) but whether it can be *attentional* in the way we are. To this extent, intentionality refers to phenomena that work as an incentive for our attention, rather than to a mental content, to which Brentano referred. Most particularly, attention ensures the integration of network activity that also underlies the emergence of the awareness of music (Loui and Guetta, 2019). The integration of network activity is achieved through the modulatory action of the "thalamus", thalamo-cortical connections and the integration of activity in both dorsal and ventral streams by the prefrontal cortex, if we want to express this in terms of a separate system (Womelsdorf and Fries, 2007; Damm *et al.*, 2020; Pouw *et al.*, 2021). It is generally assumed that our attention is directed in two ways: "bottom-up", by exogenous physical salient stimuli (like a source of pain or a loud noise) or "top-down", by voluntary targets that are internal to the observer, like wishes and beliefs (Awh *et al.*, 2012). As emphasized by Edward Awh and colleagues, this generally accepted dichotomy ignores an important gray area of biassing cues. The direction our attention takes is also tailored by events that shortly preceded the current one, the expectation of a subsequent signal, or cues whit some unconventional symbolic meaning (like loud noises or silent intervals in music). RiM is a good example of this, such as when an unexpected change in tempo or pitch occurs during a performance through the interaction with listeners. A phenomenon to which we are all sensitive, as evidenced by the fact that after the end of a concert we automatically gradually start clapping in harmony: a prototype of shared intentionality (Néda *et al.*, 2000).

Artificial and natural neural networks in a Chinese Room experiment

So far a thorough analysis of the composite elements of RiM and a representation of the associated modules and neural networks have been provided. The primary goal however, was not to explain the functioning of RiM, but to substantiate why we have *a particular sensitivity* to this kind of "aesthetically coded" and disembodied information and its Gestalt character. Ultimately the representation of RiM is determined by a spectrum of rhythmic activity in sound waves, transmitted by a sender and attuned to oscillatory activity in the neural networks of a receiver. The concept of *neural resonance* was introduced to represent this attunement, partly biologically determined but particularly also acquired and primed in our individual development. In that socio-cultural context we have learned to understand the Gestalt properties of music and RiM. This

understanding has been expressed in the concept of *shared intentionality*. The question now arises whether AI facilities will ultimately also be capable to this kind of shared intentionality. The philosophical concept of intentionality, that has stirred so many minds, may need an overhaul with the development of devices that can also efficiently engage in creative interaction with their users (Tigre Moura, 2023). Ambitious researchers in this field hypothesize that it even must be possible to develop a form of *embodied* AI (EAI) in the near future and do not avoid the discussion about their unconventional redefinition of concepts such as "embodiment" and the autopoietic principles underlying biological life (Damiano and Stano, 2021). But it can be questioned if it makes sense to develop AI devices the primary goal of which is to serve its autopoiesis, to which any further goal is subordinate. After all, that is how the development of our nervous system came about in an evolution of millions of years: to search and attend to upcoming needs and challenges of its living organism in an ever-changing environment, that finally also was co-created by it. Ultimately, the principle of shared intentionality underlies not only the operation of an interactive system, but also its creation and development in evolution.

The computations in artificial neural networks are similar in many ways to those in our CNS, like their algorithms based on deep learning, based on prediction and feedback and the composition of syntactic structures, similar to those in language. The processing of information in both the CNS and in artificial neural networks is governed by rhythmic oscillatory activity, some registration of which must be principally possible, both in the CNS and in AI devices and in their attunement to RiM (Llinás, 2014; Beste *et al.*, 2023). Following John Searle, in his well-known exploration of intentionality, an alternative Chinese room experiment might be imagined, to investigate whether the modulations of oscillatory activity in an (E)AI facility corresponds to that in our brain, during a variable interactive music performance, that meets its Gestalt character. But we can also consider another, probably more realistic path, that could lead to AI successfully passing this Chinese Room experiment. Our neural networks in the CNS are increasingly modeled in a world dominated by the representations of artificial systems. A living world in which AI dominates attunes the CNS of children by its own algorithms from birth. The oscillatory patterns of our connectome (tuned more by culture than by nature) will increasingly tend to resonate with those of EAI devices. And that might be the message. The increasing use of AI in the development of knowledge and instrumental or industrial applications is unstoppable. But the use of AI in early development, when the choice must be made between tuning by natural, socially interactive or artificial sources, can still be controlled by us, as long as we do not depend on AI. The danger is of AI is not only that of its unprecedented possibilities, but particularly also of its further application in an ever-younger stage of life. Music and language served as models for this, in the way they co-

evolved with our intelligence and the growing complexity of the connectome in our CNS.

Acknowledgements

Funding

Conflict of interest statement

The author declares that he has no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Allman J. M, Tetreault N. A, Hakeem A. Y, Manaye K. F, Semendeferi K, Erwin J. M, et al. The von Economo neurons in fronto-insular and anterior cingulate cortex. Annals of the New York Academy of Sciences 2011; 1225: 59–71.

Ahw E, Belopolsky A. V, Theeuwes J. Top-down versus bottom-up attentional control: a failed theoretical dichotomy. Trends in Cognitive Sciences 2012; 16(8): 437-443.

Bellinger D, Altenmüller E, Volkmann J. Perception of time in music of patients with Parkinson's disease – The processing of musical syntax compensates for rhythmic deficits. Frontiers in Neuroscience 2017; 11:68.

Bennett M. R. and Hacker P.M.S. Philosophical Foundations of Neuroscience. Blackwell Publishing Ltd, Oxford 2003.

Berwick R. C, Friederici A. D, Chomsky N, Bolhuis J. J. Evolution, brain, and the nature of language. Trends in Cognitive Sciences 2013; 7(2): 89-98.

Beste C, Münchau A, Frings C. Towards a systematization of brain oscillatory activity in actions. Communications Biology 2023; 6:137

Billig A. J, Lad M, Sedley W, Griffiths T. D. The hearing hippocampus. Progress in Neurobiology 2022; 102326.

Bonini L, Rotunno C, Arcuri E, Gallese V. Mirror neurons 30 years later: implications and applications. Trends in Cognitive sciences 2022; 26(9): 767-781.

Bornkessel-Schlesewsky I, and Schlesewsky M. Reconciling Time, Space and Function: A New Dorsal-Ventral Stream Model of Sentence Comprehension. Brain and Language 2013; 125(1): 60-76.

Brandt A, Gebrian M, Slevc L. R. The role of musical development in early language acquisition. In: Thaut MH and Hodges DA, eds. The Oxford Handbook of Music and the Brain. Oxford University Press, 2019 (Pb, 2021): 567-591.

Brette R. Is coding a relevant metaphor for the brain? Behavioral and Brain Sciences 2019; 42: e215.

Burkhardt P, Colgren F, Medhus A, Digel L, Naumann B, Soto-Angel J, et al. 2023 Syncytial nerve net in a ctenophore adds insights on the evolution of nervous systems. Science 2023; 380(6642): 293-297.

Buzsáki G. and Draguhn A. Neuronal oscillations in cortical networks. Science 2004; 304(5679): 1926-9

44

Buzsáki G, Tingley D. Space and time: The hippocampus as a sequence generator. Trends in Cognitive Sciences 2018; 22(10): 853-869.

Canolty R. and Knight R. T. The functional role of cross-frequency coupling. Trends in Cognitive Sciences 2010; 14(11): 506-515.

Cebolla A. M, Cheron G. Understanding Neural Oscillations in the Human Brain: From Movement to Consciousness and Vice Versa. Frontiers in Psychology 2019; 10:1930.

Cheung V. K. M, Harrison P. M. C, Meyer L, Pearce M. T, Haynes J-D. Uncertainty and surprise jointly predict musical pleasure and amygdala, hippocampus and auditory cortex activity. Current Biology 2019; 29:4084-4092.

Cheyne P. Encoded and Embodied Rhythm. In: Cheyne P, Hamilton A, Paddison M, eds. The Philosophy of Rhythm: aesthetics, music, poetics. Oxford University Press 2019: 255-271.

Chomsky N. Language architecture and its import for evolution. Neuroscience and Biobehavioral Reviews 2017; 81: 295-300.

Clark A. Surfing Uncertainty. Prediction, action and the embodied mind. Oxford University Press, New York, 2016.

Clayton M. Entrainment and the social origins of musical rhythms. In: Cheyne P, Hamilton A, Paddison M, eds. The Philosophy of Rhythm: aesthetics, music, poetics. Oxford: Oxford University Press, 2020: 183-198.

Couzin I. D. Synchronization: The key effective communication in animal collectives. Trends in Cognitive Sciences 2018; 22(10): 844-846.

Damiano L. and Stano P. A Wetware Embodied AI? Towards an autopoietic organizational approach grounded in synthetic biology. Frontiers in Bioengineering and Biotechnology 2021; 9:724023.

Damm L, Varoqui D, De Cock V. C, Dalla Balla S, Bardy B. Why do we move to the beat? A multi-scale approach, from physical principles to brain dynamics. Neuroscience and Biobehavioral Reviews 2020; 112:553-584.

Dehaene S, Meyniel F, Waconge C, Wang L, Pallier C. The neural representation of sequences: From transition probabilities to algebraic patterns and linguistic trees. Neuron 2015; 88: 2-19.

Dehaene S, Al Roumi F, Lakretz Y, Planton S, Sablé-Meyer M. Symbols and mental programs: a hypothesis about human singularity. Trends in Cognitive Sciences 2022; 26(9): 751-766.

Draguhn A. and Sauer J. F. Body and mind: how somatic feedback signals shape brain activity and cognition (editorial). European Journal of Physiology 2023; 475(1):1-4.

Drayton L. and Furman M. Thy mind, thy brain and time. Trends in Cognitive Sciences 2018; 22(10): 841-843.

Eichenbaum H. On the integration of space, time, and memory. Neuron 2017; 95(5): 1007-1018.

Eling P, Whitaker H. History of aphasia: A broad overview. In: Hillis AE, Fridriksson J, eds. Handbook of Clinical Neurology. Elsevier B.V. 2022; 185: 3-24.

Elimari N. and Lafargue G. Network neuroscience and the adapted mind: rethinking the role of network theories in evolutionary psychology. Frontiers in Psychology 2020; 11:545632.

Feinberg T.E, Mallatt J. Phenomenal consciousness and emergence: Eliminating the explanatory gap. Frontiers in Psychology 2020; 11: 1041.

Fitch W. T. Rhythmic cognition in humans and animals: distinguishing meter and pulse perception. Frontiers in Systems Neuroscience 2013; 7: 68.

Fitch W. T. Dance, music, meter and groove: A forgotten partnership. Frontiers in Human Neuroscience 2016; 10(64):1-7.

Friederici A. D. Evolutionary neuroanatomical expansion of Broca's region serving a human-specific function. Trends in Neurosciences 2023; 46(10):786-796.

Glass L. Synchronization and rhythmic processes in physiology. Nature 2001; 410(6825): 277-284.

Graybiel A. M. Habits, Rituals, and the Evaluative Brain. Annual Reviews of Neuroscience 2008; 31:359-387.

45

Greenfield M. D, Honing H, Kotz S. A, Ravignani A. Synchrony and rhythm interaction: from the brain to behavioural ecology. Philosophical Transactions of the Royal Society B 2021; 376:20200324.

Hauser M. D, Chomsky N, Fitch W.T. The faculty of Language: What is it, who has it, and how did it evolve? Science 2002; 298: 1569-1579.

Hauser M. D, Watumull J. The Universal Generative Faculty: The source of our expressive power in language, mathematics, morality, and music. Journal of Neurolinguistics 2017; 43(B): 78-94.

Hickok G. The dual stream model of speech and language processing. Handbook of Clinical Neurology, eds. Argye E. Hillis and Julius Fridriksson, Elsevier 2022; 185: 57-69.

Hodges D. A. Music through the lens of cultural neuroscience. In: Thaut MH, Hodges DA, eds. The Oxford Handbook of Music and the Brain. Oxford University Press, 2019: 19-41.

Honing, H. Without it no music: beat induction as a fundamental musical trait. Annals of the New York Academy of Sciences 2012; 1252:85-91.

Honing, H. Musicality as an upbeat to music: Introduction and research agenda. In: Honing H, ed. The Origins of Musicality, Cambridge, MA: MIT Press, 2018: 3-20.

Jacob P. "Intentionality". In: Zalta EN, Nodelman U, eds. The Stanford Encyclopedia of Philosophy. Spring 2023 edition.

Janata P, Tomic S. T, Haberman J. M. Sensorimotor coupling in music and the psychology of the groove. Journal of Experimental Psychology: General 2012; 141(1): 54-75.

Kasdan A. V, Burgess A. N, Pizzagalli F, Scartozzi A, Chern A, Kotz S. A, et al., Identifying a brain network for musical rhythm: A functional neuroimaging meta-analysis and systemic review. Neuroscience and Biobehavioral Reviews 2022; 136: 104588.

Kleisner K. The semantic morphology of Adolf Portmann: a starting point for biosemiotics of organic form? Biosemiotics 2008; 1: 207-219.

Kotz S, Ravignani A, Fitch W. T. The evolution of rhythm processing. Trends in Cognitive Sciences 2018; 22(10): 896-910.

Lakatos P, Gross J, Thut G. A new unifying account of the roles of neuronal entrainment. Current Biology 2019; 29: R890-R905.

Lake B. M, and Baroni M. Human-like systematic generalization through a meta-learning neural network. Nature 2023; 623: 115-121.

Large E.W, Roman I, Kim J. C, Cannon J, Pazdera J. K, Trainor L. J, et al., 2023. Dynamic models for musical rhythm perception and coordination. Frontiers in Computational Neuroscience 2023; 17:1151895.

Lenc T, Merchant H, Keller P. E, Honing H, Varlet M, Nozardan S. Mapping between sound, brain and behaviour: four-level framework for understanding rhythm processing in humans and non-human primates. Philosophical Transactions of the Royal Society B 2021; 376:2020035.

Llinás R. R. Intrinsic electrical properties of mammalian neurons and CNS function: a historical perspective. Frontiers in Cellular Neuroscience 2014; 8:320.

Loui P. and Guetta R. E. Music and attention, executive function, and creativity. In: Thaut MH, Hodges DA, eds. The Oxford Handbook of Music and the Brain. Oxford University Press, 2019: 263-284.

Maguire E. A, Gadian D. G, Johnsrude I. S, Frith C. D. Navigation-related structural change in the hippocampi of taxi drivers. Proceedings of the National Academy of Sciences 2000; 97(8): 4398-4403.

Matthews T. E, Witek M. A. G, Lund T, Vuust P, Penhune V. B. The sensation of groove engages motor and reward networks. Neuroimage 2020; 214: 116768.

Mehr S. A, Singh M, Knox D. M, Ketter M, Pickens-Jones D, Atwood S, et al., 2019. Universality and diversity in human song. Science 2019; 366;eaax0868:1-17.

Merchant H, Grahn J, Trainor L. J, Rohrmeier M, Fitch W. T. Finding the beat: A neural perspective across humans and non-human primates. In: Honing H, ed. The Origins of Musicality. Cambridge, MA: MIT Press, 2018: 171-203.

46

Morillon B, and Baillet S. Motor origin of temporal predictions in auditory attention. Proceedings of the National Academy of Sciences of the United States of America 2017; 114(42), E8913–E8921.

Müller E. J, Munn B. R, Aquino K. M, Shiner J. M, Robinson P. A. The music of the hemispheres: Cortical Eigenmodes as a physical basis for large-scale brain activity and connectivity patterns. Frontiers in Human Neuroscience 2022; 16: 1062487.

Néda Z, Ravasz E, Brechet Y, Vicsek T, Barabási A. L. The sound of many hands clapping. Nature 2000; 403: 849-850.

Patel A. D. Vocal learning as a preadaptation for the evolution of human beat perception and synchronization. Philosophical Transactions of the Royal Society B 2021; 376: 20200326.

Peretz I, Vuvan D. T, Lagrois M-E, Armony J. L. Neural overlap in processing music and speech. In: Honing H, ed. The Origins of Musicality. Cambridge, MA: MIT Press, 2018: 205-219.

Peters D. Rhythm, preceding its abstraction. In: Cheyne P, Hamilton A, Paddison M, eds. The Philosophy of Rhythm: aesthetics, music, poetics. Oxford University Press 2019: 110-124.

Petkov C. I, Marslen-Wilson W.D. Editorial overview: The evolution of language as a neurobiological system. Current Opinion in Behavioral Sciences 2018; 21: V-XII.

Petter E. A, Gershman S. J, Meck W. H. Integrating models of interval timing and reinforcement learning. Trends in Cognitive Sciences 2018; 22(10): 911-922.

Pinker S. The cognitive niche: coevolution of intelligence, sociality, and language. Proceedings of the National Academy of Sciences 2010; 107(S2):8993-8999.

Pouw W, Proksch S, Drijvers L, Gamba M, Holler J, Kello C, et al., 2021. Multilevel rhythms in multimodal communication. Philosophical Transactions of the Royal Society B 2021; 376: 20200326.

Pulvermüller F. How neurons make meaning: brain mechanisms for embodied and abstract-symbolic semantics. Trends in Cognitive Sciences 2013; 17(9): 458-470.

Ravignani A. and Madison G. The paradox of isochrony in the evolution of human rhythm. Frontiers in Psychology 2017; 8: 1820.

Robin R. and Ierna C. Christian von Ehrenfels. The Stanford Encyclopedia of Philosophy. Summer 2022 Edition, Edward N. Zalta.

Rosenblum M. and Pikovsky A. Synchronization: from pendulum clocks to chaotic lasers and chemical oscillators. Contemporary Physics 2003; 44(5): 401-416.

Ross J. M. and Balasubramaniam R. Time Perception for Musical Rhythms: Sensorimotor Perspectives on Entrainment, Simulation, and Prediction. Frontiers in Integrative Neuroscience 2022; 916220.

Roy A, Perlovsky L, Beshold T.R, Weng J, Edwards J.C.W. Editorial: Representation in the brain. Frontiers in Psychology 2018; 9: 1410.

Sauvé S. A, Bolt E. L, Nozaradan S, Zendel B. R. Aging effects on neural processing of rhythm and meter. Frontiers in Aging Neuroscience 2022; 14: 848608.

Schirmer A, Meck W. H, Penney T. B. The socio-temporal brain: connecting people in time. Trends in Cognitive Sciences 2016; 20(10): 760-772.

Schön D. and Morillon B. Music and language. In: Thaut MH, Hodges DA, eds. The Oxford Handbook of Music and the Brain. Oxford University Press, 2019: 592-622.

Schurger A, Hu PB, Pak J, Roskies AL. What is the Readiness Potential? Trends in Cognitive Sciences 2021; 25(7): 558-570.

Seitz R, Angel H-F. Belief-formation – A driving force in evolution. Brain and Cognition 2020; 140:105548.

Sheya A, Smith L. Development weaves brains, bodies and environments into cognition. Language, Cognition and Neuroscience 2019;34(10):1266-1273.

Sridharan D, Levitin D.J, Chafe C. H, Berger J, Menon V. Neural dynamics of event segmentation in music: Converging evidence for dissociable ventral and dorsal networks. Neuron 2007; 55:521-532.

Stolk A, Blokpoel M, Van Rooij I, Toni I. On the generation of shared symbols. In: Willems R, ed. Cognitive Neuroscience of Natural Language Use. Cambridge University Press, 2015: 201-227.

47

Tigre Moura F. Artificial Intelligence, Creativity, and Intentionality: The need for a paradigm shift. Journal of Creative Behavior 2023; 57(3): 336-338.

Tomasello M. Origins of Human Communication. Cambridge, MA: MIT Press, 2008.

Trainor L. J. The origin of music: auditory scene analysis, evolution, and culture in musical creation. In: Honing H, ed. The Origins of Musicality. Cambridge, MA: MIT Press, 2018: 81-112.

Trainor L. J. and Marsh-Rollo S. Rhythm, meter and timing: The heartbeat of musical development. In: Thaut MH, Hodges DA eds. The Oxford Handbook of Music and the Brain, Oxford University Press, 2019; (Pb, 2021): 592-622.

Vanden Bosch der Nederlanden C. M, Taylor J. E. T, Grahn J. A. Neural basis of rhythm perception. In: Thaut MH, Hodges DA, eds. The Oxford Handbook of music and the Brain. Oxford University Press, 2019 (Pb ed. 2021);165-186.

Von Schnehen A, Hobeika L, Huvent-Grelle D, Samson S. Sensorimotor synchronization in healthy aging and neurocognitive disorders. Frontiers in Psychology 2022; 13: 838511.

Wark B, Lundstrom B. N, Fairhall A. Sensory adaptation. Current Opinion Neurobiology 2007; 17: 423-429.

Weiller C, Reisert M, Glauche V, MussonM, Rijntjes M. The dual-loop model for combining external and internal worlds in our brain. NeuroImage 2022; 119583.

Wilkins R. W. Network Neuroscience: an introduction to graph theory network-based techniques for music and brain imaging research. In: Thaut MH, Hodges DA, eds. The Oxford Handbook of Music and the Brain. Oxford University Press, 2019 (Pb ed. 2021): 122-144.

Wilson M. and Cook P. F. Rhythmic entrainment: Why humans want to, fireflies can't help it, pet birds try, and sea lions have to be bribed. Psychonomic Bulletin and Review 2016; 23(6): 1647-1659.

Womelsdorf T. and Fries P. The role of neuronal synchronization in selective attention. Current Opinion in Neurobiology 2007; 17:1-7.

Zhu J, and Fox Harrell D. System Intentionality and the Artificial Intelligence Hermeneutic Network: the role of intentional vocabulary. UC Irvine 2009.

Zuidema W, Hupkes D, Wiggins G. A, Scharff C, Rohrmeier M. Formal models of structure building in music, language, and animal song. In: Honing H, ed. The Origins of Musicality, ed. Cambridge, MA: MIT Press, 2018: 253-286.

48