Radical Embodied Computation: Reproduction of Similarity by Analogy as an Order Generating Mechanism in Complex Systems

Fred Hasselman* [1]

[1] School of Pedagogical and Educational Science, Radboud University Nijmegen

*** Correspondence:**
Fred Hasselman
f.hasselman@pwo.ru.nl

Keywords: Natural Computation, Order Generating Mechanism, Multi-Fractal Spectrum, Reproduction of Similarity by Analogy, Thermodynamic Improbability,

Abstract

About two decades ago, scholars from the scientific disciplines that study human behavior and cognition, suggested an era of post-cognitivism was imminent, in which the computer metaphor, computationalism and representationalism would be discarded as viable theoretical frameworks for explaining phenomena of the body and the mind. In the present paper I argue that explanations of complex adaptive behavior require a theory of meaning mechanics that explains how complex adaptive systems can use semantic information to coordinate their behavior. This calls for a unification of sorts between the insights obtained in ecological psychology and embodied embedded cognition with principles of natural computation (cf. Decastro, 2007) in the context of explaining the behavior and properties of complex adaptive systems and networks (see e.g., Freeman et al., 2001; Chialvo, 2010; Flack, 2017a; Scheffer et al., 2018). I will refer to this framework as *Radical Embodied Computation* (REC_{++}) and discuss some of the philosophical and theoretical issues that have to be resolved. I conclude by suggesting a mechanism for the emergence of meaning that is based the conception of self-affine scaling as the reproduction of similarity by analogy.

Word count body of text & footnotes: 11,500

1 The Missing Physics of Meaning Mechanics

Classical, algorithmic and quantum information theory explicitly exclude meaningful, or, semantic information as part of their explanatory domain (cf. Desurvire, 2009, p. 38). As Shannon lucidly explained:

> "The fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another point. Frequently the messages have *meaning*; that is, they refer to or are correlated according to some system with certain physical or conceptual entities. These semantic aspects of communication are irrelevant to the engineering problem. The significant aspect is that the actual message is one *selected from a set* of possible messages. The system must be designed to operate for each possible selection, not just the one which will actually be chosen since this is unknown at the time of design." (Shannon, 1948, p. 379, emphasis in original)

The semantic aspects of a message are irrelevant for reproducing it, if they were relevant for message reproduction, a universal theory of information and communication would not be possible. In order to successfully communicate a message through a channel if semantics were involved, one would have to know particular facts about these correlations with "*some system with certain physical or conceptual entities*". Obviously, this is not necessary from the perspective of the engineering problem of reproducing a message, however, it is also apparent that complex living systems appear to make use of the meaning (semantics) associated with the configuration (syntax, or, content) of an information source to coordinate their behavior (cf. Turvey and Carello, 2012; Bruineberg et al., 2019). The semantics can be highly contextual and specify a meaningful message only for a very particular system and state of its internal or external environment. Other types of semantic information can be more general and specify a meaningful signal for a larger group of systems that share an identity, for example, because their dynamics are coupled in some way, or, because they evolved in the same ecological niche (see e.g., Grignolio et al., 2014; Flack, 2017b). In most disciplines of science, this distinction between the different aspects of information representation by physical systems is well known and plays an important role in theorizing about the behavior of thermodynamic, relativistic, and quantum systems (Levitin, 1992a; b; Wolpert, 1992), and reality as described by string theories and quantum gravity due to the Holographic Principle ('t Hooft, 1994; Susskind, 1995).

As was recently pointed out by Kolchinsky and Wolpert (2018), there aren't many theories available that formally describe how complex adaptive systems exploit the semantics represented by a specific configuration of an information source (an encoded message), to coordinate their behavior. In addition, the mechanisms and boundary conditions that allow for the emergence of meaning in complex systems are often only implicitly assumed and are rarely accompanied by a plausible account of the physical realization of the proposed mechanisms (see e.g., Boyle, 1992; Chemero, 2009). Another problem has to do with the nature of the coordination of complex adaptive behavior which appears to be based on 'personalized meaningful information', that is, behavior is co-determined by some form of memory for specific interactions with the internal and external environment. In order to predict their future behavior, one cannot rely solely on universal principles and laws that specify the temporal evolution of the system departing from some estimate of initial conditions, instead, one has to consider a unique history of experienced events that somehow specified the internal structure of a system to such an extent that its behavior will diverge from an identical system save for its interaction history. This type of behavior is fundamentally different from non-living matter. Imagine what would happen if the internal structure of the particles in a thermodynamic system would were such that the after-effects of interactions could specify their internal configuration in such a way that with each collision with another particle the probability that it would avoid a future collision would increase, a universal theory of thermodynamics (or statistical mechanics) would not be possible.

Both the lack of a comprehensive theory about meaningful information as well as the crucial role played by 'meaning' in terms of highly contextualized information needed to explaining adaptive behavior, may be the cause of the frequent conflation of the information quantity with meaningful information by the scientific disciplines that study human behavior. It is the purpose of the present paper to explore whether advancements in the study of complex adaptive systems can provide a theoretical framework that would allow a more formal treatment of the role played by meaningful information in explaining the coordination of adaptive behavior. Several recent proposals for such a framework have been made (Küppers, 2013; Flack, 2017a; Kolchinsky and Wolpert, 2018; Kolchinsky and Corominas-Murtra, 2019) Especially the behavioral-, cognitive- and neurosciences could benefit from a theory of meaning mechanics, or, at least the development of a broader theoretical framework whose concepts, principles and laws will not contradict the (formal) definitions of the highly corroborated theories of information and computation developed in other scientific disciplines. An example of an influential theoretical framework in which concepts from information theory were adopted and subsequently conflated with meaning and semantic content, is the so-called computational theory of mind (CTM), also referred to as the computer metaphor of human behavior. Before I can suggest the outlines for a framework of behavior coordination based on semantic information, I will first characterize some of the issues associated with the use of information-theoretic concepts as meaning-bearing constructs invoked to explain complex adaptive behavior.

1.1 Cognitive Science's Unauthorized Information Theory[1]

About two decades ago, many scholars expected an era of post-cognitivism[2] was imminent, in which CTM would be discarded as a viable framework for explaining phenomena of the body and the mind (see e.g., van Gelder, 1995; Freeman, 1997; Clark, 1999; Turvey and Shaw, 1999; Potter, 2000; Weng et al., 2001; Anderson, 2003). The objections to CTM, which ascribes to the body and mind an architecture that resembles an input-output machine with hardware and software, capable of processing information much like an electronic computer, are best captured by Clark's 007-principle:

> "In general, evolved creatures will neither store, nor process information in costly ways when they can use the structure of the environment and their operations upon it as a convenient stand-in for the information-processing operations concerned. That is, *know only as much as you need to know to get the job done*." (Clark, 1997, p. 46).

The argument is expressed in terms of the improbability of natural evolution giving rise to complex structures and functions, such as internal representations of the state of the external world, and processes operating on those internal representations to generate complex adaptive behavior. Here, 'costly' refers to the fact that although computational architectures could in principle emerge from evolutionary processes, often, they are not necessary (and/or sufficient) to explain a whole range of intelligent, adaptive behaviors: A skilled agent will be able to exploit the regularities of the environmental niche (often referred to as ecological information, cf. Bruineberg et al., 2019), to coordinate its behavior according to ecological principles and laws (cf. Petrusz and Turvey, 2010). Indeed, at their core, many of the objections against CTM concern the implausible biological and physical order that is imagined into existence by computational theories of cognition, without providing an account of where this order came from (Gibson, 1966; Carello et al., 1984; Searle, 1990; Varela et al., 1992).

This neglect for providing a plausible physical implementation of posited computational architectures for the mind often results in mechanisms that contradict theories or formal definitions of concepts from other scientific disciplines. For example, computational processes explaining cognitive phenomena often perform some kind of dual function that involves the manipulation (coding/decoding) as well as the representation (storage/retrieval) of information content. This mechanism of concurrent representing and processing of information based on semantics was explicitly declared as a necessary condition for explaining intelligent behavior in the Classical computational architecture for cognitive processes: *"[…] according to Classical theories, the syntax of mental representations mediates between their semantic properties and their causal role in mental processes.*" (Fodor and Pylyshyn, 1988, p. 5). This notion of a combinatorial syntax (grammar) and semantics (meaning) that govern the structure of the constituent parts of the representational system persists in many other theories of cognitive phenomena. For example, a popular model that can be fitted to distributions of response latencies recorded in binary decision tasks claims that it describes

[1] The term `unauthorized' is due to an anecdote related by Walter Freeman at the inaugural meeting of the Society for Complex Systems in Cognitive Science in Amsterdam, 2009. Freeman recalled a lab-visit by Claude Shannon. According to Freeman Shannon expressed concerns about the using the constructs from his mathematical theory of communication to describe information representation and communication in biological systems.

[2] I will use the term post-cognitivism in a very general sense, referring to any theoretical framework that rejects (parts of) the ontological and epistemological claims by the classical Computational Theory of Mind. Specifically, the metaphor of the input-processing-output machine and the internal (mental) representation.

an accumulation of information, which is guided by the information content of a stimulus: "*The diffusion model provides an account of both the stimulus information and the processing that makes use of that information.*" (Ratcliff and Rouder, 1998). A similar notion, prevalent in a well-known theory of human memory, concerns a combination of information storage and processing: "*The term "working memory" evolved from the earlier concept of short-term memory (STM) […] I will use STM to refer to the simple temporary storage of information, in contrast to WM, which implies a combination of storage and manipulation*" (Baddeley, 2012). In both examples it is unclear how and where information can accumulate or be stored as a state configuration of a physical medium. The diffusion model uses a well-known mathematical description of Brownian motion, but why would mental information accumulate according to this specific process and not anomalous diffusion, or diffusion limited aggregation, which are more common in nature (see e.g., Burov et al., 2011; Magdziarz and Weron, 2011; Metzler et al., 2014).

1.2 The Loan Crisis: Time to foreclose on virtual explanatory domains?

Theories of embodied cognition based on principles like the 007-principle seek to 'offload' as many storage and computation of meaningful information as possible onto the (dynamically) invariant structure of the body and the environment, without taking out so-called 'loans on intelligence' (Dennett, 1981, p. 12). This refers to incorporating knowledge about particular facts of the environment and body (i.e. semantic information), as well as the ability to exploit that knowledge (i.e. coordinate behavior based on semantic information), into the physical structure of a skilled agent. This is a rather conspicuous example of recruiting the "effect = structure fallacy" (see e.g., Bosman et al., 2013) to dispatch the explanatory burden of accounting for observable phenomena to some underspecified internal structure endowed with powers of causation (see e.g., Michaels and Carello, 1981; Petrusz, 2008; Chemero, 2009; Noble, 2010). For example, by referring to the activation of innate motor programs to explain the emergence of crawling and walking in young children, or suggesting a predisposition for perceiving particular features of the social environment in order to explain infants' preference for looking at human faces, one merely re-describes the phenomenon in terms of a causally potent internal structure, "walking is caused by the internal structure that causes walking", "infants prefer to look at human faces because there is an internal structure that causes them to prefer to look at human faces".

Taking out a loan on intelligence is a valid method in theory construction if it concerns a reference to an auxiliary theory or principle. Van Geert and Fischer (2009) describe how importing mathematical tools that describe the behavior of a system as a function of time at a characteristic level of description (dynamical models), involves taking implicit loans on scientific knowledge about physical composition:

> "Any of these levels can be taken as the starting point of a dynamic model, that is, a model that describes the time evolution of some variable of interest that is characteristic of the level of aggregation chosen. These levels of aggregation, or levels of description, correspond with different timescales, and different types of dynamics. [. . .] But whatever one's target level of description, there is always an implicit loan taken on other levels. In many cases, the loan is guaranteed, in the sense that the link with adjacent levels is well understood." (Van Geert and Fischer, 2009)

This scenario of loan taking is acceptable: All is well as long as the implicit loans are guaranteed by theories who have *"[. . .] lots of money in the bank*" (Meehl, 1990, p.15), which is the case for many physical and mathematical theories that describe system dynamics at the spatial and temporal scales of aggregation at which cognitive phenomena are expected to unfold. If a credible auxiliary theory goes bankrupt, it will be considered 'normal science' if theories that have benefitted from the extra credit, will have to foreclose on any property or behavior assigned to an entity, or any domain in reality claimed for proprietary explanation. The theory will be affected in terms of perceived credibility whenever some of the foundations it is built on, turn out to be false. This type of loan-taking itself may be risky, but as a scientific practice it is unproblematic and likely unavoidable.

A problem arises when these ontic links between levels are cut, but there are no consequences for the perceived scientific credibility of loan-takers, which can occur if the necessity for plausible physical implementation is not recognized. An example is the perspective of 'non-reductionist physicalism', proposed by the progenitors of CTM. It concerns a complete dismissal of the need for providing any account of physical realization of posited virtual computational architectures other than stating they are assumed to be emulated by the brain and nervous system: "*Computer theorists [. . .] often speak of identities of virtual architecture. Roughly, you establish the virtual architecture of a machine by specifying which sets of instructions can constitute its programs [. . .] the present question is why anything except virtual architecture should be of any interest to the psychologist;*" (Fodor, 1983, p.33). This has led to a new kind Mind-Body dualism, one for the digital age which I will refer to as Virtual-Physical dualism. In addition to the problem of defaulting on loans of intelligence, one can question what kind of scientific knowledge is produced by focusing exclusively on virtual architectures:

> "This interpretive view of computation is responsible for the widespread use of functional models to understand cognition and computer programs to simulate mental behavior. Such models, however, fail to tell us anything about physical characteristics of the brain's information processing, only that we can interpret the brain as computing some class of functions. Yet it is from those physical characteristics that minds emerge, underscoring the importance of understanding the basis of this phenomenon. Certainly we won't understand it by identifying the class of functions the brain can be interpreted as computing, nor is that even a remotely realistic goal given the complexity of the brain and our behaviors." (Boyle, 1994, p. 452)

At best, research programs embracing V-P dualism can discover classes of functions the internal structure of a complex adaptive system can be interpreted as computing. Boyle's argument is more than two decades old, but the history of Robotics and AI seems to have proven it is was a valid objection. There may have been benefits for the development of new technologies based on these virtual architectures, but it has not produced the advances in knowledge that would help a scientific understanding of how meaning and minds emerge from the physical characteristics of the structure and interactions between the internal and external environments of complex adaptive systems. One of the remedies suggested by post-cognitivist perspectives such as described by Van Geert and Fisher (2009) is to use the well-studied process of self-organization observed in open physical systems that exist far-from thermodynamic equilibrium, which does not require one to take any loans on intelligence to explain the emergence of new structures and behaviors (see e.g., Kelso, 1995, p. 34).

1.3 From Interpretative Computation to Radical Embodied Computation

Although there have been many advances in terms of the development of roadmaps for post-cognitivist research programs and theory development, such as *Radical Embodied Cognition* (Chemero, 2009), *Physical Intelligence* (Turvey and Carello, 2012) and *Embodied, Embedded, Extended, Enactive (4E) Cognition* (cf. Newen et al., 2018), a discipline-wide, revolution has not happened yet (te Molder, 2016). This has led some authors to suggest that perhaps the anti-computationalist, anti-representationalist foundations of the post-cognitivist movement have to be reconsidered (cf. Villalobos and Dewhurst, 2017; Golonka and Wilson, 2019; Heras-Escribano, 2019). Admittedly, a large part of the present paper provides another post-cognitivist roadmap for theory development, which I will refer to as *Radical Embodied Computation* (REC++)[3], however, I will not propose to discard the 007-principle and resort to V-P dualism, but attempt to get the best of both worlds, by embracing principles of natural principles of natural computation (cf. Decastro, 2007) in the context of explaining the behavior and properties of complex adaptive systems and networks (see e.g., Freeman et al., 2001; Chialvo, 2010; Flack, 2017a; Scheffer et al., 2018).

The notion of embodied computation is not new and can be summarized as computation by ‘respecting the medium’ and ‘exploiting the physics’ of the system:

> *“*[…] with conventional computing technology we often “torture” the physical substrate so that it implements desired computations (e.g., using continuous electronic processes to implement binary logic), whereas embodied computation “respects the medium,” conforming to physical characteristics rather than working against them. The goal in embodied computation is *to exploit the physics, not to circumvent it* (which is costly).*”* (Maclennan, 2011, p. 230)

Here, ‘costly’ refers to the efficiency of the physical implementation of a (proposed) computational architecture. The best current example of ‘exploiting the physics’ is of course the race to achieve quantum supremacy by building a machine that can exploit that nature of reality at the level of the quantum world (cf. Lanting et al., 2014). Natural computation entails much more than engineering efficient computational machines, in its most radical form it concerns a scientific realism about the concept of the information quantity. Informational Realism (cf. Floridi, 2003) and Digital Physics, can be said to *“[…] regard the physical world as made of information, with energy and matter as incidentals*” (Bekenstein, 2003, p. 59), the universe is essentially made up of physical information and lawful behavior should be considered the result of natural computations (see e.g.,Wheeler, 1989; Fredkin, 1990; Hopfield, 1994; Wolfram, 2002; Pfeifer et al., 2007; Polani et al., 2007; MacLennan, 2012).

The remainder of the paper is structured as follows, first, I provide a very brief introduction to information and entropy from the perspective of physical information and analyze what it is that may be required of such concepts before they can serve as physically plausible and ecologically useful coordinative structures for generating adaptive behavior. Second, I will propose a number of definitions of concepts that can be used to describe the behavior of complex adaptive systems in terms of natural computation and self-organization. The goal is to offer verbal definitions of theoretical concepts that should at least not contradict highly corroborated theories of other scientific disciplines. Finally, a REC++ mechanism is proposed that is based the idea that the inherent multi-scale structure of the world as described by self-affine scaling, can be considered a code that that signals similarity across scales by means of (approximate) reproduction of patterns by an affine scaling factor.

2 Physical Information and Entropy

In the most general terms, information can be defined as a quantity that resolves uncertainty about the state of an information source, a physical entity that can represent an amount of information. The

[3] The addition of ++ in REC++ not only allows easy discrimination from REC, the acronym commonly used for *Radical Embodied Cognition*, but also refers to the popular computer programming language C++ which offers a function ‘++’, the increment operator. The operation of incremental counting will be considered an important property of the computational mechanism discussed in the last section of this paper.

information-theoretic quantity known as *self-information*, or *information content* (I) quantifies the reduction of uncertainty due to the observation of a particular state of the information source. For example, an information source that can be in 2 states, a fair coin, can represent 2 bits of information (2 things to be uncertain about). Observing 1 of the 2 possible states ('heads') resolves uncertainty about the state of the system by an amount of 1 bit of information. Another important information-theoretic quantity is *entropy* (H), which can be interpreted as the expected value of I over repeated observations of states of the information source. For a fair coin, the 2 possible states are equiprobable, so the expected amount of information represented by the observation either 'heads' or 'tails' will be 1. Figure 1 displays the relation between the information and entropy represented by each state for different configurations of the coin: Biased towards 'Tails' ($Pr(X = \text{'}Heads\text{'}) < .5$); fair ($Pr(X = \text{'}Heads\text{'}) = .5$); biased towards 'Heads' ($Pr(X = \text{'}Heads\text{'}) > .5$). The entropy is at its maximum when the system is configured to be fair, in which case the observation of a particular state does not provide any information about the state that will be observed next, our best guess, is quite literally, to guess.

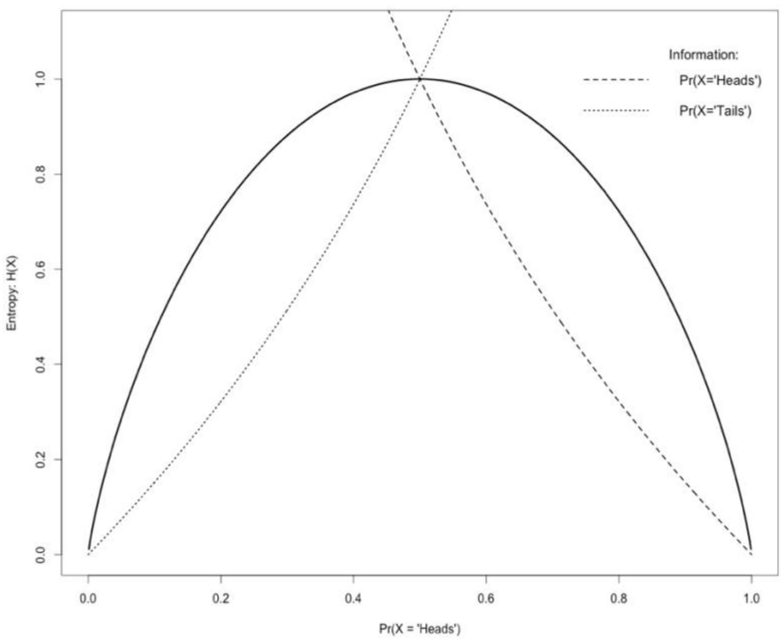

Figure 1. Entropy and information in a coin system. The solid line shows the relation between the probability of observing the state $\boldsymbol{X = 'Tails'}$ and the entropy of this 2-state system $\boldsymbol{H(X)}$. The dotted lines show the Information represented by the observation of $\boldsymbol{X = 'Heads'}$ or $\boldsymbol{X = 'Tails'}$. Maximal entropy $\boldsymbol{H = 1\ bit/symbol}$ is observed when the states of the system are equiprobable at $\boldsymbol{Pr(X = 'Heads') = 0.5}$. The observation of each state represents the same amount of information ($\boldsymbol{I = 1}$).

The maximum entropy state represents maximum unpredictability, in the case of a physical system one would refer to the maximum entropy state as a state of maximum disorder, which will be discussed in the next section. For now, it is important to note that information is dimensionless in the sense that it does not represent a characteristic scale of observation, but has to be interpreted relative to the scale at which the microscopic states of the system are observed. If a computer scientist and a particle physicist have to estimate the amount of information that is represented by a 1GB memory chip, they will give two different, but correct answers (2^{33} and $\sim 2^{76}$ bits, respectively). This is due to the scale at which the degrees of freedom of the information source are identified. The computer scientist cares about the on and off states of the transistors on the chip, whereas the particle physicist is more likely interested in the states of billions of individual particles (example taken from Bekenstein, 2003, p. 59). A user of a 1GB flash drive commonly wants to know whether there is enough free space to copy one or more files onto it. The amount of free space, expressed as a quantity of information, for example, 100MB, refers to the degrees of freedom at the level of the on and off states of the transistors on the chip that are still available to become "*correlated according to some system with certain physical or conceptual entities*". I will refer to such degrees of freedom as *non-information bearing d.o.f.*, or, *non-informative structure*, as opposed to *information-bearing d.o.f.*, or, *informative structure*. The latter would concern those transistors whose state is temporarily fixed, because they are part of a specific configuration of transistor states that is associated with one or more systems of physical or conceptual entities (i.e. the state encodes for whatever digital content was stored on the drive). The non-informative structure concerns transistors whose state is not fixed, and would be available to become associated to some other system of entities. Another way to describe this situation, is that there has been a (relative) reduction of the total degrees of freedom available to the system, from 8,589,934,592 (1GB) to 838,860,800 (100MB) on/off states. The informative structure represents systematicity, certainty, order: "*order is essentially the arrival of redundancy in a system, a reduction of possibilities*'' (von Foerster, 2003). The non-informative structure represents possibility, uncertainty, disorder.

These quantities of information can be used to distinguish one system state relative to another (e.g. less or more redundancy/possibility, order/disorder, unavailable/available storage space), and this can lead to the emergence of identity (cf. Kak, 1996). However, this is not the same as the emergence of meaning, that is, it the quantity does not account for the emergence of informative structure, which would require knowledge about the system with which the configuration of degrees of freedom became associated. The amount of information represented by an information source does not specify what it codes for, it is meaningless (cf. Küppers, 2013). To 'make use' of the meaning encoded in an information source one would need to be able learn the systemic regularities it codes for by observing how the redundancies came about, or, figure out how to translate the configuration into another, known code (an equivalent to the Rosetta stone). Compared to the concept of internal or mental representation, the physical representation of information is much more an indication of a capacity for registering an amount of information (Lloyd, 2006). It refers to a potential for establishing a new order in the configuration of the system by recruiting available degrees of freedom, whereas the mental representations the post-cognitivists seek to dispense with, refer to previously realized order that was somehow trapped, stored, or imprinted into the structure of the system.

2.1 Properties of coordinative structures for adaptive behavior

The persistent problem for any theory seeking to explain intelligent behavior by a complex adaptive systems that coordinates its behavior based on previously experienced events, is that one needs to posit the existence of highly contextualized meaningful information that concerns particular facts about the unique history of interactions the system had with its internal and external environment, the interaction biography of the system (see e.g. the case of anomalous diffusion, Metzler et al., 2014). For example, in order to understand why one of two virtually identical immune systems (monozygotic twins developing in the same environment), responds with an allergic reaction to an otherwise harmless substance, while the other doesn't, one has to trace the interaction biography to a unique event experienced by one system, but not the other (Grignolio et al., 2014). Even the early post-cognitivist theories acknowledge that meaningful information, whether this refers to an experienced event, or, embodied, embedded knowledge of an evolved agent in an ecological niche is crucial for understanding complex adaptive behavior:

> "Neither classical information theory nor the currently popular quasilinguistic view of information is acceptable […] Ecological realism imposes severe demands on the concept: Information must be: *unique* and *specific* to the facts about which it informs […], *meaningful* to the coordination and *control* requirements of the activity […], and *continuously* scaled to the dimensions of the system over which the activity is defined" (Turvey and Carello, 1981, p. 316, emphasis not in original)

These same 'demands' of information were reiterated recently in the context of the impressive research program 'intelligence from first principles', referred to as *Gibson information* (Physical Intelligence, Turvey and Carello, 2012). The need to introduce some notion of meaningful information to explain the full spectrum of complex adaptive behaviors, is exemplified by recent suggestions to consider the existence of different kinds of information (Golonka and Wilson, 2019). This often occurs when the goal is to explain 'higher cognition', which Bruineberg et al. (2019) define as coordination of behavior based on "*aspects of the sociomaterial environment that are not sensorily present*". In their framework *lawful ecological information* refers to the regularities and invariants of the ecological niche an agent can perceptually couple with to coordinate its behavior, whereas *general ecological information* refers to regularities of the ecological niche of the kind that if X occurs, it is likely that Y will occur as well (i.e., lawful versus statistical regularities). This is in fact describing an interaction biography and the distinction between the types of information refers to different types of regularities of the environment that may be used by the agent.

From the perspective of ecological realism, the properties requested of the posited information entities are reasonable and indeed necessary given the explanatory domain, however, if one examines the explanatory work they do, one must eventually conclude this is equivalent to the explanatory work done by the 'quasilinguistic information' posited to exist by the Classical account due to 'nomic necessity' (Fodor and Pylyshyn, 1988). That is, by stating an agent-environment system will make use of regularities of the environment, this means the coordination of the activity involves a process in which informative structures are involved, the internal state configuration of the agent has become "*correlated according to some system with certain physical or conceptual entities*" in the external environment. To understand how such a correlation emerges, is to understand the emergence of meaning. Stating that ecological information or Gibson information is different from Shannon's information, is a missed opportunity to make a formal connection to some of the most successful scientific theories produced by science. Notwithstanding this critique, ecological psychology and other post-cognitivist perspectives do provide many answers to the 'how' question, often based on first principles and general laws, but avoiding the use of concepts like information and computation. It is beyond the scope of this paper to discuss all such accounts, but I suggest some of the proposals made in the present paper could be incorporated into these accounts, without affecting their solid foundations: Direct Perception (Michaels and Carello, 1981), Direct Learning (Jacobs and Michaels, 2007), tau-guidance (Lee, 1998), Ecological Mechanics (Shaw and Kinsellashaw), the dynamical theory of Intentional Systems (Kugler et al., 1990; Shaw et al., 1994), the theory of Event Perception (Shaw et al., 1996), the deduction of perception-action cycles from thermodynamic principles (Swenson and Turvey, 1991) and Strong Anticipation (Stephen et al., 2008; Stepp and Turvey, 2010; Stepp et al., 2011; Marmelat and Delignieres, 2012).

To summarize, I argue that attempts to explain complex adaptive behavior often conflate information-theoretic concepts with meaningful information. At the core of the problem lies the distinction between informative and non-informative structure, of which the latter does not appear to play any role in explanatory claims of complex adaptive behavior. The former, however, is almost always confused for the meaning an information structure encodes for. A specific state configuration can only be decoded for meaning if it is known according to which regularities they became correlated, that is, the regularities of some other system. I assume most contemporary cognitive scientists would not object in principle against the definitions by Bruineberg et al. (2019) in which higher cognition refers to the coordination of behavior mediated by causes that are not part of the manifold of immediate sensory experiences. The main difference between a cognitivist account of higher cognition and a post-cognitivist account is that the former would likely posit abstract, internal mental representation as efficient causes of such phenomena (taking loans on intelligence), the latter would introduce an agent exploiting the regularities of the environment it evolved in. These different perspectives both introduce a 'stand-in' for the unique, specific, meaningful information that is required to explain behavioral phenomena that qualify as higher cognition, but neither develops a plausible physical implementation of how this meaning emerges in the agent-environment system.

3 Physical Information and Self-Affine Structure

The very existence of ordered structure in the universe seems to defy, at least locally, the 2nd law of thermodynamics, which states everything in a closed system will end up in a state of thermal equilibrium, a state of maximum entropy, minimal redundancy, maximum homogeneity, minimal predictability (Atkins, 2010). Therefore, from the perspective of lawful attraction to thermodynamic equilibrium, all living physical systems must be considered highly improbable structures of ordered mass and energy (Schrödinger, 1944). The *entropic arrow of time* and the *biological arrow of time* appear to run in opposite directions (Walker, 1972), order generation must be considered an opposing principle to entropy generation, denoted by Schrödinger (1944) as *negentropy* (also see Mahulikar and Herwig, 2009; Friston, 2012). When thermodynamically improbable order emerges in any physical system, this implies that free energy was expended to do work (create order), dissipating heat energy (disorder) back into the environment, hence the term 'dissipative systems' to denote living matter (e.g., Prigogine and Nicolis, 1977; Prigogine and Stengers, 1984; Lintern and Kugler, 1991; Schweitzer et al., 2001).

There is a formal equivalence between information-theoretic quantities and quantities of physical theories that is most apparent when they are described in terms of the degrees of freedom a system has available to generate its behavior, relative to all possible configurations of its state space. As discussed in the previous section, this information can be of at least two different kinds: the degrees of freedom that are fixed represent redundancies, the degrees of freedom that are available represent possibilities (also referred to as surprisal, or, novelty). In order to constrain or free up degrees of freedom in state space a system has to expend free energy and as a consequence dissipate heat energy to its environment, which means entropy is generated. It can be said the system is processing information, computing the new state configuration, or that it is self-organizing into a new stable state. I will refer to order generating processes (OGPs) whenever a system appears to work against the entropic arrow of time. Examples of OGPs are biological reproduction and development, self-organized criticality and computation. Examples of order destroying processes (ODPs) are diffusion, burning, and erasing information. With respect to the latter example, the physical connection to information is profound if one considers Landauer's principle: Erasing information increases entropy by converting the average amount of information that is erased into heat energy (see Landauer, 1991; Berut et al., 2012).

The reason to point to these general principles and laws of biological systems is to stress that scientific explanations of complex adaptive behavior of biological systems cannot ignore, or contradict these principles and laws. In the REC++ framework, any process that generates ordered structure must do so in a way that 'exploits the physics' and 'respects the medium'. An example of a theoretical framework that meets many of the goals of REC++, is the free energy principle (Friston, 2012; Badcock et al., 2019). It assumes that a specific set of biological systems whose boundaries can be considered a so-called Markov blanket (i.e., ergodic systems, or, random dynamical systems, cf. Pearl, 1988) display self-organized coordinated behavior, often called active inference, by minimizing available free energy. Such systems maintain their internal complexity by minimizing the difference between an internal state (a model of the world) and their perception of the current state of the world.

Theories based on free energy minimization meet the requirement of being based on physical principles and laws, however, the actual physical implementation of the mechanisms proposed to explain cognition and adaptive behavior remains unspecified, they indeed exploit the physics', but not (yet) 'respect the medium'. Notwithstanding this objection, it is clear that free energy minimization and related principles, such as the entropy maximization (Swenson and Turvey, 1991; Barab et al., 1999; Kondepudi et al., 2015; Davis et al., 2016) and least-action principles (Kugler and Shaw, 1990; Shaw and Turvey, 1999) can provide powerful explanations for the emergence of (self-organized) order in the universe. Also, the Markov blanket is an interesting explanatory tool that is related to the adiabatic boundary in thermodynamics (see e.g. Hopfield, 1994) and the phenomenon of coarse graining (see e.g. Flack, 2017a), both of which will be discussed in the next sections.

The free energy principle by itself does not provide an explanation for why highly ordered, thermodynamically improbable structure can emerge under the 2nd Law of thermodynamics, how and why such structure is reproduced and appears to be able to (locally) increase in complexity over time. One interesting conjecture by England (2013), sheds light on the question why thermodynamically improbable configurations of energy and matter are (self-)reproducing and able to increase their complexity. England (2013) provides a formal argument in which it is shown that the reproduction of thermodynamically improbable structure is in fact a highly efficient way to dissipate free energy, that is, to increase the global entropy of a closed system. If the conjecture holds, nature is full of replicating systems *because* this optimizes global entropy production and the evolution towards the thermodynamic heat-death of the universe as predicted by the 2nd law of thermodynamics. It explains why OGPs, operating within a closed system only locally, or, temporarily, appear to defy the 2nd law, in the long run it is an optimal strategy for transforming order into entropy.

Figure 2 is a schematic representation of the proposed relations between the emergence of order and disorder and thermodynamically probable and improbable structure that can emerge in a closed system. The explanatory goal of REC++ is to provide an account of complex adaptive behavior that not only respects the specific domain in reality carved out in Figure 2, but also respects the fact that such behavior is likely driven by highly contextual, meaningful information, rather than completely determined by universal physical laws. Theory generation about complex adaptive behavior, such as phenomena of the mind, should occur within the constraints these explanatory goals put on ontology and epistemology. In what follows I present a theoretical framework in which the reproduction of similarities across spatial and temporal scales plays a central role in explaining the emergence of meaning in complex adaptive systems.

3.1 Natural Codes: Reproducing Similarity by Analogy

Shannon's information theory is a communication theory and the formulation of the source coding theorem and the channel capacity theorem are its important results (cf. Desurvire, 2009; Shannon, 1948). They concern a description of the properties of messages and the way they can be transmitted (or reproduced) from one or more sources. In Shannon's terminology a message can be communicated through a channel by a transmitter that encodes it, or "*operates on the message in some way to produce a signal suitable for transmission over the channel*" (Shannon, 1948). A receiver performs the inverse operation (decoding) at a destination: "*a person (or thing) for whom the message is intended*" (Shannon, 1948). It is clear that encoding for meaning by information sources can only occur if they are part of a larger communication system of several interacting components. Another way to describe this perspective on communication is that its purpose is to detect the identity of a message across a lag of time. Information sources can potentially lose their identity in time, which implies that recognition, reconstruction or reproduction of identity must either be based on properties that are dynamical- or conformal invariants (symmetries with respect to space or time), or, the loss of identity is completely governed (predicted) by known ODPs such as decay and diffusion or simply due to sources of additive noise on the channel. Shannon's theory of communication is of the latter kind, a nomothetic theory which tells us how to reproduce a message using laws of logic and probability. Several decades of research (and indeed Shannon himself) suggests that this theory cannot sufficiently explain the complex adaptive behavior of living organisms. The solution could be to complement the nomothetic theory of communication

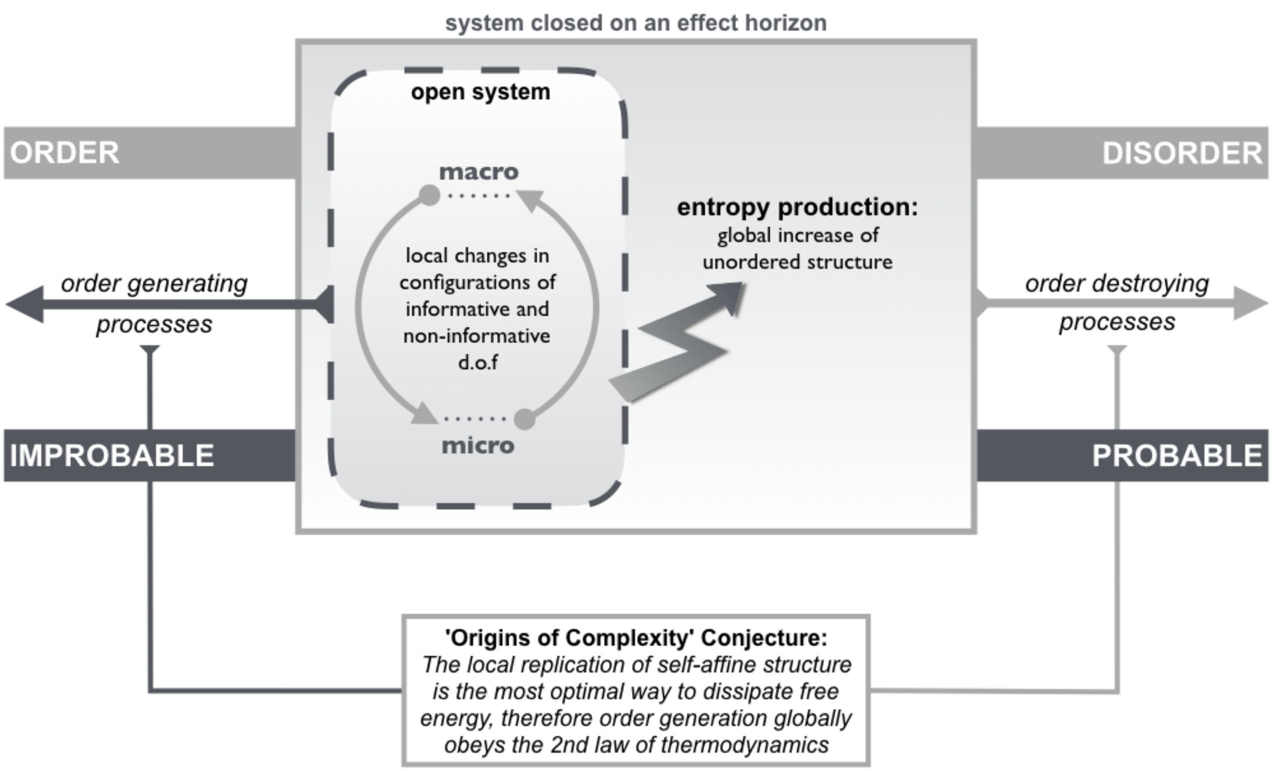

Figure 2. Schematic representation of the proposed relations between the emergence of order and disorder and thermodynamically probable and improbable structure.

with a theory about the reproduction of messages based on dynamical and physical structure that is at least partially invariant across a range of spatial and temporal scales. The post-cognitivist research programs mentioned earlier identify such invariants as the regularities of the environment that can constitute ecological niches in which skilled agents can emerge. What is needed is the conception of a communication channel that can transmit, or, recognize invariants across different scales. I will argue that the phenomenon of multi-fractal scaling observed in time series of human performance and physiology suggests that detecting similarities (identities) across different scales of fluctuation is essential for understanding the emergence of meaning associated with the adaptive behavior of complex systems.

3.2 Self-organized Codes

Meaning can be linked to properties of physical structures and to order generation:

> "Current theories that incorporate the notion of information are not theories of information content, at least not content in the form of spatially-extended structures, such as words, computer programs and pictures. Such theories refer to content only indirectly through some formulation of information measure." (Boyle, 1992, p. 52)

Boyle's conjecture again seems sound: Meaning as a topic of inquiry in physics must always refer to a spatially extended physical structure whose description in terms of 'amount of information' is subject to change in accordance with the laws of physics. This perspective is also taken in the field of biological semiotics (cf. Barbieri, 2003) in which the concept of an organic code as an information source plays an important role in understanding order generation. Physical manifestations of message coding systems are rather common in nature. In organic chemistry it is known that two information sources (nucleotides and amino acids) can be connected by a coding structure (transfer RNA) by means of the operation of *translation* which consists of two physical recognition processes. Other operations commonly identified in organic chemistry are transcription (requires only 1 recognition process), splicing, folding and assembly (cf. Barbieri, 2003). One way to re-frame what it is organic codes actually do is that they can give meaning to information structures by translating their similarities, that is, by recognizing, or exposing (signaling) the invariant structure that exists between them. The transfer RNA molecule can be said to connect the world of nucleotides (DNA) to the world of amino acids (Barbieri, 2003, p. 98). In terms of information theory, a message is transmitted from the world of nucleotides, encoded, channeled and decoded through the structure of the tRNA, to the destination world of amino acids. The system as a whole is the physical realization of the correlation between two different (id-)entities, it is the embodiment of contextual meaning.

Based on this formulation of a code in terms signaling the structural or dynamical similarities that exist between two different information sources, three ingredients can be identified that may be necessary conditions for the emergence of meaning in physical systems:

1. An identity divide: Information sources that represent 'different worlds' according to some criterion of independence (e.g., separation by temporal or spatial scales, different constituent components of the information sources, Markov blankets, adiabatic boundaries).

2. A code, or codon/analogon pair (Walker, 1983), that is able to couple the information sources across an identity divide, through its structural configuration and/or through its dynamical behavior (establishing a communication channel of sorts).

3. Processes of physical recognition that allow the analogon to translate the similarities that exist between the information sources and expose them as redundancies, or invariants.

A better way to describe ingredient 3 is that these codes give meaning to information sources by reproducing the similarities (redundancies) that exist between them by an analogy. Walker (1983) notes that in the context of reproduction, the following applies to 'analogy':

> "Two separately identifiable patterns are related by analogy if the existence and frequency of the one is correlated with the existence and frequency of the other in the absence of direct forces between the two patterns that could cause the correlation. That is, correspondence between codon and analogon came about, and is maintained, by reproduction of an initial random event." (Walker, 1983, p. 809)

The point here is that if a force would be responsible for the correlation to be captured by the analogy, there would be no need to map the correspondence between the structures. The code to establish a translation in that case, would be a universal law of physics and we already have an excellent theory for those types of correspondences. The same holds for the arrangement of code components, if some force dictates their order, they cannot efficiently code for anything, except perhaps characteristics of the force itself. From a meta-theoretical perspective, stripped from any context of application, the anologon[4] represents a structure or regularity in reality that systematically covaries with information sources.

Table 1

Definitions of Concepts to Describe Complex Adaptive Behavior According to REC++

Complexity	The effective complexity of an entity is the length of a highly compressed description of its regularities; the redundancies in a system.
Information	A measurable quantity that resolves uncertainty about a system state by specifying the state (e.g., through measurement). Information can only distinguish one thing from another within the context of uncertainty and probability defined for the level at which the degrees of freedom of the system are identified. Deciding on the difference between meaningful and meaningless information is not a subject of information theory.
Entropy	Given a macro-state description of a system, it is a quantification of the number of degrees of freedom a system has available at the micro-scale to generate the global behavior. The entropy of a thermodynamic system is maximal when there is a complete absence of order in the distribution of its energy, that is, its internal structure will have equiprobable states in maximal disorder and require maximal information to describe.
Order Generating Process	Any state propagation rule or system configuration that increases the order (= decreases the entropy) of the internal structure of a system. An OGP changes the amount of information needed for a complete description of the system. Adiabatic order generating processes allow isentropic transformation of energy. The embodiment of the introduction of redundancy into a system.
Information representation	The entropy associated with the observation of a macro state. Systems represent maximal information if the entropy is maximal.
Meaning	Semantic aspects of messages that "*refer to or are correlated according to some system with certain physical or conceptual entities*" (Shannon, 1948). Meaning is the recognition of identity between information sources that are separated by spatial and/or temporal scales. Meaning emerges due to the reproduction of similarity by analogy.
Interaction biography	An explanatory tool to refer to the interaction history of a system. Biographical events can concern interactions with the environment or interactions between subsystems. The concept is comparable to the epigenetic landscape.
Adaptive behavior	An observable response to changes in the external environment or in constituent subsystems of a complex system that appears to be (partially) coordinated by biographical events or informative structure.
Effect Horizon	A horizon that separates causal entailment of processes on different spatiotemporal scales into immediate (permissive) and mediative (causative) with respect to explaining observed behavior.

[4] I will use analogon to refer to the function performed by tRNA: Connecting two worlds by a spatially-extended structure. This can be due to a process of physical recognition, self-affine scaling or a complex of transmitting (encoding a message) and receiving (decoding a message).

The analogon can therefore be defined as a construct specifying systematic relations between things of dissimilar identity through a non-specific configuration of its own structure. As an example, the choice of symbols to mathematically represent a lawful relation between physical quantities is independent of the actual physical relation it specifies.

3.3 Beyond the Effect Horizon: Permissive and Causative Levels of Analysis

Marr and Poggio (1976) argued there are three levels of analysis for complex information processing systems: A computational level, an algorithmic level, and a level of implementation. In Radical Embodied Computation, the different levels of observation and analysis involved in describing complex adaptive behavior in fact describe the computational process itself, the generation of thermodynamically improbable order through reproduction of similar patterns at different levels of aggregation.

The concept of the analogon embodies the process of physical recognition of identity or invariant structure between things separated by an identity divide. This is relevant for understanding how coordination relative to temporal scales could self-organize in physical systems. A system can exchange energy across an *adiabatic boundary* without dissipating entropy. When applied to physical change processes, it is the phenomenon that change processes at faster time scales will take the value of processes fluctuating at much slower frequencies as their parameter settings (cf., Hopfield, 1994). The parameters act as constants, as a landmark for temporal orientation that can be used to coordinate behavior. In more established mathematical theories that seek to describe (quasi-)periodic state space dynamics as topological invariants (e.g., KAM-theory, Aubrey-Mather theory) there is a notion of separation (or partitioning) of dynamics according to iso-surfaces of some physical quantity (temperature, entropy). At even larger divides of identity, the *renormalization group* connects degrees of freedom that exist at physical scales described by different theories (e.g. scale transformations in QED, Gell-Mann and Low, 1954). From the perspective of a classical system, quantum mechanical laws operate at a scope beyond a horizon spanned by the Planck scale. The Planck scale could be described as a barrier for quantum phenomena that are coarse grained to emerge as structural or dynamical parameters for phenomena at the classical scales.

Behavior of systems that is coordinated against an apparently static background provided by an adiabatic-like separation can be understood and modeled as the coarse graining of one or more dimensions in state space (Ashby, 1947; Hopfield, 1994; Flack, 2017a). It involves removing, fixing, or averaging a range of values from the complete microscopic description of the state space see e.g., (Izvekov and Voth, 2005; Gohlke and Thorpe, 2006; Fuchslin et al., 2009). Expressed in terms of degrees of freedom in state space, coarse graining implies fixing some of the degrees of freedom from the complete micro-scale description of the system and formulated this way, it makes sense to interpret the coarse-grained dimensions to represent informative structure. Its configuration has become correlated to the regularities of some other information source, which happens to be a process fluctuating at a much slower time scale. For now, I assume it is reasonable to assume such coarse graining boundaries can emerge whenever the traces of efficient causes of observed states are "lost" across a temporal or spatial scale. This will often be more or less symmetrical to the prediction horizon of the state evolution of a system at a given point in time, but in this case obscuring the past.

I will use the term *Effect Horizon* to denote a lag of time or space beyond which efficient causes of current states are no longer identifiable, or, are no longer sensible to consider as such. Quantum physical phenomena must be considered permissive of phenomena at larger scales, but we can use thermodynamics to describe those phenomena. Likewise, it is clear that for the macro scale behavior of biological systems it will not be sensible to declare quantum phenomena as efficient causes, although there is no doubt that the macro state would not be able to manifest itself without those quantum phenomena. The same holds for the biological example of speciation. Genes are permissive of phenomena observed at the macro-scale of the organism, but they are not causative in a direct way. On a bio-chemical level of analysis, it is not wrong, but also not very informative, to suggest the adaptive behavior of an organism concerns an ancient event in the evolutionary history of its species that is in some way represented by its genome. Such nested events *mediate* causation, but cannot be considered *immediate* causes. Tracing back the efficient causes of an individual organism to generations far beyond the moment of its own conception in order to explain its current adaptive behavior might be possible to some extent, but it does not seem very sensible.

3.4 Spatially Extended Representation of Self-Organized Meaning

The placement of effect horizons at particular temporal and spatial scales on which coarse grained degrees of freedom appear as static parameters can be evident from a physical perspective, but this may be more arbitrary and depend on the specifics of the phenomena under scrutiny in cognitive science. Here, digital physics does offer a potential model, in relativistic theories of space and time, accelerated frames of reference will allow effect horizons to emerge between any separation of scales (Verlinde, 2011). All that is needed is a nested spatial or temporal hierarchy where smaller or faster scales can take larger or slower scales as a coarse-grained constant. The structures and processes on the other side of the horizon can be thought of as permissive of the adaptive behavior of interest.

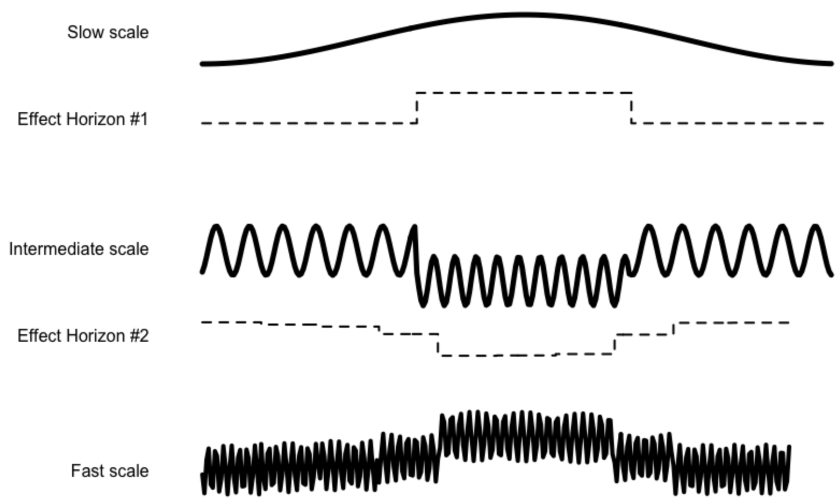

Figure 3. Coarse graining due to nesting of temporal scales. Slow processes will appear as fixed degrees of freedom, or constants from the perspective of faster changing processes, for prolonged periods of time. The figure shows how coarse-grained variables can appear as a step function input, or as a coupled variable to other nested processes. In the figure, the level of each nested process represents an inverse relation to the level of the step function (e.g. Level of sunlight: day/night → Mean level of heart-rate: awake/asleep → Mean level of cell activity due to production of neuro-mediators while asleep)

In terms of adaptive behavior and memory function, the key is to imagine the adaptive behavior occurs with respect to a scale of fluctuation. This is a temporary screen on which an invariant structure is projected that allows coordination of behavior across lags of time. The lag of time to which coordination of future behavior is observed depends on the rate of fluctuation of processes behind the screen. Circadian rhythms, seasonal phenomena, and reproductive cycles are all examples of changes in the state of the internal or external environment of a biological system that coordinate the dynamics of faster changing processes across temporal scales. The permissive structure for the coordination of some of this behavior concerns the celestial mechanics of the solar system, but it does not seem sensible to include the physics of celestial mechanics in the formal description of efficient causes of a biological cycle, other than in terms of (related) coarse-grained variables: Amount of light during the day, temperature of the environment, availability of liquid water and food, etc.

By observing the behavior of a complex adaptive system relative to an Effect Horizon, one can consider it temporarily as a subsystem of a closed system. That is, loss, gain, or transfer of energy or entropy beyond the horizon are (temporarily) considered irrelevant, because they are unknown. Moreover, such horizons provide boundaries that allow us to identify a level of analysis at which the number of degrees of freedom to consider can be quantified for a specific context of observation.

In what follows I conjecture that the concept of scale-invariance, which can be regarded as a nested causal structure, especially when time scales are concerned, can account for the 'natural' emergence of an Effect Horizon. The horizon obscures the dynamics of those processes that may be considered immediate causes of the dynamics of observable processes, due to the relatively slow scale at which they fluctuate. The multi-fractal formalism known as the multiplicative cascade (see e.g., Ihlen & Vereijken, 2010, 2013; Stephen, Anastas & Dixon, 2012) can be interpreted as a formal description of accumulations of broken symmetries, as singularities in a hierarchy of scales of fluctuation. The symmetry between spatial and temporal extent can be recovered by estimating an exponent indicating the scaling relation between the two axes. The branch-like bifurcations in Figure 5 are singularities indicating an emergent horizon beyond which the scaling exponent that is needed to detect similarity, will be different. The multi-fractal, or singularity spectrum represents the distribution of scaling exponents (singularities) in a signal. It is a representation of the entire nested, scale invariant structure of a signal.

4 Complexity Matching: A Conjecture about Reproduction of Self-Affine Structure

Evidence is accumulating that humans are able to coordinate their behavior by exploiting specific invariant properties of complex dynamical patterns either due to attraction to criticality or complexity matching. Attraction to criticality refers to the association between fractal-like scaling observed in health and well-being (cf. Goldberger, Amaral, Hausdorff, Ivanov, Peng & Stanley, 2002; Van Orden, Kloos & Wallott, 2009) and proficiency and fluency of performance (e.g., in motor learning Wijnants, Bosman, Hasselman, Cox & Van Orden, 2009; or as nested constraints on performance, Wijnants, Cox, Bosman, Hasselman & Van Orden, 2012). The Complexity Matching or Complexity Control hypothesis posits that humans make use of the invariant structure of complex dynamical patterns to coordinate their behavior in ways that are comparable to principles for optimal and maximal information transport between complex systems as posited by formal fluctuation dissipation theorems. The self-affine structure can be regarded as a complex resonance frequency (e.g., the 1/f resonance hypothesis; Aquino, Bologna, West, Grigolini, 2011) or complex stochastic resonance (Kelty-Stephen & Dixon, 2013). Evidence of selective matching of dynamical behavior to scaling exponents in different observables measured simultaneously throughout the body, suggests a complex multi-scale coupling relationship between physiological and psychological processes may exist (Rigoli, Holman, Spivey & Kello, 2014). Complexity matching has also been reported for dyadic interactions, for example interpersonal coordination of coupled movements (Marmelat & Delignires, 2012) and overt behavior during joint problem solving (Abney, Paxton, Dale & Kello, 2014). In a direct comparison of different characterizations of speech stimuli (Hasselman, 2015), a simple classifier (Quadratic Discriminant Analysis) reproduced the labeling of stimuli by participants most accurately when it was based on the multi-fractal spectrum (96.6% correct), a phenomenon that has been reproduced recently (Ward and Kelty-Stephen, 2018).

4.1 Permissive structures for Radical Embodied Computation

One of the key components to reproducing complex configurations of the internal or external environment has to be a physical principle able detect (or discriminate between) identity across lags of time and the easiest way to achieve this is counting:

> "Counting as distinction in time leads to identification. Counting, […], is based on the faculty to "remember" the former event at the advent of a later such event. […] Counting as an irreversible adding process within memory would provide a perfect absolute time axis for long term memory, against which all other time events are evaluated." (Walker, 1972)

Suppose that every time a certain event occurs, a counter is updated, then, until a similar event recurs, the former is 'remembered', or, similarity is recognized. A ranking, or a direct comparison of the values of counters associated with states, would provide an immediate nested structure for the description, or, the re-presentation of the states in terms of the 'values' of counters. Counting by summation is irreversible and like chirality for spatial orientation (left right asymmetry), it allows orientation with respect to the arrows of time. A well-known biological counter keeps track of 'the Hayflick limit' (Hayflick and Moorhead, 1961). There is an apparent limit to how many times cells can replicate (50--60 times) and this happens because structures called telomeres prevent DNA polymerase to be fully copied, each subsequent replication is a slightly degraded copy of the original with respect to the telomeres (for a historical review of discoveries associated with the Hayflick limit see Shay and Wright, 2000). Built into the heart of the genetic memory system is a replicometer and what is interesting for the story so far, is the observation that the replicometer counts down. Compared to the previous copy, the telomere attached to the DNA molecule is shorter and specifies the current copy of the molecule as a structure of lesser complexity. This reveals that order generating and destroying processes share a complementary existence proof (Kelso and Engstr\o m, 2006). The order generating process of counting DNA replications is achieved by destroying ordered structure, interestingly, Dolly the 'copied' sheep, had shorter telomeres in all the cells of her body than age matched non-cloned sheep did. It was as if the copy of the whole had been counted by each cell in the body (Shiels et al., 1999; Shay and Wright, 2000).

In the nervous system, the obvious candidate for the replicometer, which represents the embodiment of a process of summation counting, is the oscillation of neurons. Figure 4 displays some of the neural codes generated by the auditory system which is a beautiful example of how physical regularities of sound pressure waves are coded into a spatially extended structure: The cells in the cochlea are tonotopically organized and respond when the sound pressure wave hits its maximum, the 'when' occurs always at the same location because this represents the frequency of the wave oscillation. The tonotopic organization represents meaningful information about the structure of the world. The oval window at the base of the cochlea is an effect horizon, it translates oscillations of the air pressure into movements of the cochlear liquid. The entire system reproduces sound pressure oscillations by analogy. When the signals enter the auditory cortex the immediate connection to the dynamics of the wave oscillations are lost and the firing rates of the auditory neurons initiated by the movements of the hair cells in the cochlea form an effect-horizon for the oscillations of the neurons in the auditory cortex. The auditory cortices are also tonotopically organized, here different codes for frequency and intensity are generated by synchronously oscillating neurons. There are also asynchronous neurons, and their behavior resembles summation counting, their firing rate will increase if the frequency with which similar sounds are repeatedly presented increases. More complex oscillatory patterns have been found to travel orthogonally to the tonotopic organization which can be thought of as an amplitude modulating wave (Baumann et al., 2015).

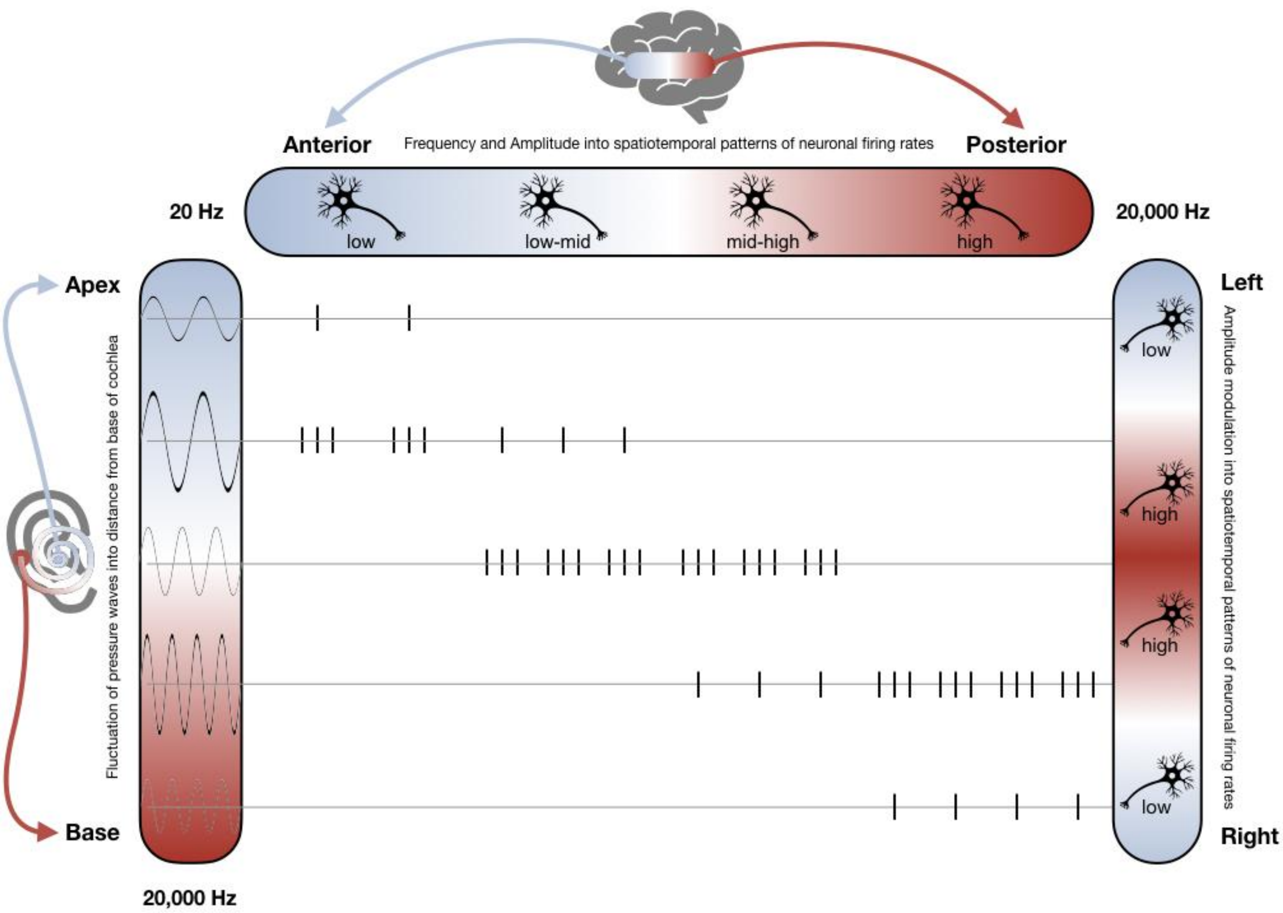

Figure 4. Coding for meaning by the auditory system, see text for details.

Figure 6 is a representation of a 6 second waveform representing the spoken word /bAk/ [vat] (the scaleogram of the envelope is displayed in Figure 5). It is a dynamic network (a hyve plot, see https://youtu.be/Fh4IMuzJUlQ for the animation) in which the lines represent scaling relations identified by a continuous wavelet transform of the signal (a scaleogram). The scales are ordered from small to large along the three axes of the plot with smallest scales in the center. Edges connecting from 're-mediate' to 'im-mediate' (the Now) affirm that a similar, scaled-down structure was perceived at an earlier point in time. Edges projecting toward what is yet to come, especially those that span a large range of scales can be thought of as anticipatory cues. The lines thus indicate whether some component of the sound pressure wave that is present right now, is in fact a self-affine reproduction of a component that was present a moment ago, or whether it is likely a self-affine copy will be present in the near future. The experiments that reveal listeners make use of aspects of the multi-fractal spectrum (Hasselman, 2015; Ward and Kelty-Stephen, 2018) suggest that at least our auditory system is somehow able to tap into this extremely rich source of meaningful information. Given its organization, it seems reasonable to suggest that information about similarities and affinities of the sound pressure wave that exist across different scales of fluctuation can be decoded into neural firing rates by the auditory system. Recently, such evidence has been provided in a study that assessed brainstem responses to the spoken syllable /da/, which found the brainstem response to be multi-fractal in nature, characterized by the presence of long-range temporal correlations (Mozaffarilegha and Movahed, 2019).

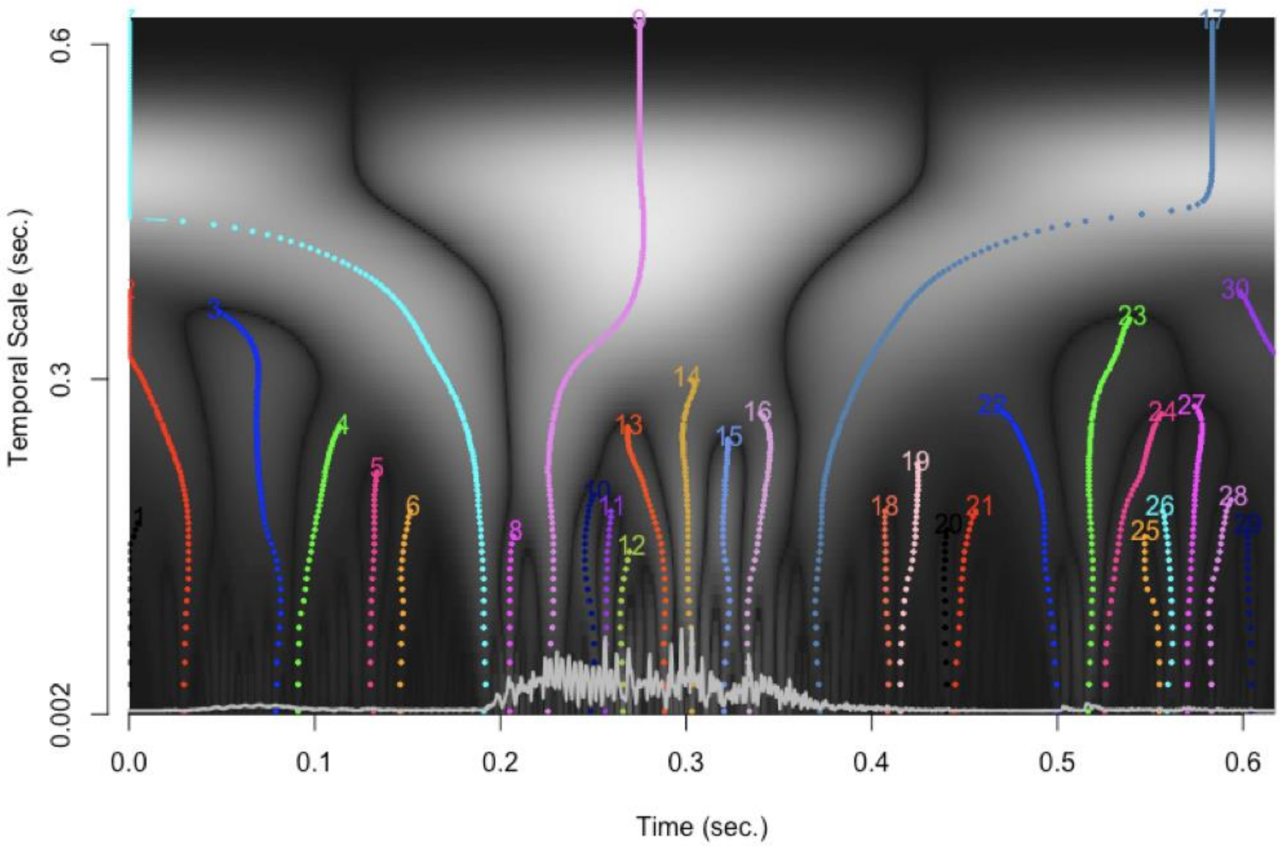

Figure 5. A scaleogram representing a Continuous Wavelet Transform (CWT) of the amplitude envelope of an audio recording of the word /bAk/ (the grey line). The colored traces that connect the different scales represent the wavelet singularity extrema obtained by the Wavelet Transform Modulus Maxima (WTMM). They indicate the association between the scaled wavelet and the envelope as it unfolds, extends across many different temporal scales as a self-affine structure. The numbers at the end-points of the traces represent singularities an analogon, connecting processes at different scales of fluctuation.

4.2 Self-organized reproduction of similarity by analogy

The examples in this section point to plausible physical implementation of the reproduction of similarity[5] (identity, redundancy) across different spatial and temporal scales, the emergence of effect-horizons, the operation of summation counting, but more importantly, that the process of reproduction of similarity by analogy itself constitutes the emergence of a mechanism that, for all

[5] The term affinity is in fact more appropriate when referring to the scaling relation, because it is not a zoom (self-similar) transformation, but a warp (self-affine) transformation.

intents and purposes can be considered a self-organized coding system. I suggest that the spectra of scaling exponents identified by multi-fractal analyses of time series of observables of complex adaptive systems reveal the richness of the nested structure that exists all around us. This structure can be detected and reproduced through very simple mechanisms of which several examples have been discussed. The oscillations of tonotopically organized neurons in the auditory cortex is an example, but any physical system that can be thought of as a self-tuning resonator or oscillator organized in some spatial or temporal hierarchy would constitute a permissive structure for the reproduction of similarity by analogy. From the perspective of the efficiency of generating entropy through self-replication (England, 2013), the existence of evolved systems that perform functions that can be considered the approximate reproduction of patterns should not be surprising.

Recently Kolchinsky and Corominas-Murtra (2019) presented a way to decompose the mutual information that may exist between two system states into copied information and transformed information. Further study is necessary in order to support the following conjecture, but it seems too obvious to not present it right now: An information theoretic analysis of our sensory systems (specifically those which are topologically organized to represent properties of the physical world), should reveal that these systems recognize similarity (quantifiable by mutual information) and partially copy and transform it (quantifiable by transfer entropy).

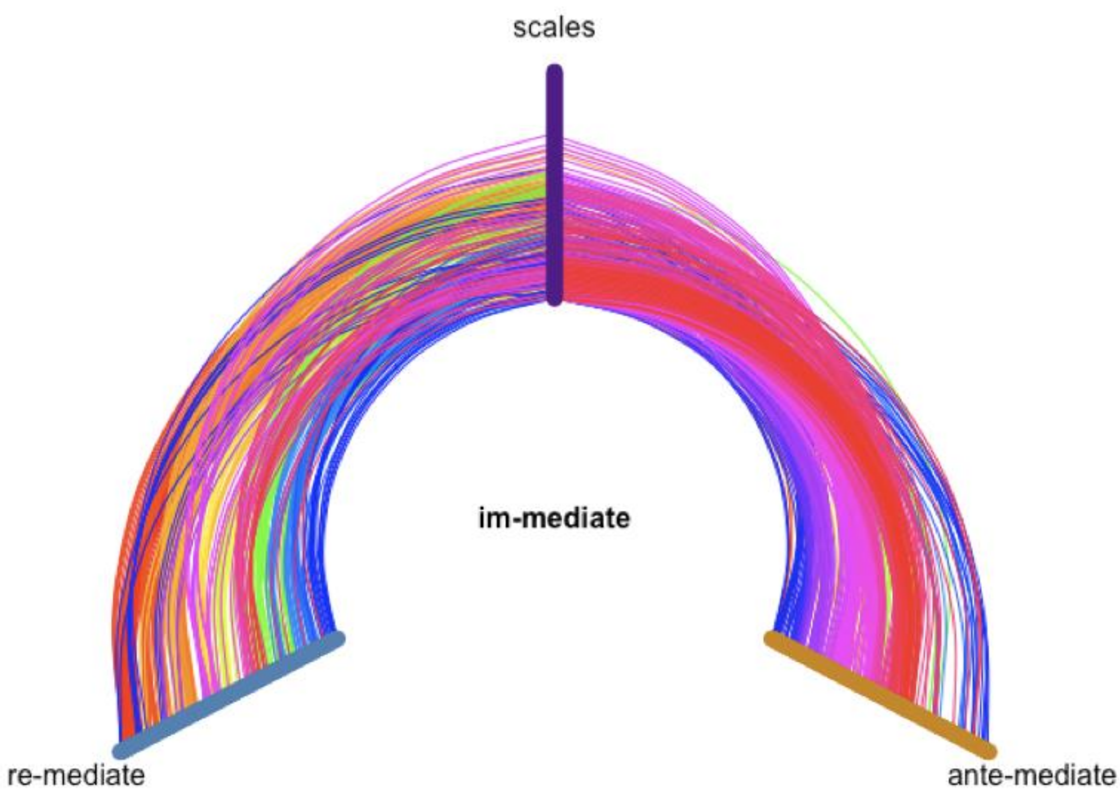

Figure 6. Animation of forward and backward reproductions of similarity by analogy in a 6 second recording of the word bAk [vat]. For an animation see **https://youtu.be/Fh4IMuzJUlQ**.

5 Discussion

To summarize, I argue that theories of complex adaptive behavior, such as human cognition, should focus on developing a physics of meaning, a theoretical framework that allows for a more formal treatment of the role of semantic information in coordinating behavior. This will require the adoption of some form of information theory for natural computation in complex adaptive systems, but at the same time, one would like to avoid a regress to theories based on an 'interpretative' view of computation, or invoking a Virtual-Physical dualism. I suggest recent developments in digital physics and natural computation can offer a solid foundation for a research program of Radical Embodied Computation in which the plausibility of the physical implementation of computational mechanisms should play a central role. There are three main principles a REC++ theory should adhere to:

1. *Physically plausible order generation*: Positing an Order Generating Process in order to explain complex behavior requires a plausible account of the generation of thermodynamically improbable order (and its destruction). Moreover, OGPs should not contradict contemporary theories of physical information and dynamics of complex systems and networks. The 'free-energy principle' (Friston, 2012) and the 'statistical physics of self-replication' (England, 2013) are examples of accounts that are plausible, in principle. The veracity of such accounts will of course have to be established by repeated application of the scientific method.
2. *Respect the medium*: Ordered mass or energy that is posited to exist as an internal structure or state of a complex system, must be (bio)physically realizable, in principle. Identifying real-world structures or phenomena that are at least permissive of representing such order through (some hierarchy of) state configurations, should be a priority.
3. *Account for the emergence of meaning*: Any explanations of the behavior of a system at the causative level that refer to specific interactions that occurred in the past of a system should be accounted for in terms of semantic information. This implies some kind of coding mechanism should be posited. The self-organized reproduction of similarity by analogy in the present paper is an example of such a mechanism. Naturally, principles 1 and 2 apply when positing such mechanisms.

The example of the auditory system provided in the previous section will need more empirical support, but as a first exercise to reframe the functions performed by the auditory system in terms of radically embodied natural computation, appears valid. No loans on intelligence are necessary so far. The auditory speech signal is a multifractal sound pressure wave, which means there is a spectrum of scaling exponents that indicate redundancies exist as self-affine copies at different scales of fluctuation. The auditory system contains several structures that can be thought of to perform a function of the analogon, connecting the world of sound pressure waves to the world of neural firing rates, for example by translating the movement of the oval window of the cochlea into movement of the inner ear fluid, into the stimulation of hair cells at a specific distance from the base of the cochlea that corresponds to the peak intensity of a frequency component of the sound pressure wave. It is likely that the multifractal structure of the wave is translated into neural firing rates through this system. One way this could be achieved is by a simple network of ensembles of neurons that perform summation counting (e.g. by so-called asynchronous neurons). Behavioral experiments have so far been able to evidence that human listeners make use of the properties of the multi-fractal structure to discriminate between different speech sounds. Further study is needed to establish whether other types of meaningful information are encoded in this spectrum of scaling exponents.

5.2 Radical Embodied Computation for Robotics and AI?

In the final activity report of the research project MACS (Multi-sensory Autonomous Cognitive Systems Interacting with Dynamic Environments for Perceiving and Using Affordances; Rome, 2008) it is mentioned that a valid starting point for the project was to draw inspiration from cognitive science, being ecological psychology and embodied, embedded cognition and situated cognition (Rome, 2008, p. 5). This is rather surprising, because cognitive scientists generally do not consider ecological psychology and related research programs to be a part of cognitive science, but apparently other disciplines disagree, and I would argue they are correct to take the knowledge generated by these disciplines as a starting point, because this concerns some of the most formal and principled accounts of explaining complex adaptive behavior, applicable to all skilled agents, not only humans. However, I also argue the post-cognitivist research programs will have to reconsider their objections to using concepts related to 'computation' and 'information processing', especially because these objections are likely due to the conflation of semantic information with the information syntax or quantity, a problem that can be resolved by adopting principles of natural computation and digital physics. This will merely translate some of the thermodynamic principles their theories are already based on into the formal language of physical information.

One could question why applied Robotics and AI should be concerned with the principles of REC++ laid out in this paper. After all, why should an engineer be concerned with the plausibility of biophysical implementation of a particular computational architecture, as long as it fulfils the functions it was designed for? This is a valid point and from this perspective an interpretative view of computation would be unproblematic. An example of differences in goals between theorists and engineers concerns an extension to the affordance ontology suggested by Şahin et al. (2007), as a result of the aforementioned MACS program, which would allow affordance-based robot control. The response from ecological psychologists (Chemero and Turvey, 2007b) was that the extension may allow for intelligent adaptive robot behavior, but could not serve as an ontology for a theory of Gibsonian direct perception. Artificial direct perception of Gibsonian affordances would require non-trivial adaptations to the computational architecture of artificial agents to deal with impredicativities and non-well-founded sets (Aczel, 1988; Chemero and Turvey, 2007a; Chemero and Turvey, 2008). This may prove to be an endeavor roboticist are currently not eager to undertake.

One way to interpret this exchange in the context of the present paper is as follows: The concept of the affordance, the action opportunities provided by the environment, that may be directly perceived by an agent given the action capabilities provided by its internal structure, is a very powerful explanatory vehicle for understanding complex adaptive behavior and therefore a very promising control mechanism for the development of artificial agents that have to navigate such environments. What is apparently missing from current theoretical descriptions of the direct perception of affordances are details related to the 3 principles of REC++, what exactly does this process of direct perception entail in terms of dynamic control mechanisms such as self-organization, multi-stability, emergence or self-affine scaling? These details will become more important and will eventually have to be specified by theorists and translated into formal models, irrespective of the goal being the advancement of scientific knowledge or the development of technology. I conjecture that any control mechanism that seeks to exploit the invariant structure of the environment for the coordination of complex adaptive behavior will have to be based on a self-organizing system that performs the simple function of reproducing similarity by analogy.

References

't Hooft, G. (1994). "Dimensional reduction in quantum gravity", in: *Salamfestschrift: a collection of talks*. World Scientific).

Aczel, P. (1988). *Non-well-founded sets. Foreword by Jon Barwise.* CSLI Lecture Notes, 14.

Anderson, M.L. (2003). Embodied cognition: A field guide. *Artificial Intelligence* 149(1), 91-130. doi: 10.1016/S0004-3702(03)00054-7.

Ashby, W.R. (1947). Principles of the self-organizing dynamic system. *J Gen Psychol* 37(2), 125-128. doi: 10.1080/00221309.1947.9918144.

Atkins, P. (2010). *The laws of thermodynamics: A very short introduction.* OUP Oxford.

Badcock, P.B., Friston, K.J., and Ramstead, M.J.D. (2019). The hierarchically mechanistic mind: A free-energy formulation of the human psyche. *Phys Life Rev* 31, 104-121. doi: 10.1016/j.plrev.2018.10.002.

Baddeley, A. (2012). Working memory: theories, models, and controversies. *Annu Rev Psychol* 63, 1-29. doi: 10.1146/annurev-psych-120710-100422.

Barab, S.A., Cherkes-Julkowski, M., Swenson, R., Garrett, S., Shaw, R.E., and Young, M. (1999). Principles of self-organization: Learning as participation in autocatakinetic systems. *Journal of the Learning Sciences* 8(3-4), 349-390. doi: DOI 10.1207/s15327809jls0803&4_2.

Barbieri, M. (2003). *The Organic Codes: An Introduction to Semantic Biology.* Cambridge University Press.

Baumann, S., Joly, O., Rees, A., Petkov, C.I., Sun, L., Thiele, A., et al. (2015). The topography of frequency and time representation in primate auditory cortices. *Elife* 4. doi: 10.7554/eLife.03256.

Bekenstein, J.D. (2003). Information in the holographic universe. *Sci Am* 289(2), 58-65. doi: 10.1038/scientificamerican0803-58.

Berut, A., Arakelyan, A., Petrosyan, A., Ciliberto, S., Dillenschneider, R., and Lutz, E. (2012). Experimental verification of Landauer's principle linking information and thermodynamics. *Nature* 483(7388), 187-189. doi: 10.1038/nature10872.

Bosman, A.M.T., Cox, R.C.A., Hasselman, F., and Wijnants, M.L. (2013). From the Role of Context to the Measurement Problem: The Dutch Connection Pays Tribute to Guy Van Orden. *Ecological Psychology* 25(3), 240-247. doi: 10.1080/10407413.2013.810091.

Boyle, C.F. (1992). "Physical Laws And Information Content", in: *Workshop on Physics and Computation*. IEEE).

Boyle, C.F. (1994). Computation as an intrinsic property. *Minds and Machines* 4(4), 451--467.

Bruineberg, J., Chemero, A., and Rietveld, E. (2019). General ecological information supports engagement with affordances for 'higher' cognition. *Synthese* 196(12), 5231-5251. doi: 10.1007/s11229-018-1716-9.

Burov, S., Jeon, J.H., Metzler, R., and Barkai, E. (2011). Single particle tracking in systems showing anomalous diffusion: the role of weak ergodicity breaking. *Phys Chem Chem Phys* 13(5), 1800-1812. doi: 10.1039/c0cp01879a.

Carello, C., Turvey, M.T., Kugler, P.N., and Shaw, R.E. (1984). "Inadequacies of the Computer Metaphor," in *Handbook of Cognitive Neuroscience*.), 229-248.

Chemero, A. (2009). *Radical embodied cognitive science.* The MIT Press.

Chemero, A., and Turvey, M.T. (2007a). Complexity, hypersets, and the ecological perspective on perception-action. *Biological Theory* 2(1)**,** 23--36.

Chemero, A., and Turvey, M.T. (2007b). Gibsonian Affordances for Roboticists. *Adaptive Behavior* 15(4)**,** 473--480. doi: 10.1177/1059712307085098.

Chemero, A., and Turvey, M.T. (2008). Autonomy and hypersets. *Bio Systems* 91(2)**,** 320--330. doi: 10.1016/j.biosystems.2007.05.010.

Chialvo, D.R. (2010). Emergent complex neural dynamics. *Nature Physics* 6(10)**,** 744-750. doi: 10.1038/Nphys1803.

Clark, A. (1997). *Being There. Putting Brain, Body, and World Together Again.* The MIT Press, USA.

Clark, A. (1999). An embodied cognitive science? *Trends Cogn Sci* 3(9)**,** 345-351. doi: 10.1016/s1364-6613(99)01361-3.

Davis, T.J., Kay, B.A., Kondepudi, D., and Dixon, J.A. (2016). Spontaneous Interentity Coordination in a Dissipative Structure. *Ecological Psychology* 28(1)**,** 23-36. doi: 10.1080/10407413.2016.1121737.

Decastro, L. (2007). Fundamentals of natural computing: an overview. *Physics of Life Reviews* 4(1)**,** 1-36. doi: 10.1016/j.plrev.2006.10.002.

Dennett, D.C. (1981). *Brainstorms: Philosophical Essays on Mind and Psychology.* Harvester Press.

Desurvire, E. (2009). *Classical and Quantum Information Theory: An Introduction for the Telecom Scientist.* Cambridge University Press.

England, J.L. (2013). Statistical physics of self-replication. *J Chem Phys* 139(12)**,** 121923. doi: 10.1063/1.4818538.

Flack, J.C. (2017a). Coarse-graining as a downward causation mechanism. *Philos Trans A Math Phys Eng Sci* 375(2109). doi: 10.1098/rsta.2016.0338.

Flack, J.C. (2017b). "Life's Information Hierarchy," in *From Matter to Life - Information and Causality*. Cambridge University Press).

Floridi, L. (2003). Informational realism. *Selected papers from conference on Computers and \ldots* 37.

Fodor, J. (1983). *The modularity of mind.* Cambridge, Mass.: MIT press Cambridge, MA.

Fodor, J.A., and Pylyshyn, Z.W. (1988). Connectionism and cognitive architecture: a critical analysis. *Cognition* 28(1-2)**,** 3-71. doi: 10.1016/0010-0277(88)90031-5.

Fredkin, E. (1990). Digital mechanics. An informational process based on reversible universal cellular automata. *Physica D: Nonlinear Phenomena* 45**,** 254-270.

Freeman, W.J. (1997). Three Centuries of Category Errors in Studies of the Neural Basis of Consciousness and Intentionality. *Neural Netw* 10(7)**,** 1175-1183. doi: 10.1016/s0893-6080(97)00054-3.

Freeman, W.J., Kozma, R., and Werbos, P.J. (2001). Biocomplexity: adaptive behavior in complex stochastic dynamical systems. *Biosystems* 59(2), 109-123. doi: 10.1016/s0303-2647(00)00146-5.

Friston, K.J. (2012). A Free Energy Principle for Biological Systems. *Entropy (Basel, Switzerland)* 14(11), 2100--2121.

Fuchslin, R.M., Fellermann, H., Eriksson, A., and Ziock, H.J. (2009). Coarse graining and scaling in dissipative particle dynamics. *J Chem Phys* 130(21), 214102. doi: 10.1063/1.3143976.

Gell-Mann, M., and Low, F.E. (1954). Quantum Electrodynamics at Small Distances. *Physical Review* 95(5), 1300-1312. doi: 10.1103/PhysRev.95.1300.

Gibson, J.J. (1966). The problem of temporal order in stimulation and perception. *J Psychol* 62(2), 141-149. doi: 10.1080/00223980.1966.10543777.

Gohlke, H., and Thorpe, M.F. (2006). A natural coarse graining for simulating large biomolecular motion. *Biophys J* 91(6), 2115-2120. doi: 10.1529/biophysj.106.083568.

Golonka, S., and Wilson, A.D. (2019). Ecological Representations. *Ecological Psychology* 31(3), 235-253. doi: 10.1080/10407413.2019.1615224.

Grignolio, A., Mishto, M., Faria, A.M.C., Garagnani, P., Franceschi, C., and Tieri, P. (2014). Towards a liquid self: how time, geography, and life experiences reshape the biological identity. *Frontiers in immunology* 5(April), 153.

Hasselman, F. (2015). Classifying acoustic signals into phoneme categories: average and dyslexic readers make use of complex dynamical patterns and multifractal scaling properties of the speech signal. *PeerJ* 3(3), e837. doi: 10.7717/peerj.837.

Hayflick, L., and Moorhead, P.S. (1961). The serial cultivation of human diploid cell strains. *Exp Cell Res* 25(3), 585-621. doi: 10.1016/0014-4827(61)90192-6.

Heras-Escribano, M. (2019). Pragmatism, enactivism, and ecological psychology: towards a unified approach to post-cognitivism. *Synthese*. doi: 10.1007/s11229-019-02111-1.

Hopfield, J.J. (1994). Physics, Computation, and Why Biology Looks So Different. *Journal of Theoretical Biology* 171(1), 53-60. doi: DOI 10.1006/jtbi.1994.1211.

Izvekov, S., and Voth, G.A. (2005). A multiscale coarse-graining method for biomolecular systems. *J Phys Chem B* 109(7), 2469-2473. doi: 10.1021/jp044629q.

Jacobs, D.M., and Michaels, C.F. (2007). Direct learning. *Ecological Psychology* 19(4), 321-349. doi: Doi 10.1080/10407410701432337.

Kak, S.C. (1996). Speed of computation and simulation. *Foundations of Physics* 26(10), 1375-1386. doi: Doi 10.1007/Bf02058274.

Kelso, J. (1995). *Dynamic patterns: The self-organization of brain and behavior*. The MIT Press.

Kelso, J.A.S., and Engstr\o m, D.A. (2006). *The complementary nature.* The complementary nature. xviii, 317 pp. Cambridge, MA: The MIT Press; US.

Kolchinsky, A., and Corominas-Murtra, B. (2019). Decomposing information into copying versus transformation. *Journal of The Royal Society Interface* 17(162). doi: 10.1098/rsif.2019.0623.

Kolchinsky, A., and Wolpert, D.H. (2018). Semantic information, autonomous agency and non-equilibrium statistical physics. *Interface Focus* 8(6), 20180041. doi: 10.1098/rsfs.2018.0041.

Kondepudi, D., Kay, B., and Dixon, J. (2015). End-directed evolution and the emergence of energy-seeking behavior in a complex system. *Phys Rev E Stat Nonlin Soft Matter Phys* 91(5), 050902. doi: 10.1103/PhysRevE.91.050902.

Kugler, P.N., and Shaw, R. (1990). Symmetry and Symmetry-Breaking in Thermodynamic and Epistemic Engines: A Coupling of first and Second Laws. *Synergetics of cognition* 45.

Kugler, P.N., Shaw, R.E., Vincente, K.J., and Kinsella-Shaw, J. (1990). Inquiry into intentional systems I: Issues in ecological physics. *Psychological Research* 52(2), 98--121.

Küppers, B.-O. (2013). "Elements of a Semantic Code," in *Evolution of Semantic Systems*.), 67-85.

Landauer, R. (1991). Information Is Physical. *Physics Today* 44(5), 23-29. doi: Doi 10.1063/1.881299.

Lanting, T., Przybysz, A.J., Smirnov, A.Y., Spedalieri, F.M., Amin, M.H., Berkley, A.J., et al. (2014). Entanglement in a Quantum Annealing Processor. *Physical Review X* 4(2), 021041. doi: ARTN 021041

10.1103/PhysRevX.4.021041.

Lee, D.N. (1998). Guiding movement by coupling taus. *Ecological Psychology* 10(3-4), 221-250. doi: DOI 10.1207/s15326969eco103&4_4.

Levitin, L. (1992a). Physical Information Theory Part I. Quasiclassical Systems. *Physics and Computation, 1992. PhysComp'92., \ldots*.

Levitin, L. (1992b). Physical information theory part II: Quantum systems. *Physics and Computation, 1992. PhysComp'92., \ldots* (4).

Lintern, G., and Kugler, P.N. (1991). Self-organization in conncetionist models: Associative memory, dissipative structures and thermodynamic law. *Human movement science* 10, 447--483.

Lloyd, S. (2006). *Programming the universe: a quantum computer scientist takes on the cosmos.* Random House LLC.

Maclennan, B.J. (2011). Bodies---both informed and transformed embodied computation and information processing. *Information and Computation*, 225--253.

MacLennan, B.J. (2012). Embodied Computation: Applying the Physics of Computation to Artificial Morphogenesis. *Parallel Processing Letters* 22(3). doi: Artn 1240013

10.1142/S0129626412400130.

Magdziarz, M., and Weron, A. (2011). Anomalous diffusion: testing ergodicity breaking in experimental data. *Phys Rev E Stat Nonlin Soft Matter Phys* 84(5 Pt 1), 051138. doi: 10.1103/PhysRevE.84.051138.

Mahulikar, S.P., and Herwig, H. (2009). Exact thermodynamic principles for dynamic order existence and evolution in chaos. *Chaos Solitons & Fractals* 41(4), 1939-1948. doi: 10.1016/j.chaos.2008.07.051.

Marmelat, V., and Delignieres, D. (2012). Strong anticipation: complexity matching in interpersonal coordination. *Exp Brain Res* 222(1-2), 137-148. doi: 10.1007/s00221-012-3202-9.

Marr, D., and Poggio, T. (1976). From understanding computation to understanding neural circuitry. *Massachusetts Institute of Technology, Artificial Intelligence Laboratory* (AI Memo 357).

Metzler, R., Jeon, J.H., Cherstvy, A.G., and Barkai, E. (2014). Anomalous diffusion models and their properties: non-stationarity, non-ergodicity, and ageing at the centenary of single particle tracking. *Phys Chem Chem Phys* 16(44)**,** 24128-24164. doi: 10.1039/c4cp03465a.

Michaels, C., and Carello, C. (1981). *Direct perception.*

Mozaffarilegha, M., and Movahed, S.M.S. (2019). Long-range temporal correlation in Auditory Brainstem Responses to Spoken Syllable/da. *Sci Rep* 9(1)**,** 1751. doi: 10.1038/s41598-018-38215-w.

Newen, A., De Bruin, L., and Gallagher, S. (2018). *The Oxford handbook of 4E cognition.* Oxford University Press.

Noble, D. (2010). Biophysics and systems biology. *Philos Trans A Math Phys Eng Sci* 368(1914)**,** 1125-1139. doi: 10.1098/rsta.2009.0245.

Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference.* Elsevier Science.

Petrusz, S.C. (2008). What Does ``Lawful" Have to Mean for Self-Organizing Systems to Qualify as Such? *Ecological Psychology* 20(3)**,** 37--41.

Petrusz, S.C., and Turvey, M.T. (2010). On the Distinctive Features of Ecological Laws. *Ecological Psychology* 22(1)**,** 44--68.

Pfeifer, R., Lungarella, M., and Iida, F. (2007). Self-organization, embodiment, and biologically inspired robotics. *Science* 318(5853)**,** 1088-1093. doi: 10.1126/science.1145803.

Polani, D., Sporns, O., and Lungarella, M. (2007). "How information and embodiment shape intelligent information processing," in *50 Years of AI*, ed. M. Lungarella. (Berlin Heidelberg: Springer Verlag), 99--111.

Potter, J. (2000). Post-cognitive psychology. *Theory & Psychology* 10(1)**,** 31-37. doi: Doi 10.1177/0959354300010001596.

Prigogine, I., and Nicolis, G. (1977). *Self-Organization in Non-Equilibrium Systems.* Wiley.

Prigogine, I., and Stengers, I. (1984). *Order out of Chaos: Man's new dialogue with nature.* Flamingo.

Ratcliff, R., and Rouder, J.N. (1998). Modeling response times for two-choice decisions. *Psychological Science* 9(5)**,** 347-356. doi: Doi 10.1111/1467-9280.00067.

Rome, E. (2008). "MACS - Final Activity Report", in: *Methodology.* EU).

Şahin, E., \cCakmak, M., Doğar, M.R., Uğur, E., and \"U\ccoluk, G. (2007). To Afford or Not to Afford: A New Formalization of Affordances Toward Affordance-Based Robot Control. *Adaptive Behavior* 15(4)**,** 447--472. doi: 10.1177/1059712307084689.

Scheffer, M., Bolhuis, J.E., Borsboom, D., Buchman, T.G., Gijzel, S.M.W., Goulson, D., et al. (2018). Quantifying resilience of humans and other animals. *Proc Natl Acad Sci U S A* 115(47)**,** 11883-11890. doi: 10.1073/pnas.1810630115.

Schrödinger, E. (1944). *What is life?* : Cambridge University Press.

Schweitzer, F., Ebeling, W., and Tilch, B. (2001). Statistical mechanics of canonical-dissipative systems and applications to swarm dynamics. *Phys Rev E Stat Nonlin Soft Matter Phys* 64(2 Pt 1)**,** 021110. doi: 10.1103/PhysRevE.64.021110.

Searle, J.R. (1990). "Cognitive science and the computer metaphor," in *Artifical Intelligence, Culture and Language: On Education and Work*. Springer), 23-34.

Shannon, C.E. (1948). A Mathematical Theory of Communication. *Bell System Technical Journal* 27(3), 379--423.

Shaw, R., Flascher, O., and Mace, W. (1996). "Dimensions of Event Perception," in *Handbook of perception and action*. (New York: Academic Press), 345--295.

Shaw, R., and Kinsellashaw, J. (1988). Ecological Mechanics - a Physical Geometry for Intentional Constraints. *Human Movement Science* 7(2-4), 155-200. doi: Doi 10.1016/0167-9457(88)90011-5.

Shaw, R.E., Kadar, E., and Kinsella-Shaw, J. (1994). Modelling systems with intentional dynamics: A lesson from quantum mechanics. *Origins: Brain and self-organization*, 53--101.

Shaw, R.E., and Turvey, M.T. (1999). Ecological foundations of cognition: II. Degrees of freedom and conserved quantities in animal-environment systems. *Journal of Consciousness Studies* 6(11-12), 111--123.

Shay, J.W., and Wright, W.E. (2000). Hayflick, his limit, and cellular ageing. *Nat Rev Mol Cell Biol* 1(1), 72-76. doi: 10.1038/35036093.

Shiels, P.G., Kind, A.J., Campbell, K.H., Waddington, D., Wilmut, I., Colman, A., et al. (1999). Analysis of telomere lengths in cloned sheep. *Nature* 399(6734), 316-317. doi: 10.1038/20580.

Stephen, D.G., Stepp, N., Dixon, J.A., and Turvey, M.T. (2008). Strong anticipation: Sensitivity to long-range correlations in synchronization behavior. *Physica a-Statistical Mechanics and Its Applications* 387(21), 5271-5278. doi: 10.1016/j.physa.2008.05.015.

Stepp, N., Chemero, A., and Turvey, M.T. (2011). Philosophy for the rest of cognitive science. *Top Cogn Sci* 3(2), 425-437. doi: 10.1111/j.1756-8765.2011.01143.x.

Stepp, N., and Turvey, M.T. (2010). On Strong Anticipation. *Cogn Syst Res* 11(2), 148-164. doi: 10.1016/j.cogsys.2009.03.003.

Susskind, L. (1995). The World as a Hologram. *Journal of Mathematical Physics* 36(11), 6377-6396. doi: Doi 10.1063/1.531249.

Swenson, R., and Turvey, M.T. (1991). Thermodynamic Reasons for Perception--Action Cycles. *Ecological Psychology* 3(4), 317--348.

te Molder, H. (2016). "What happened to post-cognitive psychology?," in *Discursive psychology: classic and contemporary issues*, eds. C. Tileaga & E. Stokoe. (Abingdon, UK: Routledge), 87-100.

Turvey, M.T., and Carello, C. (1981). Cognition: the view from ecological realism. *Cognition* 10(1-3), 313-321. doi: 10.1016/0010-0277(81)90063-9.

Turvey, M.T., and Carello, C. (2012). On Intelligence From First Principles: Guidelines for Inquiry Into the Hypothesis of Physical Intelligence (PI). *Ecological Psychology* 24(1), 3-32. doi: 10.1080/10407413.2012.645757.

Turvey, M.T., and Shaw, R.E. (1999). Ecological foundations of cognition: I. Symmetry and specificity of animal-environment systems. *Journal of Consciousness Studies* 6(11-12), 95--110.

Van Geert, P.L.C., and Fischer, K.W. (2009). "Dynamic systems and the quest for individual-based models of change and development," in *Toward a New Grand Theory of Development? Connectionism and Dynamic Systems Theory Reconsidered*, eds. J.P. Spencer, M.S.C. Thomas & J. McClelland. Oxford University Press), 313--336.

van Gelder, T. (1995). What Might Cognition Be, If Not Computation? *The Journal of Philosophy* 92(7), 345--381.

Varela, F.J., Rosch, E., and Thompson, E. (1992). *The Embodied Mind: Cognitive Science and Human Experience.* MIT Press.

Verlinde, E. (2011). On the origin of gravity and the laws of Newton. *Journal of High Energy Physics* 2011(4), 1--27. doi: Artn 029

10.1007/Jhep04(2011)029.

Villalobos, M., and Dewhurst, J. (2017). Why post-cognitivism does not (necessarily) entail anti-computationalism. *Adaptive Behavior* 25(3), 117-128. doi: 10.1177/1059712317710496.

von Foerster, H. (2003). *Understanding Understanding.* New York, NY: Springer New York.

Walker, I. (1972). Biological memory. *Acta Biotheor* 21(3), 203-235. doi: 10.1007/bf01557179.

Walker, I. (1983). Complex-irreversibility and evolution. *Experientia* 39(8), 806-813. doi: 10.1007/bf01990396.

Ward, R.M., and Kelty-Stephen, D.G. (2018). Bringing the Nonlinearity of the Movement System to Gestural Theories of Language Use: Multifractal Structure of Spoken English Supports the Compensation for Coarticulation in Human Speech Perception. *Front Physiol* 9, 1152. doi: 10.3389/fphys.2018.01152.

Weng, J., McClelland, J., Pentland, A., Sporns, O., Stockman, I., Sur, M., et al. (2001). Artificial intelligence. Autonomous mental development by robots and animals. *Science* 291(5504), 599-600. doi: 10.1126/science.291.5504.599.

Wheeler, J.A. (1989). "INFORMATION, PHYSICS, QUANTUM: THE SEARCH FOR LINKS", in: *3rd International Symposium Foundations of Quantum Mechanics, Tokyo.*).

Wolfram, S. (2002). *A new kind of science.* Wolfram Media.

Wolpert, D. (1992). Memory systems, computation, and the second law of thermodynamics. *International Journal of Theoretical Physics* 31(4), 743--785.