



MASTERARBEIT | MASTER'S THESIS

Titel | Title

The SOS of Experience: A Dialectic Approach to the Relationship
Between Sense of Self and Experience

verfasst von | submitted by
Ekin Deniz Dere BSc

angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of
Master of Science (MSc)

Wien | Vienna, 2025

Studienkennzahl lt. Studienblatt | Degree
programme code as it appears on the
student record sheet:

UA 066 013

Studienrichtung lt. Studienblatt | Degree
programme as it appears on the student
record sheet:

Masterstudium Joint-Degree Middle European
interdisc. master prog. in Cognitive Science

Betreut von | Supervisor:

Sebastjan Vörös PhD

ABSTRACT

The central aim of this thesis is to explore the relationship between experience and the sense of bodily self by focusing on rudimentary forms of living systems. Drawing on the work of the Chilean biologist and cognitive scientist Francisco Varela, our primary claim is that the sense of bodily self — the ability to distinguish the effects of one's actions from that of other agents and of agentless changes in the environment — emerges dialectically with experience in organisms with nervous systems. We further argue that at the most rudimentary level, experience manifests as directed activity: experimentation with the environment that goes beyond mere reactive movement.

KURZFASSUNG

Das zentrale Ziel dieser Masterarbeit ist es, die Beziehung zwischen Erfahrung und dem Empfinden des leiblichen Selbst zu untersuchen, indem wir uns auf grundlegende Formen lebender Systeme konzentrieren. Basierend auf den Werken des chilenischen Biologen und Kognitionswissenschaftlers Francisco Varela vertreten wir die Hauptthese, dass das Empfinden des leiblichen Selbst – die Fähigkeit, die Auswirkungen der eigenen Handlungen von denen anderer Akteure und von akteurslosen Veränderungen in der Umwelt zu unterscheiden – in Organismen mit Nervensystemen in einer dialektischen Beziehung zur Erfahrung steht. Darüber hinaus argumentieren wir, dass Erfahrung auf der grundlegendsten Ebene als gerichtete Aktivität auftritt: als Experimentieren mit der Umwelt, das über bloße reaktive Bewegungen hinausgeht.

TABLE OF CONTENTS

TABLE OF CONTENTS	3
1. INTRODUCTION	4
2. SETTING THE STAGE	8
3. A MATTER OF LIFE AND DEATH	19
3.1 The Skeletal Bio-Logic	19
3.2 From Autopoiesis To Autonomy	36
3.2.1 Constraints and Sense of Bodily Self in Multicellular Organisms	37
3.2.2 Autonomy and Control	45
3.2.3 From the Cell to the Immune and Nervous Systems	51
3.3 Living Systems and Change	54
4. THE CONCRETE AND THE VIRTUAL	65
4.1 Significance and Coupling	67
4.2 Relentless Action And Cognitive Dimension	72
4.3 Organism And Its Selves	81
5. CONCLUSION	89
REFERENCES	95

1. INTRODUCTION

“Our current notions about evolution and brain will be as distant to our grandchildren as this animistic cosmology is to as today.”
(Varela, 1987, p.49)

“Since we are free to make our own syntactic rules we are also free to interpret inherently simple events as messages in our own elaborate symbolic systems, and indeed this is what has generated all mythologies and probably several sciences.”
(Pattee, 1977, p.262)

What makes cognitive science a “collaborative effort” (Friedenberg and Silverman, 2016, p.35), a *discipline* rather than a mere *conglomeration of disciplines*? The largely unchallenged answer is that the different disciplines making up cognitive science share common concepts and methodologies, they are “glued” together by the “the topic of mind and, for the most part, the use of scientific methods” (ibid). Justly, the discipline is often called the “science of mind”. However, this unifying thread faces a significant obstacle: a quick glance at the contents of the cognitive science textbook from which we quoted reveals the uncertain, if not precarious, foundations of this relatively nascent science. Some selected chapter titles read: “Mind as Emergent Property”, “Mind as Information Processor”, “Mind as Brain”, “Mind as Emotion”, “Mind as Society”.

Mind, experience, and self—topics central to cognitive science—are often approached through what we may call “as” frameworks (e.g., mind as information processor, mind as brain), i.e., frameworks that focus on specific aspects of a given phenomenon while leaving their basic underpinnings underexplored. It is therefore our contention that cognitive science,

perhaps more so than any other science in our current times, could benefit from a clarification of its central concepts, which in turn would require that we bring into explicit focus those topics that are usually not addressed explicitly. In other words, we maintain that

“[n]o evidence about the relation between the objective and subjective realms can be provided if the initial explanandum itself (that which has to be explained), has been banished as a valid object of study!” (Rudrauf et al., 2003, p.29)

It is with this motivation that we wish to begin our inquiry: By focusing on the explanandum, on the relationship between the objective and the subjective — *the self* (we will see in the next chapter why the self is situated at the intersection between the object and the subject).

The aim of this thesis is thus to provide a conceptualization of *experience* in its co-defining relationship with what we term the “sense of bodily self (SOS)”. We pursue these terms by returning to their basic, elementary forms or manifestations, to their “roots.” In general, we could say that the thesis strives to unpack the following statement: *The sense of bodily self (SOS)—the ability to differentiate the effects of one’s own actions from that of others—emerges in mutual definition (dialectical relationship) with experience in organisms with nervous systems.*

In doing so, we will not proceed in a vacuum, but by standing on the shoulders of giants. Our giant is the Chilean biologist, cognitive scientist, and philosopher Francisco Varela (1946-2001), known for his work on biological autonomy, neuro-phenomenology and enactive cognitive science. We will, however, not be able to cover the entirety of Varela's research interests in this text. When we were deciding on the scope of the thesis, we were presented with two alternatives: either to concentrate on pre-reflective, pre-linguistic organisms; or to concentrate on the other end, i.e., human beings situated in the socio-linguistic dimension. We opted for the option that seemed more radical in the etymological sense of the

word—relating to, or proceeding from, the root—and thus chose the path that places the emphasis on *the rudimentary forms of life*. The reason for our choice to trace our phenomena of interest to their most elementary forms was to provide a solid ground that would allow us to make sense of them in their more complex manifestations. This, we believe, will allow us to highlight both the continuity of bio-logical processes as well as the observation that with each new level of analysis (e.g., that of neuro-logical), we encounter different modes in which a common logic presents itself.

The rest of this thesis is organized as follows:

Chapter 2 provides a brief outline of two separable, but not separate, modes of inquiry within cognitive science, framed as different approaches to the study of (living) systems. It provides reasons for adopting one approach rather than the other, and presents the rationale for the centrality of the self in the interdisciplinary context of the “science of mind”. Finally, it qualifies the idea of situating the notion of self in a broader context that cuts across these two approaches.

In Chapter 3, we will address what it means to conceptualize living systems holistically, followed by a characterization of the minimal living system (Chapter 3.1). We will then discuss how this characterization connects with the concept of autonomy (Chapter 3.2) and explore how living systems undergo change (Chapter 3.3).

Chapter 4, will serve as a bridge between our discussion of living systems and the notions of identity and self (Chapter 4.1). We will explore how various levels of living systems generate distinct cognitive domains (Chapter 4.2) and consider how an organism comprises multiple selves within the Varelian framework we adopt (Chapter 4.3).

By the time we reach the conclusion (Chapter 5), we will have established a pathway that began with, and culminates in, an understanding of the minimal form of the sense of bodily self (and how it relates to experience). This foundation is expected to reveal the mutual definition of self and non-self (including the environment and other living systems), affording us the means with which to explore higher levels of selfhood where subjectivity manifests itself in more complex ways, particularly with the emergence of language.

We acknowledge that we have set out to cover considerable ground, with a fair number of “weasel words” (à la von Hayek) and (Augustinian) paradoxes¹ at its heart. For this reason, we will, in the next chapter, provide a programmatic and quasi-historical account of the perspective we adopt throughout this thesis, which serves as both the overarching and underlying narrative of this work. This will set the stage for our inquiry into the nature of living systems and their identity.

¹“If no one ask of me, I know; if I wish to explain to him who asks, I know not.” (von Foerster, 2013, p.130)

2. SETTING THE STAGE

“If the heroic and unhappy history of cybernetics has anything to teach us, surely it is that, in addition to a bold willingness to go where no one has gone before, modesty, reasoned doubt, and mindfulness of tradition, nourished by critical reflection, are equally indispensable virtues in the quest for knowledge.”
(Dupuy, 2009, p. 161)

“If I am interested in doing anything at this point, it is in creating a form of culture, knowledge, religion, or politics that does not view itself as replacing another, in any sense, but one that can contain itself in a way of undoing itself. If we are not here to do that, I quite frankly would rather go skiing” (Varela, 1979b, p.19)

As already mentioned, this chapter is intended to provide the motivation for our subject of investigation as well as the choice of methods by drawing on some preliminary historical remarks about the tradition from which cognitive science developed, i.e., cybernetics². *Cybernetics*, whose foundations were laid in the 1940s, was conceived as a transdisciplinary study of “a set of problems centering about communication, control and statistical mechanics, whether in the machine or in the living tissue” (Wiener, 1948, p.11). Within this tradition, following Francisco Varela and the French physicist, philosopher, and cognitive scientist

²Wiener and Rosenblueth coined the “neo-Greek expression [...] Cybernetics, which we form from the Greek [word for] steersman. In choosing the term, we wish to recognize that the first significant paper on feedback mechanisms is an article on governors, which was published by Clerk Maxwell in 1868 [...]” (Wiener, 1948, p.12).

Jean-Pierre Dupuy, we identify two main approaches to the study of living and cognitive (i.e., “minded”) systems that underlie the formation of what Dupuy calls the “metaphysical research programme” of cognitive science³.

In this context, Dupuy draws our attention to the implicit metaphysics of cognitive technologies, embedded in the work of cognitive scientists, pointing out how their metaphysical commitments shape the field’s research direction. He writes:

“National Science Foundation devoted to the subject, entitled “Converging Technologies for Improving Human Performance,” summarizes the credo of the movement in a sort of haiku: If the Cognitive Scientists can think it. The Nano people can build it, The Bio people can implement it, and The IT people can monitor and control it. Note that cognitive science plays the leading role in this division of labor, that of thinker—not an insignificant detail, for it shows that the metaphysics of NBIC convergence [of nanotechnology, biotechnology, information technology, and cognitive science] is embedded in the work of cognitive scientists.” (Dupuy, 2009, p.xii)

For Dupuy, this situation is not only a cause for concern, but also a call to remember the origins of our field⁴. The cause for concern, writes Dupuy, has to do with the fact that, because of this metaphysical research programme inherent in technologies of cognitive science, the distinctions between life and non-life, as well that between man and machine, tends to get abolished: “[I]n the darkness of dreams, there is no difference between a living cat and a dead cat” (Dupuy, 2009, p.xiv). Humans are, as a result, placed “in the position of

³The term „research programme“ harkens back to the works of Karl Popper and Imre Lakatos. Dupuy writes: “on Karl Popper’s notion of a metaphysical research program, which is to say a set of presuppositions about the structure of the world that are neither testable nor empirically falsifiable, but without which no science would be possible” (Dupuy, 2009, p.x).

⁴For Dupuy (2009), cognitive science needs to learn from the mistakes of the cybernetics tradition that on the surface seems long gone, but underneath is still alive in the form of a general orientation or the dominant attitude, a framework of underlying assumptions and ambitions.

being the divine maker of the world, the demiurge, while at the same time [being condemned] to see themselves as out of date” (ibid). Alongside life and mind, the self, too, found itself mechanized and gave way to “a rigorous formulation of the notion of a subjectless process, whether it is a process of behavior or of thought” (Dupuy, 2009, p.112).

Dupuy traces this metaphysical research programme of cognitive sciences (with many nuances that we have to omit here) back to the so-called first-order cybernetics, inaugurated by, amongst others, Norbert Wiener, Warren McCulloch, Walter Pitts, John von Neumann: the tradition that, in his view, launched the grand project of the wholesale mechanization of man⁵. In contrast, the second order cybernetics, as developed by Heinz von Foerster, Ross Ashby and Gregory Bateson (all of whom have had a major influence on our main source in this text), had a different metaphysics guiding its research programme, a metaphysics that “had finally succeeded in raising itself to the level of self-consciousness, and so in becoming epistemology” (Dupuy, 2009, p.88). The difference between the two traditions - phrased in the form of a (semi)programmatic statement - reads: While the interest of first-order cybernetics was in “observed systems,” that of second-order cybernetics lay in “observing systems,” with an explicit focus on the role of the *observer* in describing systems.

As we will see more clearly in the following chapters, we are convinced that raising our inquiries to the level of epistemology is not only fruitful, but necessary for a science concerned with the mind, which in turn cannot be (especially in the field of cognitive science), done without turning the *subject* into an explicit focus of our studies. Or, if we phrase this in Varela’s words:

“[I]n the pursuit of cognitive science, that tradition of objectivity becomes a difficult stumbling block.” (Varela, 1979a, p.277)

⁵The tradition is often wrongly criticized as representing an anthropomorphization of the machine.

It is in this context that we take Dupuy's remarks as relevant: his warnings against a subject-less (cognitive) science, which amounts to a framework that isolates living from the lived, knower from the known, organism from the environment, without foreseeing a trail for their reunion. Now, although cognitive science presents itself as having brought back into the domain of science the ancient philosophical questions about the nature and structure of the human mind and its relations with the body, the other minds and the world, the (still) young science tends to overlook that its identity, that of a self-proclaimed science of mind, remains nevertheless philosophical (in the sense of being based on very specific philosophical assumptions that it tends to leave implicit):

“The science that speaks on behalf of the various disciplines that make up the field (mainly neuroscience, artificial intelligence, cognitive psychology, and linguistics), that provides these sciences d'esprit with the soul that ironically they would otherwise lack, is in reality none other than philosophy.” (Dupuy, 2009, p.90ff)

As we will see in more detail in the course of our introduction, there is an alternative – *subject-ful* – framework in cognitive science, one that manages to raise itself to become epistemology. Let us, at the risk of oversimplifying, trace this alternative approach back to the same parent: (first-order) cybernetics.

As Varela (1986a) sees it — or, in his own words, *dulls*⁶ it to make a point — there may be one occasion to split the cybernetics group in two perspectives: between Norbert Wiener representing the capacity of living systems to make sense and mechanisms of *autonomy*; and John von Neumann representing the establishment of decision rules and mechanisms of *control*⁷.

⁶When Stafford Beer asks why Varela has chosen only two names out of the many in the field, Varela replies, "se non è vero, è ben trovato" (even if it is not true, it is a very good fabrication). (Varela, 1986a)

⁷Note how the division of the world (into two) comes so easily to humans, as if it were a basic cognitive ability.

“It seems to me that it is quite clear - looking back from the 1980ies - that von Neumann actually prevailed. Cybernetics and systems theory developed most of their effort into the now familiar characterization of a machine (or an automaton) as an input-output device: you have inputs, transformations, and some kind of output, which, of course, can be made very precise with the notion of a Turing machine.”
(Varela, 1986a, p.117)

Why is it relevant for us that cognitive science inherited the cybernetics terminology, and thus perspective of, *control*? Because, as we already suggested above, we are convinced that raising our inquiries to the level of epistemology is a necessity in (cognitive) science. Also, because words come with their heritage, they carry their own (dynamic) context, bringing into the discussion—or unravelling within it—the perspective the speaker as well as the listener is taking. If we use a word in a discussion without (re)defining it, or qualifying what we mean by it (which we cannot always do, as it leads to infinite regress and would render language an impractical tool for communication), we assume that our interlocutor shares, in one way or another, a common tradition with us. In the context of cognitive science, this situation is salient with words like *information*, *processing*, and *representation*. Their usage, to invoke von Foerster, tells one about who the user, the explain-er is (von Foerster, 2013). In Varela's words:

“In discussing the future of the cognitive sciences and artificial intelligence, such terms are taken as granted, and any discussion about them assumes already some common background. In other words such discussions exists [*sic*] within a given tradition, they are historical and not mere 'objective' facts.” (Varela, 1986b, p.211)

Varela's point is that the cybernetics terminology inherited by cognitive science already brings with it a certain worldview, which is particularly suitable for engineering purposes. Adopting this terminology and the worldview that accompanies it in dealing with living

systems is at best incomplete and at worst antithetical. Speaking as a biologist, Varela argues that an over-proportionally strong preference for control perspective overlooks “the quality of living systems of having assertion of their internal coherences, their internal determination, as well as the fact that it is this internal determination [...] that shapes or imbues a world with meaning” (Varela, 1986a, p.117).

Neither meaning nor identity-assertion from within can be captured by the framework of control, or what Varela also calls the “input-process-output” characterization of a system. To characterize living systems, which are fundamentally – as we will explain in greater detail later – meaning-making systems, we need an emphasis on how the internal regularities of the system specify what is *of relevance* in the environment *for* the system. A Turing-type (input-output) system relates with its environment in a straightforward manner: “items of the environment become instructions that act on the structure of a system - that's why we call them inputs” (Varela, 1986a, p.119). A living system, on the other hand, interacts with its environment in a different fashion: through its self-determined behavior, the possibility of making or creating meaning and sense. Thus, we have “a laying down of a world, a laying down of a relevant 'Umwelt'” (ibid) (more on this later; see esp. Chapter 4.1).

For sense-making organisms, then, our mode of inference is from the inside, from the internal dynamics towards the outside, towards the world that is endowed with meaning *by the organism*. The “outside and thereby objectivity is immediately bracketed” (ibid). And this is what Varela means by objectivity being a stumbling block for cognitive science: it puts the emphasis on the outside, on the *information* to be processed by the system and thus misses the crucial point that it is the internal dynamics of the organism that decides what information is in the first place. For this reason, Varela (1979a) prefers the term *in-formation*, which underscores that something is formed from within, to *information*, which designates something coming from without.

At the center of this difference of looking at systems, we find the notion of the *self*. Wherever we, the observers, make use of this notion, there is an entity that forces us to oscillate between the autonomy and the control approach, the organism and the environment, the knower and the known, the subject and the object. We believe that especially within the field of cognitive science developing an understanding of the self is crucial because the field is inherently *self*-referential (minds studying minds, living beings studying living beings) and is situated at the center of the oscillation between the knower and known, subject and object.

It is with these considerations in mind that we attempt at bringing the *subject* to the fore, to see the organism as an active agent of concern that “creates a perspective which changes the world from a neutral place to an Umwelt that always means something in relation to the organism” (Weber and Varela, 2002, p.118).

Let us underscore: our strong focus on the subject does not mean it is our *sole* focus, it does not mean we are sailing towards *solipsism*. We see ourselves situated in the middle, i.e., in the intersection between the subject and the object. More accurately: we are seeing the subject and the object as co-specifying or co-defining each other, necessitating an oscillation in their respective conceptualizations. We will come back to this point later, but for now, let us make clear that:

“Our point is a plea for a dialectical middle way. We wish to walk the razor’s edge between the Scylla of representationism and the Charybdis of solipsism, by planting ourselves firmly in the middle.” (Varela, 1984a, p.217)

In what follows, we focus on the philosophical topics that are claimed to have been naturalized, demystified, and (seemingly) reclaimed by cognitive science from speculative philosophical discourse: life and self. To repeat: our aim is to conceptualize the relationship between (sense of bodily) self and experience. In attempting to account for the self, the

subject, we try to adopt a *subject-ful* stance, as opposed to what Dupuy (2009) criticizes as a stance that puts an overemphasis on subject-less processes. We do this by remembering “everything [that is] said is said by someone” (Maturana and Varela, 1987, p.26) and “[n]o content should be divorced from where this content has been produced.” (Varela, 1979b, p.13)

In our discourse, the primary focus is the organism; the principal fields of interest are biology, philosophy and cognitive science; and the principal source of reference is, as mentioned above, Francisco Varela. As such, this thesis is an interdisciplinary undertaking, in keeping with its situatedness in cognitive science. We believe no further justification on this point is necessary.

Now, the choice of focus on primary Varelian literature reflects this interdisciplinarity and is grounded in several reasons. First and foremost, in Varela, we discovered a lifelong interest on that which is the core of this work. As he puts it:

“I guess I’ve had only one question all my life. Why do emergent selves, virtual identities, pop up all over the place creating worlds, whether at the mind/body level, the cellular level, or the transorganism level?” (Brockman, 1996, p.233)

Another, not less important, reason is the alignment of our preferred approach: we wish to develop a theory of experience *rooted in biology* (a “naturalist” theory, so to say). In Varela’s words:

“Our approach will be mechanistic: No forces or principles will be adduced which are not found in the physical universe.” (Varela, 1979a, p.6).

In Chapter 3.1, we will explore how Varela understands the notion of *mechanistic*; above all, we will underscore how his approach contrasts with the Cartesian or “classical” conception to mechanical systems. At this juncture, let us merely note that the Cartesian view sees systems as collections or assortments of causally interrelated discrete parts, whereas Varela’s perspective, influenced by cybernetics and complex dynamic systems, conceives of mechanical systems as autonomous, i.e., self-organizing and self-producing processes, in which parts and the whole mutually specify each other. Thus, the term *mechanism*, when applied to living beings, refers to their autonomous constitution, meaning their classification as autonomous machines, focusing not on the causal interactions of individual components but on the whole they constitute through their interrelations (we will return to, and explain, these terms in the next chapter).

Before bringing these brief introductory remarks to a close, let us foreshadow what we will be doing in the upcoming chapters.

The rest of this thesis, as mentioned above, serves to unpack the following statement: The sense of bodily self—the ability to differentiate the effects of one’s own actions from that of others—emerges in mutual definition, i.e., dialectical relationship with experience, in organisms with nervous systems. We use, with Varela (1991), the definition of dialectics from Levins and Lewontin (1985).

“These are the properties of things that we call dialectical: that one thing cannot exist without the other, that one acquires its properties from its relation to the other, that the properties of both evolve as a consequence of their interpenetration.” (Levins and Lewontin, 1985, p.3)

To take an example: An organism *experiencing* a piece of wood, we argue, is an organism *experimenting*⁸ with it. For the organism to be able to experiment with the piece of wood, it needs to have a frame of reference based on which the experimentation takes place. This is to say: it needs to tacitly⁹ *know* (we handle knowing, knowledge and cognition in Chapter 4.2) that if it exerts pressure on the wood, the wood will move and this is not the same as the wood moving due to the wind or due to another organism exerting pressure on it.

The question here is: How is this organism supposed to have the frame of reference that it is the source of the wood moving without any prior experiment (or experience)? How is it supposed to have a sense of bodily self without experience? Our argument is that sense of bodily self and experience mutually define each other: an organism needs to be able to experience in order to develop the frame of reference which is the sense of bodily self, but it cannot experience without a sense of bodily self.

The development of the sense of bodily self, we argue, is encountered in multicellular organisms with a nervous system. Unlike unicellular organisms, whose activity is essentially re-active in the sense of being subordinated to their own (single-layered) identity maintenance, multicellular organisms with a nervous system exhibit directed activity. This directed activity owes its possibility to the constraints that the interdependent functioning of multiple autonomous systems¹⁰ within a coherent whole establish. Each subsystem not only maintains its own identity and coupling, but also contributes to maintaining the identity, coupling and adaptation of the organism as a whole. As a result, the organism operates

⁸The word experience comes the Latin *experientia*, i.e., *trial*, *proof*, experiment, which is derived from *experiens*, i.e., to try, to test. The Latin root of the word is interesting as it suggests that organisms experience through a stabilized history of activity, since (i) a trial, an experiment is rooted in the background of past activity, and (ii) oriented toward the horizon of future activity.

⁹We follow Polanyi in his conceptualization of tacit knowledge, which refers to the implicit, unarticulated knowledge that underpins our understanding and actions, encapsulated in his famous statement: "I shall reconsider human knowledge by starting from the fact that *we can know more than we can tell*" (Polanyi, 2009, p. 4, emphasis in the original). We extend his notion to incorporate non-linguistic animals and thus modify it as follows: the organism knows more than it is or can be aware of.

¹⁰These are: cellular, immunological, sensorimotor (Varela, 1991).

under dual constraints: each subsystem operates under its own (environmental and internal) while also operating within the overarching constraints of the whole.

“When [evolution of multicellulars] occurs, the [...] systems necessarily become subordinated [...] to the constraints [...] of the higher-order [...] unity which they, through their coupling, define topologically in the physical space.” (Varela, 1979a, p.52)

This multi-layered functioning of autonomous subsystems within the unified living system (i.e., the multicellular organism) is what makes directed activity possible, and it is this directed activity that forms the rudimentary level of experience. This raises two fundamental questions: what exactly constitutes a living system, this whole within which autonomous subsystems operate? And what exactly is an autonomous system? We address both questions in the next chapter, first by characterizing living systems in holistic terms, then by discussing what autonomy entails, and finally by explaining how living systems undergo change while preserving their identity.

3. A MATTER OF LIFE AND DEATH

“While a solipsist would argue that mountains and trees only exist in the mind, I know of no philosophy supporting the view that telephones and typewriters could have originated without the mind.” (Pattee, 1977, p.260)

“The fundamental fact of life—a fact of such universality that it may be used to define life—is that living systems contain their own descriptions.” (Pattee, 1977, p.259)

Having outlined our motivation and thematic arc, it has become clear that the first notion we need to address is that of life. In this chapter, then, we will direct our focus to living systems. We will proceed as follows: In Chapter 3.1, we will discuss the organization of living systems in holistic terms. In Chapter 3.2, we will explore how this type of organization connects to autonomous organization, and in Chapter 3.3, we will focus on the notion of change in autonomous systems.

3.1 The Skeletal Bio-Logic

What is common to all living things? They all reproduce, they all move and change and grow, they are all carbon-based, they have a metabolism, and so on. But: how long should this list that we are producing be and how can we know when (or if at all) the list will be completed? Further, is it not the case that everything we can put on the list implies a corresponding entity, i.e., an entity that can reproduce, grow, change, have a certain chemical composition and undergo internal chemical transformations? The question we should be asking, then, is, “how

can one characterize this entity?” and not, “which characteristics does this entity have?” In the words of Humberto Maturana:

“[We can] enumerate features of living systems such as reproduction, heredity, growth, irritability, and so on; but, how long a list was necessary?; when would the list be completed? In order to know when the list was completed I had to know what a living system was, which was, in fact, the question that I wanted to answer in the first place by producing such a list. I could speak about adaptation and evolution, about development and differentiation, and show how all these phenomena were tied together by the phenomenon of natural selection; but the question: 'What was the *invariant feature* of living systems around which natural selection operated?', remained unanswered.” (Maturana and Varela, 1980, p.xiii ff, my italics)

This means: In order to be able to characterize living systems, we are required to see them in holistic¹¹ or systemic terms. Taking this approach excludes the possibility of relying on an exhaustive list of its components for the characterization we are after. We need to be both more abstract and more concrete, to look for a commonality between all living beings, something that is more general than their capabilities or their component parts, but still able to account for their being concrete topological¹² entities in physical space (and in time). And furthermore, these abstract and concrete aspects need to be intertwined in the conduct of the entity. We seek a generality that contains and *encompasses*¹³, or *encircles* the particulars, the ones we know and the ones we might come to know in the future (e.g., in the search for artificial life, we might want to go beyond the reach of carbon-based characterization of life).

¹¹“[T]he key to understanding the holism of such systems, the whole-ness of systems, is to understand that they are organized, their parts are organized, in a circular form. That is, every part interacts with every other part.” (Varela, 1976a, p.26)

¹²Topological means: unchanging while undergoing transformations.

¹³Late Middle English term *encompassen*: *en*, i.e., to put in and *compass*, i.e., space or circumference.

In carrying out such a project, the “Chilean school” (of autopoiesis), as Dupuy (2009) names them, Maturana and Varela start from experience. They write:

“[W]e have to be aware that merely asking the question of how to recognize a living being indicates that we have an idea, even if implicitly, of its organization.” (Maturana & Varela 1987, p.42)

This amounts to saying, in the words of the German philosopher Hans Jonas, that we can recognize living systems because we are living systems:

“The observer of life must be prepared by life. In other words, organic existence with its own experience is required of himself for his being able to make that inference, which he does make all the time [...]. Thus we are prepared by what we are.” (Jonas, 2001, p.82)

When we see an organism, a living being, we do not see their ability to reproduce nor do we see their ability to grow and change; we do not see their being carbon-based or their metabolism. In fact, the only thing in that list of characteristics that we observe is their ability to move without being pushed or pulled by another agent, their *autonomous* conduct (we will return to this in the next section). Thus, what we see, when we observe a living being is its *organization* as a uniform whole.

Generally speaking, the term “organization” refers to the necessary relationships within an entity, or in other words, the relationships that must be present in order for something to belong to a particular category of systems. The term “structure”, in turn, refers to the components themselves and their relationships that constitute or realize a particular entity in physical space. In brief: structure is what actualizes the organization.

“[T]he organization of a machine is independent of the properties of its components, which are arbitrary, and a given machine can be realized in many different ways by many different kinds of components. In other words, although a given machine can be realized by many different structures, for it to constitute a concrete entity in a given space its actual components must be defined in that space, and have properties that allow them to generate the relations that define it.” (Varela, 1979a, p.9)

The “Chilean School” was working on the classification of living systems based on their organization as early as the 1970s, sharing the “same dissatisfaction with the dominant understanding of the living as molecular-genetic, as well as of the process of mind and cognition as information processing” (Weber, Varela, 2002, p.115). They introduced the term “autopoiesis¹⁴” in 1974¹⁵ to express “the mechanisms of self-production [that] are the key to understand both diversity and the uniqueness of the living” (Varela, 1981, p.14). The term, derived from Greek *auto*, i.e., self, and *poiesis*, i.e., creation, captures a fundamental characteristic of living systems: *self-production*.

According to the first definition of autopoiesis¹⁶, which has, in the course of time, sometimes been supplemented and sometimes consolidated, *autopoietic organization*:

“is defined as a unity by a network of productions of components which (i) participate recursively in the same network of productions of components which produced these

¹⁴The genesis of the term: “[O]ne day, while talking with a friend (Jose Bulnes) about an essay of his on Don Quixote de la Mancha, in which he analyzed Don Quixote's dilemma of whether to follow the path of arms (praxis, action) or the path of letters (poiesis, creation, production), and his eventual choice of the path of praxis deferring any attempt at poiesis, I understood for the first time the power of the word 'poiesis' and invented the word that we needed: autopoiesis. This was a word without a history, a word that could directly mean what takes place in the dynamics of the autonomy proper to living systems” (Maturana and Varela, 1980, p. xvii).

¹⁵It is in the title of a joint article: “Autopoiesis: The Organization of Living Systems, Its characterization and a Model” by Varela, Maturana and Uribe (1974).

¹⁶In this definition, neither the components nor their interactions are specified, and in this form, it is almost identical to the definition of autonomy proposed by Varela (1978, 1979a, 1981).

components, and (ii) realize the network of productions as a unity in the space in which the components exist.” (Varela, Maturana and Uribe, 1974, p. 188)

In 1974, the cell is discussed as an example with the components specified as molecules but autopoietic organization applied to the living systems in general. For Varela (1978, 1979a, 1981), however, autopoiesis becomes subordinated under the broader category of autonomy (as mentioned, we will explore autonomy in greater detail in the next section), and restricted exclusively to the minimal living system, the cell, with components being molecules and their interactions being chemical productions (Varela, 1978):

“Both in the early stages of the origin of life, and in their current terrestrial form, the basic feature of the cell as a unit is a circular concatenation of processes of chemical productions generating their own boundary in space.” (Varela, 1978, p.79)

The boundary in question is the *semi-permeable cell membrane* and the chemical processes leading to the constitution of this boundary (and thus bounded unit) together make up *the metabolism*. Thus, the cell, the minimal living system, is

“[a] network of chemical reactions which produce the molecules such that (i) through their interactions generate and participate recursively in the same network of reactions which produced them; and (ii) realize the cell as a material unity.” (Varela, Maturana and Uribe, 1974, p.188)

The cell’s capacity to produce the components that constitute itself, being the product of its own operation distinguishes it from the so-called “allopoietic¹⁷ systems”, those that are produced by the processes which are not a part of their organization.

¹⁷The word comes from Greek *allo-* meaning other and *-poiesis*, meaning creation .

“For example, although the ribosome itself is partially composed of components produced by ribosomes, as a unity it is produced by processes other than those which constitute its operation.” (Varela, Maturana and Uribe, 1974, p.189)

We will see below that this criterion of the boundary and thus the unity being determined by the internal operation becomes a criterion for autonomy, but the boundary loses the requirement of being topological in nature¹⁸ (Varela, 1981). For now, let us simply state in a cell, the boundary creates and limits the conditions for the metabolism. Within the boundary we have a circular concatenation of processes, or a network in which every production process is related to every other.

Crucially, the boundary, the cell membrane, does not precede the chemical processes for which it is a boundary. It is created by the chemical processes within the cell (within the unit encompassing the membrane and the metabolism) and it participates in those chemical processes; it is not separate from what it is bounding. Thus, we have the *fundamental circularity*¹⁹ of living systems:

“A cell stands out of a molecular soup by defining and specifying boundaries that set it apart from what it is not. However, this specification of boundaries is done through molecular productions made possible through the boundaries themselves. There is,

¹⁸Varela (1978), for example, claims that the nervous system and immune system are “informational” systems, and as such still “closed” but not “physically bounded”.

¹⁹Perhaps the best way to illustrate such circularity is to visualize it. Consider M.C. Escher’s famous engraving “Drawing Hands” from 1948, which depicts two hands emerging from a flat surface, each drawing the other into existence. Through their drawing, they create a closed structure by intercrossing or entangling two levels (Varela, 1984b) and this means that what they are drawing is not *each other*, as having two separate levels would have it, but *themselves*, as the collapse of two levels would have it. In this situation where the two levels are intercrossed or entangled, there is no first hand, there is no beginning of the drawing, there is no initial cause, no effect. There is only the level of mutual drawing, which mixes cause and effect (one level and another) together, so that the hands “mutually specify their conditions of production [and] bootstrap themselves out of the engraving to constitute a separate entity” (Varela, 1984b, p.310). By bootstrapping themselves out of the background, the hands create an *inside* and an *outside*. In other words, what sets the hands apart from the background, what creates an inside and an outside is their mutual specification: them drawing themselves. The act of drawing: an act that delineates a boundary, creates the possibility of a distinction, a differentiation in the otherwise unified background.

therefore, a mutual specification of chemical transformation and physical boundaries; the cell draws itself out of a homogeneous background. Should this process of self-production be interrupted, the cellular components no longer form a unity and gradually diffuse back into a homogeneous molecular soup.” (Varela, 1984b, p. 312)

In order for the cell to remain separate from the homogeneous molecular soup, the physico-chemical background, and remain a delineated unity, it is necessary for it to be, at the same time, related to that environment, that is, to be in constant material exchange with the environment within the range of *permissible transformations* determined by its organization. In other words, the autopoietic unit is *structurally coupled* to its environment, and through this coupling, it *keeps* its organization invariant.

“[O]ne way to spotlight the specificity of autopoiesis is to think of it self-referentially as that organization which maintains the very organization itself as an invariant.” (Varela, 1991, p.94)

This is important: *the cell realizes its organization continuously under permanent material turnover*. In other words, the physico-chemical constitution of the cell — its structure — is in a state of continuous change, and this is precisely what the word “metabolism” means: to change. It comes from the Greek *meta*, i.e., after or beyond, and *ballein*, i.e., to throw. The root of the word, we believe, is appropriate in at least three ways: 1. The organism selects and takes in matter from its environment, from what is *beyond* it, and *throws* it into its metabolic pathways. 2. This ingested matter is transformed by biochemical processes, thereby being *thrown beyond* its original state. 3. Finally, once transformed and utilized, it is *thrown* back to the environment, to the *beyond*. For any of this “throwing” and “beyonding” to take place, the cell needs to exist as a *dynamical* unity: Only by actively maintaining its organization can the cell undergo structural change, in fact, it is precisely for this

maintenance that the cell undergoes structural change. This is to say, the structural changes are subordinated to the *maintenance of organization*.

Now, the *keeping* of the organizational invariance is crucial as it is what makes the difference between a dynamical system and a static one. Unlike, for example, a crystal, which is naturally “frozen”, an autopoietic system is not freeze-able, it is necessarily *in continuous change*. But the crucial difference lies not in change itself but in the *subordination of change to organizational invariance*. One might object: but crystals, too, undergo structural change, they are not the same from one moment to another, they are not really frozen. To this possible objection, let us respond with elaborating what we mean by frozen.

We can say that the crystal changes its structure due to environmental factors, and if the environmental factors exert such strong pressure that the organization changes, then it ceases to be a crystal. But we are not saying that it *dies*. We can even argue on this basis that all structural changes in a crystal are subordinated to its organization, but we do not say that the structural changes are subordinated to *keeping* the organization invariant; “for this to happen a perspective *from* an actively constituted identity is essential” (Varela, 1997, p.79, my italics). What we are underscoring here, what we have hinted at in Chapter 2 and what we will expand on in Chapter 4 is precisely this *perspective*, the perspective that turns an *environment* into a *world* that makes the organism a locus of *concern*:

”The organic coupling and change must, according to its self-constitution, be always directed to maintain the process of self-realization. An autopoietic system is necessarily referred to itself: its actions consist in establishing the dynamical processes of staying alive. Stimuli from outside enter the sphere of relevance of such a unit only by their existential meaning for the keeping of the process of self-establishment. They acquire a valence which is dual at its basis: attraction or rejection, approach or escape.” (Weber and Varela, 2002, p.117)

This relating to the world through value or meaning - through attraction or rejection, approach or escape - is our natural, intrinsic way of making sense of the behaviour of living systems. In other words: Before any scientific or philosophical inquiry, we already understand living systems as meaning-making systems²⁰.

“If I observe a dog walking on the street and it suddenly changes direction and moves toward me, I generally impute an intention in the dog to greet me. [...] That is, the dog's behavior is difficult to account for unless I observe that the dog confronts its environment not as if it were receiving instructions from this environment for particular outcomes but, rather, as if these instructions were mere disturbances that the dog interprets and constructs according to its sense of regulation and balance.” (Varela, 1984b, p.311)

At this point, let us pause to summarize what we have discussed so far regarding the autopoietic organization: it is a dynamic process by which an organism maintains its identity—its "sameness"—amidst change. When we shift our focus from autopoiesis to autonomy in the next section, we will see that the *mechanism* of identity maintenance, the mechanism of wholeness at the heart of autopoiesis (as well as autonomy), is what Varela calls *operational closure*. In this section, we have introduced autopoiesis as the simplest case of autonomy, where its logic is most readily visible and comprehensible. Now that we are familiar with the skeletal bio-logic, and before we move on to autonomy, we want to use the insights from our previous discussions to address the question of "mechanism" that we have been alluding to from Chapter 2 onwards. This will provide an opportunity to synthesize

²⁰Remember here Varela's point about assuming background knowledge, as well as what Maturana and Varela emphasize in both 1980 and 1987: that even our ability to recognize a living thing when we see it points to its underlying organization. In short, while we seek a way to distinguish between the living and the non-living, we must realize that, in practice, we can distinguish between them quite effectively. This points towards a circularity, or rather a complementarity— between intuition and logic: “[T]he qualities of being can be thought of as perceived by intuition better than by discourse, and in a simultaneous whole, rather than in piecemeal approximation in time. [I]ntrospection says that it is out of this background of being that thought occurs, it is the archipelago of ideas in the giant sea of feeling and experience. Similarly I assumed this intuition is the earth from which discourse can be built, and to where it eventually returns” (Varela, 1976b, p.66ff).

the basic insights about living systems using a different set of terms, establishing a stable ground for exploring the more abstract tenets of living systems that we will encounter in the next section and in Chapter 3.

We encounter the term “mechanism” throughout the Varelian corpus, together with the common thread of drawing parallels between living beings and machines. It is important to emphasize that what Varela means by a mechanistic approach, as we see in familiarizing ourselves with the entirety of his thinking, is not a way of explaining phenomena by breaking them down into their basic components. It is a worldview based on this latter sense of mechanism that he criticizes as “a kind of worldview based on a purely mechanistic, observer-free science and philosophy” (Varela, 1979b, p.13). This dual use of the term mechanism — *positively*, as in Varela's own approach, and *negatively*, in his critique of reductionist science — together with our claim of alignment of preferred approaches, calls for a brief elucidation as to what we mean by “mechanistic” and “mechanism.”²¹

Mechanism comes from Latin *mechanismus* (Greek *mēkhanē*) and refers to a machine, instrument or device. A mechanism is then related to a machine, or rather how a machine works or is expected to work. It is, in a more abstract sense, *the means by which an end is to be reached*. This correlates with how we use it in science today: as a process or system that achieves a certain result. Embedded in the very word mechanism is the idea of what the result of this mechanism is supposed to be. This is not a trivial point that we will get back to.

Words morph in meaning and use over time, keeping pace with changing cultural and intellectual patterns, while often retaining traces of their origins. This is the case with the term

²¹This brief exercise is also necessary as we began this thesis with a critique of mechanisation.

mechanism, as well. Varela (1971) briefly describes how the term has changed over time, gradually but fundamentally²²:

“Descartes used the term exclusively to describe simple devices by means of which externally initiated actions could be transmitted (e.g., “levers,” “pulleys,” etc.), while at present “mechanism” refers in general to processes with an autonomous constitution (e.g., “self-assembling mechanism,” “intelligent machine,” etc.).” (Varela, 1971, p.426)

Mechanism, for Varela, is thus not on par with the classical (Cartesian) conception, which sees a system as a collection of individual, causally interrelated parts. His mechanism, influenced by the work done in cybernetics and dynamic complex systems, sees mechanical systems as self-organizing and self-producing (that is, as we will see in the next section, as autonomous). When applied to living beings, the term mechanism refers to their characteristic organization, i.e., their belonging to a certain class of machines.

Secondly, and following directly from what we have said, Varela's mechanical conception of living systems (in the specified sense we discussed above) differs from their classical mechanical conception. In his mechanical understanding, organisms are akin to von Foerster's non-trivial machines, i.e., as machines that change their inner condition—their constituents and interrelations—after every operation, i.e., self-referential machines:

“Now there are operators that are simply fixed and finished—and always do the same thing. Such an operator could be a multiplier by 2. You give it 4 and out comes 8. You give it 5, you get 10; you try with 1 and the operator responds with 2, and so forth. In this case you have what I call a trivial operator [...] But now I take an operator that, once it has carried out an operation, changes its own operation— we might also say,

²²This shift can be seen as one from a control approach to an autonomy approach, to use the later Varelian terminology or from digital computers to self-organizing systems, to use cognitive science terminology.

that changes its inner condition. After a multiplication by 2 it becomes a multiplier by 3. And if it was with 3, it goes on to multiplying by 5 or something similar.” (von Foerster, 2013, p.15 ff)

What is crucial for us is that, mechanistic or mechanical systems (or machines) are dynamical, they have a mechanism, i.e, they have processes that produce a certain behavior (effect or output).

Now, while we hold that living systems are machines, we note that they are not the same type of machine as man-made ones, which can also be self-referential (in the above sense) but have a predetermined goal imposed externally. In living systems, the goal to be achieved is self-determined and generated internally, from within, through their organization. To quote Hans Jonas's criticism of the first-order cyberneticians Rosenblueth, Wiener and Bigelow for adopting the kind of mechanistic approach that Varela criticises—or, in his own words, for submitting to the “almost irresistible tendency in the human mind to interpret human functions in terms of the artifacts that take their place, and artifacts in terms of the replaced human functions²³” (Jonas, 2001, p.110):

“In the case of the machine "missing the goal" means, of course, missing *our* goal, the goal for which it has been designed, namely by us, it "having" none itself; whereas

²³“The modern servomechanism is described as perceptive, responsive, adaptive, purposive, retentive, learning, decision-making, intelligent, and sometimes even emotional (but this last only if something goes wrong), and, correspondingly, men and human societies are being conceived of and explained as feedback mechanisms, communication systems, and computing machines [...] The classical mechanist, dealing with matter and motion, was content to speak of the "machine of the body, and, in method if not in metaphysics, to follow Descartes, originator of the notion of the "animal automaton," who had removed mind from the very terms of physical science. Later materialists repudiated the Cartesian dualism in name only. That mind is an epiphenomenon of material processes in the brain, as they held, remained a summary assertion not leading to such actual correlations and transformations as would invade the field of mind itself with the symbolism of physical science. Now [with cybernetics] is offered, for the first time, a mechanistic model which, it is claimed, applies to material and mental phenomena at once, not only equivalently but identically—that is, without involving passage from field to field.” (Jonas, 2001, p.110)

[a patient with a purpose tremor who fails to carry a glass of water to his mouth]²⁴ truly misses his goal, which is his not because he has been designed for it but because he has formed and entertained the design.” (Jonas, 2001, p.112, my italics)²⁵

In light of this, we say: The result of the mechanism of a living system is the continuous achievement of that very mechanism. More specifically, in terms of the dynamical processes of the cell: “the achievement of metabolism is metabolism itself” (Weber and Varela, 2002, p.113). Thus, we have here a dependency, a circularity between the mechanism and its result, or in the words of Weber and Varela (2002), an interrelation between means and goals:

“Because of this self-organizing circularity [...] all relations of cause and effect are also relations of means and purpose.” (Weber and Varela, 2002, p.106)

Let us now retrace our steps: chemical processes of production closed onto themselves creating a physically bounded unity, and through this self-creation put forward the conditions for its (no longer “their”) continuation, or better: established its domain of interactions, what it

²⁴Jonas uses the example from Rosenblueth, Wiener, and Bigelow (1943): “[I]f a cerebellar patient “is asked to carry a glass of water from a table to his mouth . . . the hand carrying the glass will execute a series of oscillatory motions of increasing amplitude as the glass approaches his mouth, so that the water will spill and the purpose will not be fulfilled.” The condition described is called “purpose tremor.”” (Jonas, 2001, p.111ff). For a more detailed discussion of the topic, refer to the Fifth Essay of Jonas’s *The Phenomenon of Life*.

²⁵Of course, it would be unreasonable to claim that Wiener and others were unaware of this claim, for the obviousness of which Jonas apologizes: “These elementary distinctions, for whose probably exasperating obviousness I apologize” (Jonas, 2001, p.112). Discussing the background of same work, the 1943 paper, in his famous book on cybernetics, Wiener writes: “Now, suppose that I pick up a lead pencil. To do this, I have to move certain muscles. However, for all of us but a few expert anatomists, we do not know what these muscles are; and even among the anatomists, there are few, if any, who can perform the act by a conscious willing in succession of the contraction of each muscle concerned. On the contrary, what we will is *to pick the pencil up*. Once we have determined on this, our motion proceeds in such a way that we may say roughly that the amount by which the pencil is not yet picked up is decreased at each stage. This part of the action is not in full consciousness” (Wiener, 1948, p.7, italics in the original). However, being aware of a rather obvious situation does not mean that this awareness is part of the research agenda or, to use our terminology from the introduction, the metaphysical research programme. We are reminded here of a passage by the late Wittgenstein, which Varela quotes on several occasions: “One is unable to notice something—because it is always before one’s eyes” (Wittgenstein, 1986, p.29).

can interact with without losing its stability, without its organization being damaged. Upon the formation of the unity, the processes that form it by closing on themselves function as part of a whole (their status is changed). Once we have the bounded unit, we have a re-ordering of components (in our case: processes): from floating freely into being system-parts. Once we have the unit, in the case of the cell, we have self-perpetuation (auto-induced-continuity). That is, the processes that created the unity now acquire a new dimension, that of maintaining the chain, the internal coherence, the stability.²⁶ And this is, in full, how the machine that is the cell works: its result is the continuous achievement of its very mechanism:

“[L]ife appears in its very means, that is, in its structural equipment for living, as its own achievement, or at least result, instead of being simply endowed with its means and faculties.” (Jonas, 2001, p.45)

This is the very point where we start talking about organisms being a center of concern (Weber and Varela, 2002). A self-created concern, a self-created identity “perpetually at risk of breakdown, of divergence; [...] “operating at the edge of chaos.” [...] intrinsically precarious and intrinsically fragile” (Rudrauf et al., 2003, p.38). This is precisely, to tie in with

²⁶In our accounting for wholes, the terms (i) closure, (ii) distinguishability, (iii) circularity, (iv) self-referentiality, (v) stability (coherence), (vi) reciprocity, (vii) identity, (viii) simultaneity come together. To understand systems as wholes means to understand how their parts are organized in a (iii) circular form, in which (vi, viii) every part interacts with every other part, so that we have a (v) stable (iv) self-referential system (Varela, 1976a). Describing such systems, i.e., organizationally (i) closed systems, coincides with describing a process of (vii) identity assertion, i.e., a system which is (ii) distinguished as belonging to a class of systems. Or in Varela's own terms: “The fact that wholes have this closed organization implies that in order to describe them we have to deal with self-referential descriptions [...] with functions that are functions of themselves, or interactions that interact with themselves, properties that compute themselves, and so on. [B]ut we were forbidden to deal with that in all its formal precision by the Russell-Whitehead Theory of Types, which requires us to completely break up any circularity by establishing a hierarchical form. This is the convention which is predominant, in classical science and therefore in system theory. It's hierarchical; you have input – system – output. You have the beginning and the end, and the beginning is the beginning and the end is the end; there is no confusion between the two. But if you are to really look at wholes, you can't do that. You have to look at the beginning and the end as a circle. So you cannot deal with them hierarchically, you have to look at them as self-referential situations” (Varela, 1976a, p. 27).

our earlier discussion, the difference between living matter, which keeps its identity, and inanimate matter, which lacks this quality.

“The material particle, identifiable in its space-time position, is simply what it is, immediately identical with itself without need to maintain that self-identity as an act of its existence [...] Organic identity, however, must be of a different nature altogether. In the precarious metabolic continuity of organic form, with its perpetual turnover of constituents, no inert substratum, no single "path" and no "bundle" of parallel paths of coteremporaneous [*sic*] members, is available as referent for external identity. [...] An identity which from moment to moment reasserts itself, achieves itself, and defies the equalizing forces of physical sameness all around, is truly pitted against the rest of things.” (Jonas, 2001, p.82ff)

What Jonas is pointing out in the above quote is that the organism must constantly strive to maintain its sameness—that is, its identity, its existence. This is in stark contrast to inanimate matter, which remains unchanged from one moment to the next, i.e., remains identical with itself. The organism doesn't simply *remain* what it is, but actively *achieves* what it is (Vörös, 2023)²⁷:

“[A] living being [...] can never coincide with itself, for if it did, it would no longer be living. [...] The organism as a becoming being [...] is both behind its boundary (separated from the environment, related to its dynamics) and outside of it (separated from its dynamics, related to the environment). It can, in other words, posit its boundary only by crossing it; and it is precisely in this act of (self-)crossing that it (re)posits it. Thus, [...] the theory of autopoiesis is not so much a theory of self-identity - and thus open to criticism for its perceived isolationism, stasis, etc. - as it is a theory

²⁷Our interpretation here is borrowed from the following talk: Vörös, S., & Halák, J. (2024, August 25). *Life as the advent of the problematic* [Video]. The Organic and the Normative conference. <https://www.youtube.com/watch?v=IVXt8K0rY8s>

of self-difference (Bortoft 2012: 71), a theory of how a living being can “become other without becoming an-other” (Ibid.) or, put more concisely, how it can self-differentiate without self-annihilating.” (Vörös, 2023, p.14)

Precisely because the organism is partially separated from the environment, in order to “remain”, it struggles against environmental pressures (i.e., “equalizing forces²⁸”) and for this reason it “can never coincide with itself”. There would be no need for self-maintenance and thus no concern if the system that emerged from the environment did not remain linked to it, i.e., if it completely severed ties with the environment and became self-enclosed (or in other words: if it differentiated itself *from* an environment *into* a vacuum). The link with the environment, which the organism does not sever as long as it is alive, is both a threat and a source²⁹. In other words, by the same act of self-differentiation from the environment, the organism creates the need to continue this act of differentiation, by the same stroke it creates the conditions for what counts as a threat, its own possible end by diffusion into the environment, but also its resources. Then, it becomes impossible to consider the organism in full in its own right because “the world for a living system is inseparable from its own living” (Varela, 1986b, p.213).

In other words: If the result of the mechanism of a living system is the continuous achievement of that very mechanism, and if this result is both made possible by the environment by virtue of the organism’s act of differentiation from it and is threatened as well

²⁸A short citation from Polanyi, we believe, makes Jonas’ expression more clear: “Inanimate nature is controlled by forces which draw matter toward stabler configurations. This is equally true in mechanics and thermodynamics and it applies also to open systems, like flames or flows.” (Polanyi, 2009, p. 89)

²⁹ We have considered that it may be more appropriate to describe this link as a *resource*. The word “resource” comes from the Latin *resurgere*, i.e., to rise again. The organism rises and rises again (continuously) as long as it maintains its identity, thus remains what it is. However, we have opted against this term because it immediately follows from its usage that when the organism dies, it rises again too but this time as another organism, and it is this loop that remains constant. In this interpretation, the loop becomes the context, while the organism - dying and rising again - becomes the structure or content. However, this aspect is beyond the scope of our present discussion.

as supported by the environment through the maintained link of the differentiated unity to it, environment becomes part of the system's definition³⁰.

As we will see in the next section, different autonomous systems exhibit different modes of coupling with their environment. While the cell and the macroscopic organism share a common underlying logic of organization, the manifestation of this logic is distinct in each case. In what follows, we will, in order to better explain this property of autonomous systems, first address the question of why autopoiesis remains a theory of the cell. Then, as our main argument for distinguishing between unicellular and multicellular organisms, we will introduce the concept of "freedom through constraint", drawing on the work of physicist Howard H. Pattee. Once these basic points have been established, we will provide a detailed account of autonomy and its mechanism, with a particular focus on operational closure. (We introduce Patteeian constraints before Varelian operational closure because we believe it clarifies the underlying idea common to both scientists more explicitly). Finally, we will conclude this rather lengthy section with a discussion of two other autonomous systems of interest: the immune and the nervous systems.

³⁰In the light of this discussion, we believe, we see more clearly one of the central issues of looking at living systems through the lens of input-output approach. The issue is in the question: Where does the *in* end and the *out* begin in this picture of mutual definition between the environment and the organism? Where do we find the distance between the organism and its environment? "There is no distance, not even the distance between an it and its picture, which makes it possible to ask how accurate a representation the picture is." (Varela, 1987 p.48). This is where the Escher illustration fails to account for the circularity of living systems. While it does a good job of underscoring how, through mutual drawing, the hands (the metabolism and membrane) separate themselves from the background and create a clearly delineated inside and outside; this works for the cell only in strictly topological terms. However, if we bring into the play the active linking of the organism with the environment, and note that its demarcation from the environment is not a full separation, not a division into a vacuum, and that when it comes to living systems, the line between in and out becomes rather fuzzy, the parallel we drew with Escher hands breaks down.

3.2 From Autopoiesis To Autonomy

Through a series of steps from single cells to multicellular organisms, one might contemplate expanding the concept of autopoiesis to encompass not only ourselves but also (insect) colonies or societies. However, Varela explicitly warns us against such a conflation of terms:

“[T]here have been some proposals suggesting that certain human systems, such as an institution, should be understood as autopoietic [...] I believe that these characterizations are category mistakes: they confuse autopoiesis with autonomy.”
(Varela, 1979a, p.55)

Even though institutions, systems or societies are beyond the scope of this thesis, the multicellular organism is not only well within its bounds, but is crucial for our central notions. The question here is whether it follows from the basic logic of autopoiesis that a multicellular organism is a second-order autopoietic system. The answer, according to Varela, is: no, “since its organization does not follow the same self-constructing principles” (Weber and Varela, 2002, p.115).

Although autopoiesis logically applies to all living organisms (since all living systems are made up of cells) we will treat it as a theory of the cell only. The reason: its logic gets fuzzy when extrapolated outside of the dimension of the minimal living system. For example, it would be a stretch (an unacceptable one) to claim that a baby differentiates itself from its environment purely through internal activity; this is where the logic of autopoiesis begins to waver. Moreover, the boundary of this self-created unity, if applied to multicellularity, remains ill-defined. If we consider multicellular organisms, including humans, is the boundary simply the skin? What metabolism creates and maintains this boundary? Can we even say that the skin, as a boundary, is actively maintained? More pointedly, how does this boundary enable

the very processes it is supposed to protect and maintain? For example, would a substantial disruption to the skin really lead to the collapse of the organism?

Thus, when it comes to multicellular organisms, the focus needs to shift from autopoiesis—the self-production of life—to *autonomy, self-regulation* of life:

“[M]ulticellular organism inherits its autonomous nature and sense-making qualities through the configuration of its neural identity.” (ibid)

But, what really happens when we shift from the single cell to a multicellular organism? Is there a form of discontinuity within the continuity of bio-logical complexity? If so, how can we characterize it?

3.2.1 Constraints and Sense of Bodily Self in Multicellular Organisms

Multicellular organisms with nervous systems, we argue, are unique in (at least) three aspects: (1) the self becomes stabilized through the constraints produced by the “meshing” of different autonomous subsystems (more on this in Chapter 4); (2) this meshing imposes additional constraints on the organism, in contrast to the single-layered cell; and (3) the organism exists in an external environment that offers inexhaustible possibilities, thereby both enabling and necessitating experimentation (we offer a more detail argument for this aspect in the closing paragraphs of this section)³¹. In contrast, a unicellular organism lacks this layered, interdependent structure, and its form of autonomy (i.e., autopoiesis) is therefore self-contained; it maintains itself not through the dynamics of interacting subsystems, which

³¹This is what distinguishes between the self associated with the immune system and the one associated with the sensori-motor body in space: the immune system exists within an intra-organismic environment and its environmental coupling is in terms of molecular interactions in this medium whereas the sensori-motor body couples with a macroscopic environment that is proper for cognitive behavior (Varela, 1991).

is the condition that produces constraints, but through the dynamics of its *own* self-creating and self-perpetuating processes. It is through the constraints created by the inter-meshing of different autonomous systems that the organism evolves to be capable of directed, experimental interactions with the environment rather than merely re-active movements. This directed activity, we argue, is the rudimentary form of experience.

As we move up layers of biological complexity, we have autonomous systems working together as another autonomous system and thereby creating new constraints that enable new properties and give rise to new degrees of freedom. This is crucial for our definition of the sense of bodily self: it is through the added constraints due to the meshing of autonomous systems within a coherent whole, that the movement of the self can get stabilized with repeated activity and acquires depth.

Now, what is a constraint? How can it limit a system and at the same time open up new possibilities for it?

“[C]onstraints can be understood as *relational properties* that the parts possess in virtue of their being integrated or unified (not aggregated) into a *systemic network*. [...]. The form, configuration, or topology of a system limits or prevents certain possible behaviors the parts could have on their own, while simultaneously opening up new possibilities for them in virtue of the states the system as a whole can access.” (Thompson, 2007, p.424, my italics)

In Chapter 3.1, we distinguished between autopoietic and allopoietic systems. We posited the cell as the minimal autopoietic system, and the ribosome serving as an example of an allopoietic system. Now, we can explain the concept of constraints using these examples, which should, by now, be quite clear to us. Within the context of the cell, we can think of the constraints as the relations between cellular processes (the parts) that realize the cell as a

unity (a whole) in space. In other words, in autopoietic organization, where every process is connected to every other, the constraints are precisely this total relational configuration of the whole that creates (itself as) a network. It is due to these constraints that the ribosome, an allopoietic system, is integrated into the processes that sustain the whole, the network. In other words, it is due to the *exhaustive component relations*—constraints arising through the cell's circular organization—that the ribosome becomes part of a chain, with its status subordinated to the continuation of the chain by virtue of the system's configuration. Consequently, the ribosome functions as a component, contributing to and realizing the states that are only attainable by the cell in its entirety.

Note that the cell is a living system, whereas (for our purposes and in our framework) the ribosome is not, and thus, the constraints that apply to these two systems are different in nature. For the physicist Howard H. Pattee³² (2012), there are two kinds of constraints: *structural* and (hierarchical) *control*³³. The fundamental difference between them, and the one that matters most for our purposes, is that the latter—which is unique to living matter—are the ones associated with new degrees of freedom and thus give rise to *biological hierarchy*.

“In the inanimate world, there are certain types of constraints that produce structures that we recognize as atoms, molecules and crystals, and eventually mountains,

³²In the context of complex systems theory, hierarchy theory, nonholonomic constraints, and modes of complementary explanations (symbolic and dynamical); Varela (1979a) draws on Pattee's ideas. While there are certain overlaps between the two scientists, their terminologies diverge substantially. Varela approaches life, autonomy, language (and epistemology) from a biological perspective, while Pattee approaches these same topics from a physicist's viewpoint. Although Varela cites Pattee on a few occasions, he remains largely unacknowledged in the latter's work. In Evan Thompson's (2007) book, which originated as a collaboration with Varela before Varela's passing and was completed by Thompson alone, credits Pattee with a number of points crucial to the enactive approach to complex systems, this is however, the only mention Pattee receives in the book: “Pattee emphasizes the complementarity of the linguistic and dynamical modes of description, but also suggests that symbolic information emerges from and acts as a constraint on dynamics. This idea is important for embodied dynamicism and the enactive approach” (Thompson, 2007, p.55).

³³Note the difference between Patteeian and Varelian terminology. For Varela, control systems or the control perspective are appropriate for engineering purposes, but in Patteeian terminology, control constraints are unique to living matter.

planets and stars, but these constraints are not associated with new freedom in the collections [...]” (Pattee, 2012, p.92)

For our purposes here, the notion of *hierarchy* refers to the autonomous systems interacting within the whole, or in Varela’s words to “the interdependence of different levels of system descriptions.” (Varela, 1979a, p.91). It is in the context of such a hierarchy of systems that we observe the property of “enabling through constraining,” with its manifestations ranging from the perspective of a bacterium (that we will encounter in Chapter 4) to the linguistic human cognition (which we omit in this thesis). A constraint can be seen, then, as that on which *emergent*³⁴ properties rest.

“[T]here is the apparent paradox that hierarchical controls both limit freedom and give more freedom at the same time. The constraints of the *genetic code* on ordinary chemistry make possible the diversity of living forms. At the next level, the additional constraints of *genetic repressors* make possible the integrated development of functional organs and multicellular individuals. At the highest levels of control we know that *legal constraints* are necessary to establish a free society, and constraints of spelling and syntax are prerequisites for free expression of thought.” (Pattee, 1973, p.73 ff, my italics)

Pattee’s (2012) hierarchy theory expands on the “boundary conditions” of Hungarian-British polymath Michael Polanyi, best known for his work on “tacit knowledge” (we refer to this type of knowing in our definition of the sense of bodily self) and his idea that *all knowing is*

³⁴Translating in Varelian terminology of closure: “[T]he kind of self-referential organization that has provided the stable properties [...] is what gives the system its nature. When you have a closed interaction of chemical productions, you can have a cell, and not before that. When you have a closed interaction of descriptions, you can have self-consciousness, and not before. When you have a closed interaction of species, you have an ecological system, and not before. That is, the closure, the self-referential-ness, seem to be the hinges upon which the *emergent properties* of a system turn.” (Varela, 1976a, p.27, my emphasis)

personal knowledge (we will cover knowing and personal knowledge, within the confines of our context in Chapter 4.2).

Polanyi's (1968, 2009) "boundary conditions" are conditions through which the higher-level systems "harness" the dynamics of lower levels. To illustrate his idea, Polanyi starts off with the case of a man-made machine: while its physical components obey the laws of physics and chemistry, their organization and function depend on higher-level design principles that are not determined by these laws (this relates to the second aspect of the mechanism we mentioned in the previous section, i.e., that classical (Cartesian) machines have an external goal). Similarly, the emergence of life and its higher-order functions cannot be fully explained by chemical reactions at the molecular level alone.

"If each higher level is to control the boundary conditions left open by the operations of the next lower level, this implies that these boundary conditions are in fact left open by the operations going on at the lower level. In other words, no level can gain control over its own boundary conditions and hence cannot bring into existence a higher level, the operations of which would consist in controlling these boundary conditions. Thus the logical structure of the hierarchy implies that *a higher level can come into existence only through a process not manifest in the lower level*, a process which thus qualifies as an emergence." (Polanyi, 2009, p.45, my italics)

In other words, Polanyi's *boundary conditions* are conditions that a given level leaves unspecified (to be controlled by the higher level), those that lie at the boundary of a system's charted territory, touching or intersecting it, without being integrable into its dynamics—conditions that are tangent to the system's boundaries³⁵. In this framework, the

³⁵In a (loose) sense, they are the conditions that the lower level *outsources* to the higher level. These conditions are marked but not harnessed without the higher level, not fully realized until the higher level *engages* with them. The way the higher level engages with the boundary conditions left unspecified by the lower levels by "controlling" them.

higher level in a hierarchy controls the boundary conditions of the level below it, but the reverse is not allowed: the lower level can neither create nor control the conditions required for a higher level since a higher level emerges through processes that cannot be fully accounted for by the operations of the lower levels.

We can apply these considerations to our case, and see that the *sense of bodily self* that emerges at the level of organisms with nervous systems cannot be fully accounted for by the single-layered perspectivity (or the self prior to its stabilization) of the single cell because it constrains or controls precisely that what the perspectivity (*lower level*) “enabled” but left unspecified.

Polanyi (1968) describes both machines and living organisms as operating under “dual control”³⁶: They obey and “harness” the lower level, i.e., physical laws while functioning within boundary conditions set by the higher level, i.e., their design (for the machine) or organization (for the living system). The “first level of emergence”³⁷, or the first level of “new freedom” that we are interested in, is created by the boundary condition “expressly left open by physics and chemistry [and] controlled by principles foreign to physics and chemistry” (Polanyi, 2009, p.41)³⁸.

“Obviously, the two critical events of evolution - first the appearance of life and later that of thought and language - led to phenomena that previously did not exist on the earth. To describe and to interpret these phenomena, new concepts, meaningless at

³⁶It is with the DNA that Polanyi has support for this dual-control: “What better instance of a boundary condition harnessing “lower” level processes could Polanyi have wished for? [...] DNA and its linear sequences of bases were thought to arise through a long evolutionary history of being subject to selective forces and, in that way at least, became a boundary condition not vulnerable to reduction by physics or chemistry” (Goldstein, 2012, p.142).

³⁷“The first emergence, by which life comes into existence, is the prototype of all subsequent stages of evolution, by which rising forms of life, with their higher principles, emerge into existence.” (Polanyi, 2009, p.49)

³⁸“The most common misconception of physics is that because laws are inexorable— that is, no event can disobey laws—the implication is that laws determine all events. That is not the case. Most of the structures in the universe are undetermined by laws or are accidental” (Pattee, 2013, p.15).

the previous level, are required. What can the notions of sexuality, of predator, or of pain represent in physics or chemistry?” (Jakob, 1977, p.1162)

We must add that, in addition to being left *undetermined*, these boundary conditions must be *determining*, in the sense of being and constraining to give rise to emergent properties. Between two extremes of determination and indeterminacy we find ourselves, once again, situated in the middle:

“From this physical point of view, life can be described as the result of an optimum level of constraint—a fortuitous balance of determined and undetermined behaviors that is available to matter only within a very narrow range of cosmic energies that allows local stable heritable structures (memory) to constrain state-determined (memoryless) dynamical laws. Life arises and evolves from those undetermined self-organizing constraints that (using Polanyi’s metaphor) propagate by harnessing the lawful ordered structures of their surrounding.” (Pattee, 2013, p.16)

To conclude this section, we will use the case of individuality to illustrate the implications of these technical considerations for the central focus of this thesis: the sense of bodily self. This will also provide an opportunity to comment on the experimental aspect of the sense of bodily self.

According to Polanyi (2009), life appears primarily in the form of *individuality*, which is rather weak at the vegetative level, i.e., in plants. With the emergence of *animal behaviour* (in our terminology, with the possibility and necessity of directed activity), individuality (in our terminology, the sense of bodily self) becomes more pronounced and peaks with the human level intelligence, culminating in *personality*:

“The vegetative system, sustaining life at rest, leaves open the possibilities of bodily movement by means of muscular action, and the principles of muscular action leave open their integration to innate patterns of behavior. Such patterns leave open, once more, their shaping by intelligence, the working of which offers, in its turn, wide-ranging possibilities for the exercise of still higher principles in those of us who possess them.” (Polanyi, 2009, p.41)

This increase in salience in individuality brings with it its own newly emerged threats as “every additional function with which the center [of the individual] is credited exposes it to new reproaches in respect of new failures” (Polanyi, 2009, p.50). It is this increase in the possibilities of failure, together with the opening up of the domain of sense of bodily self (the domain that lets the organism know — tacitly— that its activity leads to a certain change in the environment), that we can speak of the (possibility and) necessity of experimentation. This is why, in our framework, the sense of bodily self necessarily coincides with experience, and cannot be fully explained by taking into account the basic sensitivity of minimal life forms, but it is based on it, i.e., it “harnesses” its dynamics.

To sum up: In our framework, the mechanism through which we obtain the constraints that give rise to the stabilization of the self and thus emergence of the sense of bodily self is the mechanism of the inter-functioning of autonomous systems within a coherent whole (which is the organism). This requires us to shift our focus from the autopoietic system, i.e., the single cell, the minimal living system, to the autonomous system, i.e., the multicellular organism.

As we shift our focus to autonomy, it is important to note that this broader framework encompasses non-topologically bounded systems, such as the nervous and immune systems (discussed further in Section 3.2.3). This amounts to a “distinction between autopoiesis as proper to the unitary character of living organisms in the physical space, and autonomy as a general phenomenon applicable in other spaces of interactions” (Varela,

1981, p.14). This means that autopoiesis is a paradigm to think about autonomy, its simplest case, or the case in which the logic of autonomy is most easily grasped. In the next section, we will discuss *autonomy* in greater detail together with its counterpart, which is, as we touched upon in Chapter 2, *control*.

3.2.2 Autonomy and Control

The universal characteristic of autonomous systems is the so-called organizational or operational closure, i.e., “an indefinite recursion of component interaction” (ibid). This is what Varela (1978, 1979a, 1981) puts forward as the *Closure Thesis*, which states that every autonomous system is operationally closed. The notion of operational closure is borrowed from the mathematical context, and should not be interpreted to mean that the system is closed for interaction, rather it means that the system is characterized by the ongoing recursivity of organization. In more formal terms³⁹, the thesis states that we are dealing with situations of the form $F=\varphi(F)$, where F can be any process, and φ is the relationship between such processes, i.e., their form of interdependence (Varela, 1978).

“Such fixed-point rendering of organizational closure shows that, in a foundational sense, we are dealing with typically [*sic*] self-referring situations: F says something about itself, namely that $\varphi(F)$ is the case [...] [T]he notion of self-reference (circularity, indefinite recursion) is at the core of natural systems.” (Varela, 1978, p.81ff)

Note that the concept of constraints, which we discussed in the previous chapter, is reflected in how the relational properties of the whole constrain the degrees of freedom of individual components while allowing for the autonomy of the system (i.e., enable the states that can only be reached by the system as a whole). In other words, constraints are the rules that govern the relational interdependence of components, i.e. the recursive organization of an

³⁹This does not take into account the boundary requirement or aspect, as Varela (1978) also notes.

operationally closed system. Importantly, these rules are not imposed externally (as in the case of man-made machines) but arose internally, due to the systems self-creating and self-maintaining processes. Constraints and closure can be thought of as two sides of the same coin: while constraints emerge from closure, they also sustain it by structuring the rules for the interdependence that emerges from closure, which both limits and enables the system's processes.

Varela characterizes autonomy as the “assertion of the system’s identity through its internal functioning and self-regulation” (Varela, 1978, p.77), and adds that this assertion of identity is in a way that “observation proceeds through the coupling between the observer and the unit in the domain in which the unity’s operation occurs” (Varela, 1981, p.14). This is how the observer is brought into the definition of the system: the observer and the unity become coupled, and furthermore, this coupling occurs in the same domain as the unity’s operations occur. This means: there is always a criterion of distinction that the observer falls back to and the case that the system is autonomous, this criterion of distinction is linked with the organization of the system⁴⁰:

“We can interact with a recognized autonomous system because there is a criterion for distinction in some space. However, if such a distinction is, at closer inspection, not associated with the system’s operation, then either the unity is not operationally closed, or else the observer is describing it in a dimension that is not the one in which the organizational processes occur. Only when organization and distinction are linked, do we have an autonomous system, and this can only arise through organizational closure.” (Varela, 1981, p.16)

In nuce: autonomous systems are not mere conceptual constructions. There has to be a system that is different from the observer, an observed system that does not fuse with the

⁴⁰Note how this is akin to the organism establishing self-reference in the environment.

observing system, but stands independently so as to allow for the observer to enter into a mutually defining relationship with it. It has to exist *independently* so as to *depend* on the observer for its definition. We have here a loop, a circularity, a self-referentiality: the description of the system changes the system. Further, since the act of description changes the system, this change and thus the new state of the system calls for another description (infinite regress). Still further: since we are a defining component of the system, this also changes us (feedback, recursivity).

Now, it is for this reason that the autonomy alone cannot adequately account for the behavior of a system, or to use Varela's choice of words: its phenomenology⁴¹. Our conception linked the observer and the systems he is observing, the observer that cannot "step outside of the unity to consider its boundaries and environment simultaneously" (Varela, 1981, p.16). Another approach, perspective or focus, that of control⁴² (i.e., allonomy), is, then, necessitated. While this way of approaching systems initially appears to be in opposition to autonomy, it is in Varelian framework its counterpart—both are complementary aspects necessary for a full understanding of a system. Autonomy, thus, does not replace control, it extends, supplements it, which is in line with Varela's interest in creating a framework that, instead of viewing itself as replacing another, (here: the computer gestalt (Varela, 1981)), contains a way of undoing itself within itself.

Taking the autonomy approach, the observer focuses on the system's internal constitution and functioning which gives it a unique way of behaving which *appears*, to the observer, as autonomous. The reason is that the observer, if he aims at making sense of the system's phenomenology, has to consider the internal constitution of the system. But, since any (natural) system always exists in an environment, and since the environment is not only a

⁴¹The words come from the Greek *phainómenon*, i.e., that which appears, and *lógos*, i.e., study.

⁴²This is what Varela (1979a, 1981) calls the *computer gestalt*: a perspective based on information-and-representation or input-and-output.

neutral shell but a crust that exerts control on the system in question; taking the autonomy falls short of a full account.

The autonomous system is in fact subject to certain constraints that *appear*, to the observer, to direct its behaviour, as the observer has to look at the perturbations that the environment places on the system in order to make sense of its phenomenology. The "crust" exerting control on the system can be understood as imposing external constraints, which limits its behavior in interacting with its surroundings (and through this limitation open up a domain for experimentation which is crucial for a sense of bodily self)—much like internal constraints govern the system's internal dynamics. This can be understood as adopting a control approach. The observer, in this case, can focus on how the system appears to be controlled by external constraints, without needing to deal with its internal dynamics.

“For example, for the cell biologist, the cell is the autonomous unit he is interested in, his attention is turned to its internal working, and the environment disappears as a background of unconnected events. [...] For the physiologist, however, the cell is a node in a network of relations among cells, and thus each cell is not so much an autonomous entity but a component to be controlled [...]” (Varela, 1978, p.78)

Varela's point is that to account for the cell's phenomenology, both the autonomy and the control approach is necessary. It is useful to concentrate either on the system as functioning autonomously, bringing forth and maintaining its identity through its own internal functioning, but since the system cannot be separated from the environment that encompasses it, one cannot disregard that it is in fact constrained by its surroundings. In other words: An autonomous system, as we have discussed in Chapter 3.1 at length, does not exist in a vacuum, in isolation. It is always situated within an environment and this means that the internal constraints (the system's organization and processes) and the external constraints (environment's influence) are mutually-specifying.

Now, even if we take the autonomy and control approaches as complementary, and thus have access to both the operation of the system and its interaction with the environment, we could still fall into the trap of seeing the regularities or patterns that the system exhibits in its environment as external information that is being processed by the system. However, since an autonomous system establishes and maintains its own identity and thus its own domain of interactions—in other words, since it determines what it will and will not interact with, what is a stimulus and what is noise—the environment does not possess the information, and the system does not process it. What happens is rather that the environment perturbs the system, and the system compensates for these perturbations against the background of its stability, closure, identity.

"Such regularities [...] always refers us [*sic*] back to the unitary character of the system at hand, whether a cell, a brain or a conversation. [...] A given structure determines what constitutes the system and how it can handle perturbations from its ambient, but in [*sic*] needs no reference whatsoever to a mapping or a representation for its operation." (Varela, 1981, p.21)

Crucially, this property of being constrained does not need to be understood as being limited to a passive registering of external inputs to produce a behavior determined by these, i.e., it does not force us into an input-output characterization of the system. However, since an input-output characterization can also prove useful, especially in engineering, it also is not a perspective that can be eliminated from our discourse. As long as the researcher is aware that taking only one of these approaches is only one side of the coin, and does not tell us about the coin itself, the observer is "free"⁴³ to choose how he will characterize the system.

⁴³Let us note that, in his attempt to bring the observer to the foreground, early Varela seems to have overshot his target by allowing the observer to draw distinctions "wherever he pleases" (Varela, 1975, p. 22). This, then, taken to the extreme, would amount to saying that the observer can even distinguish a hierarchical system as being circular if he pleases, which, of course, amounts to "anything goes." But anything does *not* go. If it is the case that we take self-reference as "our filium Ariadnis through a succession of levels, [and] we dwell upon the re-union of the constituents of these levels up to our own union with the world" (Varela, 1975, p. 23), then the world has to be part of our definition just as much as we are part of the definition of the world. If we indeed stand in mutual

“To do this on purpose is quite useful, to forget that we did so is quite dangerous.”

(Varela, 1976b, p.64)

Now, for our purposes, i.e., for a characterization of living systems, we cannot lose sight of either of these sides: we always need to keep in mind that we are dealing with autonomous unities that find themselves within an environment. In fact, as we mentioned in the previous section, the environment becomes part of the definition of an organism: the organism is characterized by its organizational closure and structural coupling.

Let us close this section by returning to the closure thesis. It is a heuristic guide⁴⁴, which tells us how to proceed if we are interested in the autonomous aspects of the system: we should make ourselves familiar with how its organization closes onto itself, or in other words, how its components are in recursive interaction, or how the organization is circular. Crucial for us here is what we mentioned at the beginning of this section, namely, now that we have moved to autonomy from autopoiesis, we find ourselves in the space not only of spatially/topologically bounded (or physically independent, in Varela's words) systems, i.e., the cell, but also of those systems that express their autonomy in an *informational* way⁴⁵, i.e., the nervous system and the immune system.

negation with the world, as Varela (1975) claims, or by a reversal of terminology, if we stand in mutual affirmation (or definition) with the world, the observer cannot be allowed to draw distinctions wherever he pleases. Our account of the observer cannot be unilateral; it has to be *relational*. In 1981, the more mature Varela makes the observer part of the definition of an autonomous system and grants the autonomous system an onto-epistemological status. The system remains related to the observer's act of dividing the world into what he considers to be separable units, but the observer is no longer given full freedom to divide the world in any way he pleases, for although there are different ways of performing this act of dividing, there is always some criterion of distinction to which he must fall back: the system is distinguishable by observers because of its closure, which allows us to enter into repetitive interactions with it (Varela and Gougen, 1978).

⁴⁴Varela (1978) argues that this is based on empirical evidence, but it is not entirely clear what empirical evidence entails in this context.

⁴⁵Considering his later modified use of the word as *in-formation* in referring to “the admissible symbolic descriptions of the cognitive domains of an autonomous system” (Varela, 1979a, p.266), it is possible he is hinting at a focus on the internal coherence instead of a mapping of what-is-out-there, and on structural-coupling instead of external definition. In nuce: a focus on in-formation as interpretation rather than on information as instruction. Considering his cybernetics influences, it is also possible that he means systems that are *information-tight* in the cybernetics sense: “Cybernetics might, in fact, be defined as the study of systems that are open to energy but closed to information and control—systems that are “information-tight”.” (Ashby, 1957, p.4)

Since it is not entirely clear why non-topological systems should be called informational, we opt for adopting the terminology Varela (1991) uses in referring to these systems as "behavioural" units. This is to emphasize that they constitute *stable patterns of behaviour* that distinguish them from the environment. These patterns are "virtual" in the sense that they are not given at the level of physicochemical (i.e., "real" in the sense of empirically observable) constituents and processes. The virtual dimension underlines the eliminated necessity of a physical or topological boundary and the idea that such systems behave or function as if they were physically bounded. Therefore, we refer to these systems also as "virtually bounded" units.

3.2.3 From the Cell to the Immune and Nervous Systems

In order to discuss the autonomy of a system in a reliable and robust manner, we need to explicitly identify the components of the system in question, as well as their interrelationships. With this aim, in this section, we will explain in more detail the nature (or mode) of autonomy proper to the immune and nervous systems. Together with the minimal autopoietic system, these two autonomous systems will be our focus when examining the different levels of "self" within the Varelian framework, which we will discuss in the final chapter of this thesis.

In the case of the autonomous nervous system, the components are the neurons and the interrelations are "states of relative activity brought about through synaptic coupling." (Varela, 1981, p.18) According Maturana's theory of the nervous system as a closed neural network, its fundamental anatomical and functional organization is uniform, i.e., as in with Varela (1971), the same functions and operations take place in various parts, but in different contexts and integrated in a different manner. For the nervous system, then, there is no *information* coming from the environment that is being processed (Varela, 1977). The picture is rather that of perturbations from the environment which are compensated through

recursive interactions among neurons. Relating with our previous discussion of autonomy/control complementarity, Varela (1977) argues that the observer may adopt a broader perspective, that of control, and against this background see perturbations as stimuli or inputs, and compensations as representing behaviors or outputs. However, in actuality, rather than producing outputs based on inputs, “[a]ll the nervous system does is to specify a possible domain of interactions through its anatomy, and all it knows how to do is to maintain a recursive history of compensations through its own states.” (Varela, 1977, p.66)

It is the observer that decides on what is to be considered as *outside* and, to repeat, this can be quite useful as long as he does not incorporate such considerations into explanations of the operations of the system, i.e., as long as he is aware that he has taken the control approach and has neglected the autonomy of the system. In other words, as long as he is aware that “[t]here is nobody in the brain to whom we can refer for an assignment of correspondences” (Varela, 1981, p.21).

The same reasoning applies to the immune system. Here, the components are the clones of lymphocytes and interrelations are those of “molecular co-adaptation between the surface determinants on the lymphocytes” (Varela, 1981, p.18) Now, the domain in which the immune system exists is not the same as the one in which the nervous system exists: the closure of the autonomous unity defines its domain of interactions through its autonomy. In the previous section, we mentioned that Varela (1978) termed both nervous and immune systems informational and we pointed out the vagueness of his terminology. In 1981, he accounts for the boundary, or rather the space in which the immune system creates a boundary more precisely:

“[T]he immune system defines a boundary not in the topological sense, but rather a space of molecular interactions, by specifying what shapes can enter into the ongoing interactions of the system at every point of the time.” (Varela, 1981, p.18)

Varela and Vaz (1978) propose an understanding of the immune system not as a collection of antibodies and cell clones directed outwards, representing a large set of external antigens, but rather as a network of cooperative interactions between these cell clones:

“The basic idea is to look at the immune system as a closed network of interactions which self-determines its ongoing pattern of stability and its capacities of interaction with its environment. Thus, all immune events are understood as a form of self-recognition, and whatever falls outside this domain, shaped by genetics and ontogeny, is simply nonsensical.” (Varela and Vaz, 1978, p.231)

To repeat: an immune response is not directed towards the antigens but towards the components of the system itself. It in-forms (itself) when encountered by an antigen, i.e., an antigen is recognized on the basis of its internal representation, its similarity to a component within the network. Thus, what the immune system distinguishes between is not self and not-self, but rather self and non-sense⁴⁶.

Crucial for us is that the establishment of an autonomous unity, thus an organizationally closed system, specifies by this same act “a domain with which it can interact without loss of identity” (Varela, 1981, p.16). This, Varela calls a “cognitive domain” for the unity: the domain of “descriptive interactions relative to the environment as beheld by the observer” (Varela, 1981, p.17).

As we will discuss in more detail in Chapter 4.2, the term cognitive (phenomenic) domain relates to the mechanism of identity: the mutual-definition of the bounded unit and the internal dynamics of the whole (i.e., the whole and the parts). The creation of identity is the same process by which the unity determines its domain of interactions, or, to use the older

⁴⁶Varela (1991) actually argues the immune system brings forth a self, but his argument focuses on the autonomy of the system. Thus, the immunological self is due to the immune system closing onto itself, since new, emergent properties, such as the self, hinge upon closure.

terminology, “the permitted interrelations of machine components which define the possible transitions that the machine can undergo” (Maturana and Varela, 1972, p. 378).

Up to this point we have placed considerable emphasis on the invariance, sameness, stability of living systems in particular and autonomous systems in general. We have largely omitted to discuss the change they undergo throughout their history. In the next chapter change will be our focus. This will be the last missing piece before we move on to our discussion of the self.

3.3 Living Systems and Change

Organism is in constant change and this change has two aspects: (i) it changes to stay the same, i.e., is subject to continuous change in structure in order to keep its organization invariant, (ii) the lineage of organisms changes through *natural drift* to conserve adaptation, i.e., historical change to keep adaptation invariant.

“The notion of natural drift defends an ahistorical notion of adaptation: instead of considering that adaptations are traits evolved by natural selection, all existing organisms are well adapted to their environments in that they constitute viable ways of living and interacting with their environments.” (Etxeberria, 2004, p.16)

We will in this chapter focus on (ii) natural drift⁴⁷ and its path as specified by “history of internal validations of interactions by an autonomous entity” (Varela, 1984b, p. 219). We should note that, taking the risk of stating the obvious, (ii) is subordinated to (i), since for the organism to reproduce and evolve, it has to *be*.

⁴⁷ We note that “Maturana and Mpodozis’s theory of natural drift (TND) is independent from the observations on evolution by natural drift presented by Varela and colleagues (1991), even though both these versions are based on the same original idea (Maturana & Varela, 1987).” (Raimondi, 2021, p.4). Here, we focus on the Varelian framework.

Evolutionary biology as Varela (1987) argues, shares the common dominant perspective with neuro- and cognitive sciences, i.e., an over-proportionally strong focus on the control approach: the main characteristic of thinking in evolutionary biology, he claims, is the search for the *optimal mechanisms* with which the organism adapts to the environment, to its niche. The strategy in this search is isolating specific traits and looking for what this trait is optimum for and how.

This focus on *optimality* is the analogue in evolutionary biology to what Varela (1986b, 1987) calls the *representationalist programme* in neuro- and cognitive sciences which focuses on input coming from the *outside* and being processed *inside*. The pattern is repeated: there are properties of the outside that the organism adapts (responds) to in an optimal way. In contrast, the autonomy perspective shifts the focus, as we have seen in the previous chapters, on *internal coherence* of the whole organism.

“The perspective shifts from the sub-organismic realm to a systemic and relational one. It focuses on how invariant relations underlying a given life cycle [...] operate as an operational constraint for adaptive changes within a lineage, and how, in turn, the emergence of new relational invariants gives rise to the constitution of new lineages.”
(Raimondi, 2021, p.8)

The details, the nuances of different approaches in evolutionary biology do not concern us here. What we want to, following Varela, point out is the context, the “assumed, but usually invisible, tradition” (Varela, 1986b, p.211) against the background of which we took the autonomy approach. In Varela’s terms, what he (we) has (have) been attempting, is “a research programme where objective is an *alternative* understanding” (Varela, 1986b, p.209, my italics), in contradistinction to being the only, ultimate understanding possible.

These short remarks above should serve as a description of the control approach as *alternative* to autonomy; and autonomy approach not as *replacing* the control, but “one that can contain itself in a way of *undoing itself*” (Varela, 1979b, p.19, my emphasis).

Let us, then, restate: by focusing away from a control perspective to the autonomy perspective, we have moved away *from* structures capable of mirroring an external world and the changes in it *to* those that in-form and thus shape their environment “into a *world* through a history of structural coupling with it” (Varela, 1987, p.52, my italics). In nuce, we are moving away from mirroring or re-presenting towards *in-forming or presenting*.

Now, simply following our steps from the previous chapters, we can see that taking the autonomy approach entails at least two crucial points with respect to the above-mentioned search for optimal adaptation mechanisms. The first point is that traits cannot be neatly isolated, since the organism *is* a whole, thus lives, acts and evolves as a whole. No part of the organism can be separated from the whole and privileged over another part.

“[E]volution takes place only if adaptation is conserved by the unities that embody the invariant organization of the evolving lineage.” (Maturana and Varela, 1980, p.105)

And the second point is that the environment cannot consist of instructions for the organism because it is the organism that, based on its internal coherences and history of structural coupling, decides what is relevant for it in the environment and what is not. Thus: instead of adapting optimally to an environment, the organism functions within the framework set by the environment⁴⁸, conditions that must be satisfied in order for it to survive. In other words: The external and internal constraints, as we mentioned in section 3.2.2, are mutually specifying and together define the limits of the organism's possibilities, opening up the domain of

⁴⁸ Or more accurately: the organism functions within the framework set by the interplay of its organization and the structure of the environment.

experimentation. Within these constraints, the organism is a “free” agent, in the sense that its living, its very activity determines the path it will take: there are many possible paths, and it *drifts* naturally along them.

“The quality of laying down a sense is not given by what you give [to the system], but by the structure of the system. It can be arrived at by many possible rules. To say it metaphorically: you can lay down a path by walking in many different ways - the important thing is that you keep walking. That is what the [...] system does - it keeps walking. You walk differently than I do, we both walk - that is what matters. “ (Varela, 1986a, p.120)

This is a powerful metaphor: organisms (or species in the evolutionary context) walk differently, but they all walk, and that is what matters. It is the dynamics that is crucial, it is what distinguishes a cat from a crystal. What matters is to change continuously to keep constant what needs to be kept constant: organization and adaptation.

“For evolution to take place as a history of change in the realization of an invariant organization embodied in successively generated unities, reproduction must allow for structural change in the sequentially reproduced unities.” (Maturana & Varela, 1980, p.105)

Let us, at this point take a reflexive turn in Varelian spirit to note:

“[W]e stand in relation to the world as in a mirror that does not tell us how the world is: neither does it tell us how it is not. It reveals that it is possible to be the way we are being, and to act the way we have acted. It reveals that our experience is viable.” (Varela, 1984b, p.322)

For it to be possible to be the way an organism is being, organization and adaptation have to be kept invariant through continuous variance, by the reciprocity between structural and environmental changes:

“It is not a matter of the survival of the fittest; it is a matter of the survival of the fit. It is not the optimization of adaptation, but the conservation of adaptation that is central: a path of structural change of a lineage congruent with its environmental changes. [...] In moving from an adaptationist view to an understanding of evolution as natural drift, we have also moved from a logic of *correspondence* to a logic of *coherence*.” (Varela, 1987, p.56ff, my emphasis)

Thus, the perspective we took with regards to information, we extend to optimality and this entails: away from evolution as a (i) *prescriptive* rule, i.e., what is not allowed is forbidden, to evolution as a (ii) *proscriptive* rule, i.e., what is forbidden is not allowed (Varela, 1987). The approach (i) entails that there is a right and thus an optimal way to be in the relationship of the organism with its environment. The approach (ii) states that it is possible to be the way an organism is out of many possible ways of being.

We have, then, established that, essential for the existence of living organisms is, besides changing to keep the organization invariant (*ontogeny*), also changing to conserve adaptation (*phylogeny*).

“[E]very ontogeny as an individual history of structural change is a structural drift that occurs with conservation of organization and adaptation.” (Maturana & Varela, 1987, p.102ff)

Let us now add: Organism and environment change together as mutually-defining systems. We have seen how the environment provides constraints for the organism, let us now point

out that while changing, the organisms also change their environment (for the other organisms) in different ways: they deplete resources, they excrete into the environment, and their mere presence attracts other organisms with the same activities into the environment. The multicellular organism with a nervous system changes its environment through directed activity (*experimentation*) but an organism also modifies its environment by its mere “presence”. For example:

“At the surface of leaves of green plants is a film of air about a millimetre thick, which is richer in oxygen and moisture and poorer in carbon dioxide than the free atmosphere. For certain fungi and small insects, this surface film is their whole habitat; longer-legged leafhoppers’ bodies are above it.” (Levins & Lewontin, 1985, p.54)

The way the organism changes historically is in harmony with the way the environment changes historically, not because the organism adapts optimally to it, but because the organism and environment are *mutually specifying* (Maturana & Varela, 1987). Instead of having separate existences that give them their own independent properties in a vacuum; they have an *interpenetrating existence*: both are the object and the subject of historical processes at the same time (Levins & Lewontin, 1985).

“Organism and environment together form a system, and this henceforth determines the basic concept of life. Living, then, is the behavior induced by this bipolar system in one of its poles; and the typical patterns of living, the relative stability and specificness of behavior in any given species, represent the equilibrium achieved between the two factors that make up the situation. But so does organic structure itself: this too, not only its current behavior, represents an equilibrium relative to the long-term generic situation, being the outcome of the previous working of that situation on the substratum of life.” (Jonas, 2001, p.46)

Let us also remember at this point that the organism itself is an environment for its parts, and that even though the distinctions between inside and outside, between unit and environment, are not completely dependent on the observer and certain criteria can be stated to distinguish a system so that not everything goes; within these broad constraints, the observer can oscillate between considering the cell as a stable unit in itself and the organism as the environment; and considering the organism as a stable unit and its (external. macroscopic) environment as the resource (and threat)-bearer; or considering the organism and environment as a single stable unit.

In all of these cases, we see that— to reiterate—similar agents (whether other cells, other organisms of the same species, or organisms of different species) or agents of the same structure are always part of the environment and the mutual-definition process of environment and organism. This is no trivial point. Remembering our statement that the sense of bodily self and experience mutually specifying each other and that the sense of bodily self is precisely the organism's ability to sense how its actions modify the environment and how the actions of other agents modify the environment. Not only are the organism and environment mutually specifying, but also the *self* and the *other*.

Before closing this chapter, let us also note two important points: natural drift applies (i) not only to phylogeny but also ontogenic learning, (ii) not only to living systems but autonomous systems in general. We will not consider (ii) in any detail here, but (i) is of crucial importance.

Let us remember: The constraints set by both the organism itself and the environment not only limit certain degrees of freedom for the organism but also, at the same time, open up new dimensions, bring forth new degrees of freedom: the constraints act as enablers.

We have already stated that the *sense of bodily self arises through the stabilization of the self* (which is the identity-in-the-world). Now, with natural drift applied to ontogeny, we have

the missing piece, the “*how*” of this stabilization process. Let us, provisionally, call this process “learning” and characterize it as:

“a phenomenon of transformation of the nervous system associated to a behavioral change that takes place under maintained autopoiesis, occurs due to the continuous dynamic coupling of the state-determined phenomenology of the nervous system and the state-determined phenomenology of the ambient.” (Maturana and Varela, 1980, p.33)

The above quote, in Maturana's characteristically dense prose, suggests that what we call learning consists of changes in the nervous system that lead to new behaviour, provided the organism maintains its organization and adaptation. The nervous system, as an autonomous entity that conserves its own identity, undergoes change through natural drift owing to its plasticity⁴⁹:

“Transformation means that natural drift becomes possible due to the plasticity of the unit's structure. In this drift, adaptation is an invariant. Many paths of change are potentially possible, and which one is selected is an expression of the particular kind of structural coherence the unit has, in a continuous *tinkering*.” (Varela, 1987, p.61, my emphasis)

Just as the lineages of organisms change through natural drift in phylogenetic evolution, the same logic applies to learning when the focus shifts from a population to the brain as the unit of change (Varela, 1987).

⁴⁹ Notice that the notion of representation is nowhere to be found in this conceptualization of what learning is.

We noted that multicellular organisms with nervous systems are unique in that the movement of the self is stabilized by the constraints produced by the integration of various autonomous subsystems and the overall constraints imposed on the organism by this global whole. We added the requirement that the organism exists in an external environment of constraints and possibilities, which both necessitates and enables experimentation. What we have done in this chapter is to supplement a missing piece to our equation: natural drift serves as the underlying mechanism that sustains this interplay, and thus the process by which the self gains depth. It is the process that allows changes in the nervous system and the organism as a whole, while maintaining the main constraints (self-) imposed in living systems: organization and adaptation.

The way in which the organism comes to know (again: tacitly) the difference between the consequences of its actions and those of others, the process by and through which it experiments and thus acquires or brings about a sense of bodily self, can be characterized by a process of “tinkering⁵⁰”, to use Varela’s analogy we introduced above.

“The origin of knowledge (and the making of sense) therefore does not resemble the design of a system which is optimized to match a given external standard. We would say it resembles, rather, a tinkering, a dynamic sculpting, a building of structures from the materials available to an organism that it puts together as they appear in a drift

⁵⁰Varela borrows the term tinkering from the French biologist François Jacob. In addressing the evolution of complex objects, hierarchical levels and constraints, he writes: “Complex objects are produced by evolutionary processes in which two factors are paramount: the constraints that at every level control the systems involved, and the historical circumstances that control the actual interactions between the systems. The combination of constraints and history exists at every level, although in different proportions. Simpler objects are more dependent on constraints than on history. As complexity increases, history plays a greater part. [...] [A]s complexity increases, additional constraints appear - such as reproduction for living systems, or economic requirements for social systems. [...] The rules of the game differ at each level. New principles have, therefore, to be worked out at each level. [...] Novelty comes from previously unseen association of old material. To create is to recombine.[...] [Natural selection] works like a tinkerer - a tinkerer who does not know exactly what he is going to produce but uses whatever he finds around him whether it be pieces of string, fragments of wood, or old cardboards; in short it works like a tinkerer who uses everything at his disposal to produce some kind of workable object” (Jakob, 1977, p.1163).

which follows one of many possible paths. In this process, any actual path of tinkering will arise from the process of natural drift. The key to this process is that the consequences of any interactions are to be found, not in the nature of the perturbation that triggered them, but in the way the structure compensates for such interactions according to its dynamic landscape; the overall result is change of structure in a continuous maintenance of the integrity of the system in its medium.”
(Varela, 1984a, p.219)

The organism experiments continuously, always against the background of its sense of bodily self: a spiral that loops through a continuous trial and error, with the stabilized self as its dynamical core.

To repeat: The self is the identity-in-the-world and the sense of bodily self is what emerges when the self's movement is stabilized, through history of directed activity, i.e., experimentation, with the environment. Self is a dynamical process and once its movement is stabilized (while remaining dynamical), through natural drift and addition of new constraints (history serves as the third layer of constraint, the relational, or the virtual. The first two constraints are, to repeat: environment, i.e., external and organism, i.e., internal), a sense of bodily self acquires depth and deepens it in continuous tinkering.

Living for the multicellular organism with a nervous system (i.e., sensorimotor organism) is an ongoing experiment, ongoing as long as the organism is what it is, as long as identity is preserved. Living beings tinker with and through their actions, i.e., their activity is directed. Each new piece of tinkering, becoming part of the organism's history, influences the path along which it drifts. It becomes another constraint, giving the possibility to acquire depth to the spiral of the sense of bodily self. It is this tinkering that makes life unpredictable.

This implies, if we take a broader perspective: Life, whether in its ontogenic or phylogenic context, is not a set of instructions that a living system follows towards survival or death. There is no "correct" way to tinker⁵¹. Organisms may walk differently, but they all walk, and that is what matters.

⁵¹ "Unlike engineers, tinkerers who tackle the same problem are likely to end up with different solutions." (Jakob, 1977, p.1164)

4. THE CONCRETE AND THE VIRTUAL

“[The] precarious existence is always menaced by concern (Sorge), the need to avoid perishing, and to do this, it [the organism] is again completely dependent on matter whose characteristics are the reason for its concern. [A]lready the simplest forms of life have thus a subjective perspective as a result of this existential need.” (Weber and Varela, 2002. p.113)

In the preceding chapters, we covered autopoiesis, cell's characteristic self-organizing mechanism, it's mode of being autonomous, which is underlined by a reciprocal causality through which “the local and the global are braided together” (Varela, 1991, p.84): the local level is the chemical interactions of the metabolism and the global is the whole cell self-separating from the environment as a distinct entity thereby creating the conditions for the local interactions. Crucially, this entity cannot fully demarcate, it cannot sever completely its link with the environment as it depends on it, on the medium from which it arose. Its dependence is both for the creation and for the maintenance of its identity, so what we have at hand is a “dialectics of this mutual definition: neither rigid isolation, nor simple continuity with physical chemistry.” (Varela, 1991, p.86).

In this chapter, we revisit ground that is by now familiar, approaching it with a new focus (on identity and self) and terminology to bring together our discussions from the previous chapters. We start by looking at the organism as being characterized by a double dialectics (dialectics à la Levins and Lewontin, 1985):

- (1) a *dialectics of identity* emerging from bootstrapping of a local and a global term creating the entity, and

(2) a *dialectics of knowledge* emerging from bootstrapping of a significance term and a coupling term creating a world for that entity (for that identity).

The two dialectics, just like the terms of these dialectics, cannot be thought of as being separate processes. It is with the emergence of the entity and its paradoxical separation from the environment that the significance term is created and the environment is turned into what Varela terms a “world”⁵². Due to this paradoxicality of the organism and the environment—neither being the same nor being fully separate—that we, in our turn, need to consider the dialectics of identity separately from the dialectics of this identity’s being-in-the -world, i.e., its cognitive dimension (knowledge dimension, or world-making dimension or sense-making dimension).

“[A] living system makes itself into a [*sic*] entity distinct from its environment through a process that brings forth, through that very process, a world proper to the organism.”
(Varela, 1991, p.79)

In what follows, we will unpack both these dialectics, which together give rise to the organism and its “selfless selves”:

“Both dialectics give rise to the shifting nature of organism, ineluctably forming itself and informing where it is, and equally ineluctably implicated in the background from whence it springs forth. Organisms, those fascinating meshworks of selfless selves, no more nor less than open-ended, multi-level circular existences, always driven by the lack of significance they engender by asserting their presence.” (Varela, 1991, p.104)

⁵² With “world”, Varela is referring to “Umwelt”: “So, what you have is a laying down of a world, a laying down of a relevant ‘Umwelt’” (Varela, 1986a, p.117), and “[t]he notion of Umwelt is due to J. v. Uexküll” (Weber & Varela, 2002, p.122).

4.1 Significance and Coupling

(1) *Dialectics of identity* establishes the autonomous entity when two terms—(1a) *dynamical term* (the metabolic network, the chemical reactions within the cell, i.e., “parts” of a network) and a (1b) *global term* (the unity, the cell, that conditions components)—are bootstrapped (Varela, 1991). This has been sufficiently discussed in Chapter 3, albeit with different terminology. Here, having established the necessary background, we will be exploring in greater detail the domain of world-making of the organism.

(2) *Dialectics of knowledge* is the second tier of the organismic dialectics which covers the organism’s mode of relationship with its world and is a bootstrapping of

(2a) *significance term*: a surplus meaning the organism’s presence lends to the environment, transforming it into a world and

(2b) *coupling term*: permanent connection the organism maintains with its environment (Varela, 1991)

We will use different terms for this tier of organismic dialectics: dialectics of knowledge, cognition, world-making, sense making, identity-in-the world. We will see in what follows how they refer to the same dimension.

Let us unpack these two terms. The entity (organism, bacterial cell) separates itself from the environment through the mechanism of identity and yet remains connected to it. It is thus constantly confronted with encounters, occasions for coupling with its environment. These encounters are meaningful *for it*, *for itself*, i.e., they are not neutral occurrences waiting to be

harvested from the environment, but encounters that exist *for* and *from* a perspective that is not intrinsic to them. (Varela, 1997). This is to say:

“Whatever is encountered must be valued one way or another—like, dislike, ignore—and acted on some way or another—attraction, rejection, neutrality.” (Varela, 1991, p.97)

This is what we mean by a surplus of meaning or significance. Let us take, following Varela (1991), the example of a bacterium swimming in a sugar gradient. From an external perspective, a view from the outside, this situation can be analyzed in terms of the local impact of sucrose on membrane permeability, the flow viscosity, and the hydromechanics of flagellar movement. However, we analyze the gradient and flagellar movement because they are relevant to the bacterium as a whole, an autonomous unit. Thus, the external view inherently conflates with the bacterium's own perspective, shifting from physico-chemical principles and properties to interpretation and significance from the viewpoint of the living system:

“Remove the bacterium as a unit, and all correlations between gradients and hydrodynamic properties become environmental chemical laws, evident to us as observers but devoid of any special significance.” (Varela, 1991, p.85)

In Polanyi's words:

“[W]e accept the interests of the animal as the standard by which our own interest in the animal is determined. All biology is, in this sense, convivial.” (Polanyi, 2009, p.51)

We then have a distinction between the environment of the organism as we, the observers, perceive it, and the environment that the organism defines through the very act of creating

itself—its identity: the environment for the organism, i.e., its world. At the core of this distinction lies a "surplus of meaning," a surplus the observer recognizes when considering the organism-environment coupling—something absent when observing the environment as such (as-itself), without a for-itself, but added the moment the concrete organism enters into the picture. The organism's world is built upon the regularities of the environment, and as it remains linked to it, added meaning emerges by virtue of its perspective: sucrose gains significance as food because the organism's chemical processes—its metabolism—can utilize the sucrose to sustain its identity (Varela, 1991). The world for an organism is thus "dressed" in contrast to the raw, naked environment we would observe in its absence.

The identity of a living system is then inextricably linked to an active struggle for the sameness of the organization and thus to a perspective, a surplus of meaning. It "copes, relates and couples with the surroundings, and inescapably provides its own world of sense" (Weber and Varela, 2002, p.117).

Recall that we discussed the nullified need for the organism to have a perspective in the case of its full-blown demarcation from the environment: it would have no *need* and no *concern* for identity maintenance if it were not exchanging matter with the environment. As the "purpose" of maintaining identity vanishes in the absence of an environment, there would, of course, be no world and no perspective.

"[E]nvironment provides the "excuse" for the neural "music" from the perspective of the cognitive system [...] without the possibility of coupled activity the system would become a mere solipsistic ghost." (Varela, 1991, p.97)

All of these *surplus-terms* (need, concern, purpose, perspective, world) arise in the relation of unity to an environment and are as such *self-terms*.

Without the link that is being maintained between the organism and the environment, there would also be, as we will discuss below, no action. For the acting organism, the environment is not only a threat but also an open horizon for it to continuously *fulfill* itself. We can then, with Varela (1991), turn the idea of a surplus upside down into the idea of a lack:

“[T]his constant bringing forth of signification is what we may describe as a permanent *lack* in the living: it is constantly bringing forth a signification that is *missing*, not pre-given or pre-existent.” (Varela, 1991, p.86, italics in the original)

This lack, this *what-is-missing*, we believe, should not be interpreted as a pre-determined *item* for satisfaction of organismic needs. It should, for example, not be limited to hunger-food dynamics. It should be understood as an open horizon, it should be seen as a process without an end, a process for which the end would mean death. Constant bringing forth a world, bringing forth new contexts, performing new experiments: all of this is possible because the organism is *not fully realized* in any given point of time, its organization is never invariant, but it is continuously *kept invariant, i.e., maintained*. It is in this dynamic sense that we say the organism is neither a finished sculpture nor a finish-able one, but one that is in constant flux in order to attain and sustain a dynamic coherence which is its identity.

To revert back to our terminology of Chapter 3: It is in this sense that Varela argues for autopoietic organization being understood in a circular fashion as the organization that keeps itself invariant. Now, this active component of *keeping itself* invariant can only be thought of if there are disturbances from the environment. If the organism demarcated into a vacuum, its organization would be statically invariant and that would be the end of its living: from dynamism in an environment to stasis in a vacuum.

Let us bring in the example of a crystal: there is an environment that changes it, pushes it and pulls it but not an environment that *disturbs* it. Because the crystal does not bring in a

surplus and a lack into the environment, it cannot be disturbed. It is fully realized, it has no dimension of active keeping of its organizational invariance, it has no *sustaining* of its identity because it is not *lacking* in anything. It is at any point fully realized in its static organization determined by spatial relations of components. This is not the case with the organism: it is not fully realized. Its dynamic organization is determined by relationships between processes that in their closing on themselves open up the dimension of lack.

“[W]hat is meaningful for an organism is precisely given by its constitution as a distributed process, with an indissociable link between local processes where an interaction occurs (i.e., physico-chemical forces acting on the cell), and the coordinated entity which is the autopoietic unity, giving rise to the handling of its environment without the need to resort to a central agent that turns the handle from the outside [...]” (Varela, 1991, p.86)

This is also what we were hinting at in Chapter 3 as we were discussing the environment being both the threat and the resource for the organism. Now, we can recapture the idea with our current terminology: Self-creation coincides with the creation of the lack and the surplus; from incompleteness stems a move towards resources, a perspective aiming at continuous *rising again*. In other words: an organism's world-making stems from the continuous breakdowns in its organization, which is the source of the need of maintaining it. The breakdowns can be “minor, like changes in concentration of some metabolite, or major, like disruption of the boundary” (Varela, 1991, p.87), in the latter case the organism dies.

Crucially, every breakdown, every perturbation can be seen as the initiation of an action (Varela, 1991). Not a *movement*, but an *action*, precisely because the organism has a direction: towards what is missing, what is lacking *for* maintaining its identity. In Varela's terms this action is the *cognitive activity* of the system. To this, we will come back in the next chapter.

Let us close this chapter by means of a brief summary: The concern for identity maintenance directly emerges from the autonomous entity remaining linked to the environment from which it self-separated, and “directly emerging from the aspect of concern to affirm life, [...] meaning comes to its surrounding, introducing a difference between environment [...] and world [...]” (Weber and Varela, 2002, p.117).

4.2 Relentless Action And Cognitive Dimension

With the term *cognitive*, Varela covers a wide ground. It applies to every autonomous system as a cognitive domain, i.e., the *domain of interactions* that the system can engage in without losing its identity (Varela, 1981). As a cognitive system, it describes “an active, self-updating collection of structures capable of informing (or shaping) its surrounding medium into a world through a history of structural coupling with it” (Varela, 1987, p. 52), i.e., the system capable of creating a cognitive domain through its internal mechanisms. As a cognitive point of view, it applies to “a particular set of presuppositions and attitudes, a perspective [...] associated with some notion of value, or interest” (Goguen and Varela, 1979, p. 32). And as we will discuss in the next, and final chapter in this part, as a cognitive self, it describes “the manner in which the organism, through its own self-produced activity, becomes a distinct entity in space, but always coupled to its corresponding environment from which it remains nevertheless distinct” (Varela, 1991, p.94).

All these phenomena to which the word cognitive applies have, in Varela, the same underlying logic: that of *autonomy* and the *domain of interactions* an autonomous system specifies through its very operation.⁵³

⁵³We have already covered the rudimentary form of cognition in our discussion of the double dialectic of the organism: that of identity and that of knowledge. We can now refer to the latter as the dialectic of cognition or the dialectic of world-making.

Everything *cognitive* is what the observer perceives the autonomous system to be doing when it is coping with its environment, a domain of interactions defined by the organizational closure, where the system's actions are necessarily subordinated to maintaining identity⁵⁴. In short, the cognitive system self-determines a capacity to operate in and with its environment, be it systemic, organismic (including cellular) or linguistic (including our own cognition).

The common logic of Varela's use of the word cognition is crucial to note as it points towards the common logic of living systems. In other words, the reason why Varela is opting for extending the word *cognitive* to include the minimal living unity is because, as he himself puts it,

“the more traditional level of cognitive properties, involving the brains of multicellular animals, is in some important sense the *continuation* of the very same basic process.”

(Varela, 1991, p.88, italics in the original)

In line with our project in this thesis of tracing phenomena back to their “roots”, to their rudimentary manifestations, this brings to light the fact that “we cannot step outside the domain specified by our body and nervous system. There is no world except that experienced through those processes given to us and which makes us what we are” (Varela, 1984b, p.320).

Varela anticipates some disapproval for his use of cognition for the cell and the immune system—anything that comes before an organism with a nervous system—and he reckons that the objection would vanish if he used the word “information” instead (Varela, 1991). As we have previously discussed, this is certainly a more problematic choice of words. We will not here repeat our line of argumentation against the deployment of information and opting

⁵⁴ This comes very close to our concept of the self. The focus with Varela's cognitive domain is on the phenomenological domain created by the organism's identity, i.e., the world; whereas for us, the focus with our notion of the self is the overarching and underlying dynamics of the identity-in-the-world.

for in-formation. We wish only to emphasize that the word cognition fits—better than information—what we are trying to convey: *continuity*. Note: We are not referring to perfect linearity, i.e., to a linear transformation through the accumulation of phenomena on the same level, but rather to *jumps and leaps* in the history of cognitive systems, which we associate with domains of closure, and the constraints created by their functioning within the whole.⁵⁵

Before moving on to the cognitive domains of the three autonomous systems of interest to us and the inter-active level of *self* associated with each of them—making the organism a "network of selfless selves" (Varela, 1991, p.79)—we see merit in devoting a couple of paragraphs to the notions of cognition and knowledge.

Cognition comes from the Latin root *cognitio*, i.e., knowledge or the act of knowing; derives from the verb *cognoscere*, i.e., to get to know, recognize. With cognition, then, we have an *act* (of knowing), a *process* (of getting to know) and a *recursivity* (of re-cognizing).

The recursivity is particularly interesting. The prefix re-indicates repetition: to recognize something is to cognize it again. But how can cognition operate without re-cognition? We have a paradox, where X enters into the constitution of X: re-cognition enters into the constitution of (re)cognition and we have a process without beginning or end, a cognizing-recognizing loop: Perceiving, engaging, interacting leads to familiarity, identification, and familiarity and identification enable perceiving and engaging.

We believe this serves our argument, or the Varelian line of thinking, which posits that for us to be able to recognize (distinguish) a system, we need to be able to enter into repetitive interactions with it. The *mechanism* for cognitive processes is a mechanism "by which unities

⁵⁵For us, the three new degrees of freedom introduced by constraints - the three leaps in the history of living systems made possible by systemic closure - are: (i) minimal autopoietic systems, (ii) a sense of bodily self, and (iii) linguistic capacities. In each of these cases we observe a looping structure, a stable circularity generated by continuous dynamics. In this thesis we will address the first two levels, with the aim of revisiting (iii) in a later, more comprehensive work, drawing on (late) Varela as a basic reference.

can endow a world with a sense through their structure and history of interactions” (Varela, 1987, p.53) In other words, world-making or sense-making, or cognitive domain is open only to an identity-in-the-world that engages in repetitive interactions with its environment thereby creating a world.

A cognitive process, we argue, requires two kinds of history: *phylogenic* (i.e., evolutionary, ancestral) and *ontogenic* (i.e., developmental, experiential), both of which are covered within natural drift (cf. Chapter 3.3). Our interest will be in the latter: the ontogenic, experiential history, experimenting, learning. Cognition requires a *history of interaction with the world*, a history that is preserved so that the identity's actions in the world are not mere movements but directed activities, i.e., (what we have called) experiments. Now we are in the territory of memory, yet another important line of inquiry that we will have to leave by wayside. We will, here, have to take it as given that the organism learns and that memory is essential for learning, a rather uncontroversial point. How this memory works will not be our concern.

The prefix *co-* in *cognoscere* means together or with. This serves our argument of the object and subject being co-defining, and autonomy having an onto-epistemological status; our opting, with Varela, for the middle way between “the Scylla of representationism and the Charybdis of solipsism”. (Varela, 1984a, p.217). Similarly, if we move towards *knowledge*, we encounter the Old English suffix *-ledge*, meaning to be concerned⁵⁶ with or to act and this serves our argument that the act of living coincides with the act of knowing.

⁵⁶ The word *concern* is also of crucial importance for us, as we pointed out above. The organism is an autonomous center of concern (Weber and Varela, 2002). The etymology, again, offers a ground for discussion: it comes from Latin *concernere*, where *-cernere* means to separate or to sift. This, we believe, makes our argument for organismic concern more intelligible. The organism as a center of concern means that it is a center of distinguishing between relevant and irrelevant for itself, between useful or neutral for itself. Thus, from the onset, starting with the skeletal bio-logic of autopoiesis, we are confronted with concern, and the basic cognitive act of distinction, separation, sorting, sifting. The minimal living system continually develops knowledge of its surroundings in this sense, in the sense of being concerned with and acting according to the concern. This acting, for us, acquires a new dimension with the lifestyle of the organism becomes one of changing the environment by its action instead of changing the environment by mere presence.

We are concerned with the inter-active level, which involves mutual action between two systems: the environment and the organism at the level where identity begins to in-form *itself* as the inter-active level tilts the focus slightly more towards the organism, and we observe that the organism's mode of living evolves to involve changing the environment not only through its existence but through its directed activity. At this point we identify the *sense of bodily self*, which requires the organism to be able to experiment, which, in turn, requires the ability to distinguish the consequences of its actions from those of others, which, again, requires a sense of bodily self—the circle closes and the beginning becomes the ending. Here we encounter *experience* rather than mere receptivity, a sense of bodily self rather than mere perspectivity or point of view.

The organism knows with the world through the process of inter-acting with it in a repetitive fashion as to produce a history of couplings through which it induces structural change in the environment while changing itself structurally in its concern to keep its identity going, its organization invariant. The action called forth by perturbations triggered in the organisms, causing structural change, is, for the observer, *cognitive activity*:

“[E]very interaction of an organism, every behavior observed, can be assessed by an observer as a cognitive act.” (Maturana and Varela, 1987, p.174)

It is in this context we are using the term *to know*: Living is the act of maintaining organizational invariance at the face of perturbations, and this organization or identity, the sameness can be maintained only as the coupling with the environment is also maintained (the system is not alive in a vacuum). The living system, from simple cells and bacteria to humans, is a cognitive system and “the fact of living—of conserving structural coupling uninterruptedly as a living being—is to know in the realm of existence. In a nutshell: to live is to know (living is effective action in existence as a living being)” (ibid).

Let us now take a step back. The living system that knows is a meshwork of different autonomous systems each having, in Varelian terminology, a cognitive domain. For our purposes, three different autonomous systems of interest are: the minimal autopoietic system, the immune system and the nervous system. In the next chapter we will see how Varela (1991) assigns a *self* to each of these *self*-governing systems within the context of the organism. At this point we would like to give an account of each of these systems and their cognitive domains as to ease our way into the very last chapter of this thesis.

In Varelian (1991) framework, the organism is a meshwork of multiple “selfless” selves, where the minimal living system (the cell or bacterium) constitutes the first level: that of minimal self. At the second level, the immune system serves as the foundation for the bodily self and the third level is the nervous system giving rise to the cognitive-perceptual-motor self that we associate with animal behaviour.

The first case, that of the cell or bacterium, has already been sufficiently covered. We will omit the Varelian bodily self based on immunological foundations, and instead focus on activity associated with nervous system closure. This is the layer in which we formulate our idea: the sense of bodily self as the ability to distinguish between the effects of one’s actions from that of others and from that of nature’s moves. Our focus is on pre-linguistic organisms, and for this reason we will also omit the next two levels of (Varelian) self, the *socio-linguistic* and the *collective social* levels. Thus, from now on, our focus will be on the third level of Varela's meshwork of selfless selves—the *cognitive perceptuo-motor self*.

We have seen above that the word cognitive applies to the cell as a minimal living system as well as the nervous and the immune systems designating the *domain of possible interactions* specified by the internal dynamics of the system and its history of coupling with the environment. This is to say, each autonomous system specifies its own, unique cognitive

domain, that of the cell is different than that of the immune system and the domain in which the immune system exists is not the same as the one in which the nervous system exists.

In the context of organisms endowed with a nervous system, we see a circularity, a coupling between movement and sensory modulation, with the net result being “perception-action correlations arising from and modulated by an ensemble of intervening neurons, the interneuron network” (Varela, 1991, p.89). Just as the metabolism-membrane circularity of autopoiesis is the basic bio-logic, the perception-action circularity is the basic neuro-logic. The traditional cognitive systems involving nervous systems is a continuation of the cognitive domain of the basic bio-logical system and that of the immune system. In all these cases, the system *decides* what is relevant in its environment and what can be done in the environment:

1. Through its internal coherence, the nervous system defines the range of admissible sensorimotor interactions the organism can engage in. (Varela, 1984b)

2. The bio- logic of the cell, i.e., the circularity of its organization, establishes a "categorical classification" of events or objects in its environment and thus endows them with significance (i.e., sugar is significant for the bacterial cell because it can be metabolized). (Glasserfeld and Varela, 1985)

3. Through operational closure and its recursive history, the immune system identifies what is relevant to it, i.e., what is recognized as an antigen is recognized as such because of its similarity to a pre-existing set of determinants in the network. (Vaz and Varela, 1978)

The details of the functioning of these systems do not concern us here. For our purposes, the common logic of a creation of cognitive domain through closure of organization is crucial.

We are now sufficiently familiar with the idea of a dialectical relationship between the sense of bodily self and experience and with the idea that the former rests on the organism's ability to perform experiments in an environment in a historical fashion. We also have seen that an experiment is a directed activity, a sort of testing or tinkering that in-forms the organism for past as well as future conduct, and that: for this, (1) an external environment as opposed to an intra-organismic one, and (2) a coupling between perception and action is necessary.

“Whenever motion is an integral part of the lifestyle of a multicellular, there is a corresponding development of a nervous system linking effector (muscles, secretion) and sensory surfaces (sense organs, nerve endings).” (Varela, 1991, p.89)

We might ask, since the bacteria also has a coupling between its sensory and effector surfaces, what is special about the nervous system? What does it bring into the picture? The main difference is *the interneuron network*. In the bacteria, the coupling is situated at the first layer of the organismic selves: that of minimal autopoietic unity. It is *an* autonomous system that is maintaining its identity, one might call it, with limited accuracy, the first-order autopoietic system.

In multicellulars, we see several autonomous systems together giving rise to the organism. Remaining at the third layer of the self, that of cognitive, we see that we have three autonomous systems within the autonomous system that is the organism: the cellular, the immunological, the nervous. As we have seen previously, each time we have a system curling onto itself, i.e., each time we have closure of the system, we have emergent properties. An autonomous system made up of closure of three systems, then, is bound to give rise to a quite extended array of properties compared to an autonomous system that is the closure of itself. This aspect we have already covered in our discussion of constraints, boundary conditions and emergent properties in Section 3.2.1. Let us here add: The emergent property in the case of nervous system closure is the *behavior* of the organism.

“[T]he nervous system is organized by the operational closure of a network of reciprocally related modular sub-networks giving rise to ensembles of coherent activity such that: (i) they continuously mediate invariant patterns of sensory-motor correlation of the sensory and effector surfaces; (ii) give rise to a behavior for the total organism as a mobile unit in space.” (Varela, 1991, p.94)

The activity for the multicellular organism endowed with a nervous system produces and is produced by the autonomous activity of the nervous system. In other words, the activity of the nervous system is brought about by the activity (or motion, as directed-ness is not a necessity) of the whole organism. Varela (1991) points out that to a large extent, it is the case that behavior⁵⁷ is a regulation of perception. This is a non-trivial but, we believe, for contemporary cognitive science a rather uncontroversial point, that we, once more will have to omit.

“The perceptuo-motor coherencies we describe externally as behavior disguises the arising, within the interneuron net, of a large sub-set- an ensemble as is usually said- of transiently correlated neurons. These ensembles are both the source and the result from the activity of the sensory and effector surfaces.” (Varela, 1991, p.89)

Besides the sheer amount of connections within the nervous system and its closure leading to the consequence of it only having access to its own internal activity, we also have the dimension of the uniqueness of neurons in that “in their axonal and dendritic ramifications permitting multiple contacts and extending for large distances (relative to cellular soma sizes) providing the essential medium for this intra-organismic sensor-effector correlation” (Varela, 1991, p.89). Let us stop here as the neuroscientific details do not concern us, but let us

⁵⁷ Behavior, as we see it, is intrinsically associated with a sense of bodily self, it is the counterpart of what we call directed activity. We opt for using the latter in what follows.

mention that the best characterization of the different ways in which the nervous system closes onto itself can be found in Varela (1984b).

4.3 Organism And Its Selves

In Chapter 4.1, we discussed *the local* (i.e., dynamical) element where interaction happens and *the global* element that guides or conditions the local are mutually-defining terms that together account for the cell's identity: the metabolic net and the cellular membrane bootstrap to account for the partial demarcation of the cell. Once this identity produces *itself*, partially *self-separates* from the environment and enters into mutual definition with it, it must continually maintain this link with the environment, just as identity is continuously maintained. The mechanisms for identity maintenance and the maintenance of structural coupling with the environment are mutually defining: because the organism only partially separates from the environment, the possibility (surplus-horizon) and the necessity (lack-need) of active identity maintenance emerges.

In other words, the coupling with the environment is what makes it possible and necessary for the organism to preserve its identity. It is due to this "need" that the organism acquires—or is said to have—a perspective (for instance, sugar is meaningful for the bacterium). This is what Varela terms the surplus of significance or importance: the sugar is important to the bacterium because it signifies food, and it is significant because the organism perpetually lacks sustenance (support, maintenance). Thus, we encounter the second dialectic in this mutual definition between the environment and the organism: a bootstrapping of the significance term (e.g., cellular semantics in case of the cell) and the coupling term (e.g., physico-chemical laws for the cell).

In this chapter, we will approach the organism, the autonomous center of concern, as an assemblage of *regional selves* with their particular mode of constitution (Varela, 1991; Weber and Varela, 2002).

Let us begin with the account of the *self* in the Varelian framework, which is a dynamic concept emerging in the movement between the two dialectics:

“[W]e have on the one hand the dialectics of identity of self; on the other hand the dialectics through which this identity, once established, brings forth a world from an environment. Identity and knowledge stand in relation to each other as two sides of a single process that forms the core of the dialectics of all selves.” (Varela, 1991, p.102)

Self-referential situations, i.e., situations where an X says something about *itself*, are in our opinion a good place to start thinking about the notion of self. Those situations let us realize that the concept of self-reference itself is self-referential, and the *self* as a concept is hidden in the phenomenon of self-reference: it is not the self that is engaging in the referencing, it is the X. We can extend this remark on self-reference into all those situations in which the observer is forced to consider a dynamical entity within an environment, or one that enters in mutual definition with it. In other words, in all those situations where the entity in question establishes self-reference in its surrounding medium. In such situations, the observer is forced to continuously oscillate between the entity and the environment unless the decision is made to let go of one end of the oscillation and focus on the one term or the other. The self arises in this oscillation and it is what necessitates this oscillation, it is nowhere to be found except for in a relational context of two mutually defining systems.

Now, the crucial point is that the self is not an epistemological concept that forces an observer to oscillate between two perspectives. For us it is an onto-epistemological concept, just as autonomy is for Varela. This is to say: A living organism creates a self-referential

situation within its environment by distinguishing *itself* from its background while remaining linked to it, thereby creating a world, i.e., treating the environment from a perspective (Varela, 1991). Through this perspective, it classifies environmental events as beneficial or detrimental, thereby forcing the observer to consider both *it* and its environment in order to understand its functioning.

In the Varelian (1991) framework of the organism as a *meshwork of selfless selves*, the minimal autopoietic unit, the cell, occupies the first level. The fundamental bio-logic of this *minimal self* finds its continuation towards the more traditional level of cognitive properties: that of neuro-logic. The shift from minimal to a *cognitive self* of the animal endowed with a nervous system bypasses the intermediate stage, that of the *immunological self*. It is here that Varela (1991) introduces a bodily self: clonal antibody networks (dynamical term) and self/non-self⁵⁸ discrimination (global term) bootstrap for the immune system distinguishing itself from the (intra-organismic) environment and self as somatic assertion (significance term) for the interaction with intra-organismic molecules (coupling term) bootstrap for the system's sense-making or world-making.

As organisms become multicellular and we shift our focus to a collective of cells (and selves) the intermediary step of the immune system underscores an *internal regulation of coexisting cells*, giving rise to a bodily self:

"The shift from minimal cellularity to organism with nervous system [...] is a transition in units of selection, and one that implicates the somatic balance of differentiated populations of cells in an adult organism, as well as crafty development pathways to establish a bodily structure. As Buss has stated recently: "The evolution of development is the generation of a 'somatic ecology' that mediates potential conflicts between cell and the individual, while the organism is simultaneously interacting

⁵⁸ We prefer the "self vs non-sense" as was used in Vaz and Varela (1978).

effectively with the extrasomatic environment" [...] For most vertebrates, this "somatic ecology" is bound together through the network of lymphocytes that constitute the core of the immune system." (Varela, 1991, p.88)

When we refer to a (sense of) bodily self, however, we are not referring to this intermediary step of immune system closure, but rather to the sensori-motor closure. The reason, as we have pointed earlier, is that our focus is primarily on the organism that is capable of *directed activity*, i.e., the animal that performs experiments. In other words, we are interested in organisms that do not only alter their environment by their mere presence but also by action. The global term we require for this is a *body-in-space* and the dynamical term is the *nervous system activity* (neural assemblies).

This means that we situate the sense of bodily self at the level of the organism that is no longer physically bounded⁵⁹ and no longer confined in the somatic environment. We will come back to the former statement below; as for the latter, a few words: We are interested in the level of the organism that is perpetually getting to know the contingencies of the extra-organismic environment through getting to know its invariance, or re-discovering its own *certainties* through continuously getting to know the *uncertainties* of the environment.

The important difference between the intra-organismic and extra-organismic environments, in our view, lies in the richness of the latter, in its invitation for exploration. It is the combination of the organism who can sense and act as well as the environment that in Varela's words, "lends itself" to (perception and) action, lends itself to meaning and thus creates the necessary background for a sense of bodily self. To reiterate: the organism capable of action does not modify the environment by its mere presence, but rather through its directed activity, its experimenting with the latter's myriad of possibilities. To borrow Varela's words:

⁵⁹Skin does not function as the membrane for the cell.

“[T]he nature of the environment for a cognitive self acquires a curious status: it is that which *lends itself (es lehnt sich an...)* to a surplus of significance. Like jazz improvisation, environment provides the "excuse" for the neural "music" from the perspective of the cognitive system involved. At the same time, the organism cannot live without this constant coupling and the constantly emerging regularities; without the possibility of coupled activity the system would become a mere solipsistic ghost.”
(Varela, 1991, p.97, italics in the original)

Let us here qualify our remark about the *certainties of the organism* and *uncertainties of the environment*. Surely, the organism is not a pre-determined, fixed entity, rather, it is one of a perspective, creation or laying down of a world. Furthermore, the environment surely offers plenty that is certain, it is how the world of the organism is built: on the regularities of the environmental conditions. For the organism as well as the environment, our terminology has been revolving around words like coherence, pattern, stability. So, what do we aim at when we talk about certainties of the organism and uncertainties of the environment?

We are aiming at emphasizing two crucial points. The first is that *the organism is constrained by its constitution*. For it, certain domains are open, others are possible, yet others are forever closed. This is to say:

“We cannot step outside the domain specified by our body and nervous system. There is no world except that experienced through those processes given to us and which makes us what we are. We find ourselves in a cognitive domain, and we cannot leap out of it or choose its beginnings or modes.” (Varela, 1984b, p.320)

Though these remarks refer to the subject-object relationship at the level of the observer—the human being—we believe they also apply to the domain of the non-linguistic cognitive self: the organism cannot step out of the confines specified by the closure of its

nervous system. The world of the cognitive self is what makes the organism what it is, while at the same time, it is given to the organism through the latter's constraints. The organism cannot but maintain its identity or its coupling with the environment. The environment as uncertainty (maintaining the coupling), enters into the definition of the certainty (maintaining the identity) of the organism.

The second point we wish to emphasize is that *the certainties created by the dynamics of the organism are nevertheless neither exhaustible nor pre-determinable*. We have here, once more, a case of constraints creating new degrees of freedom, closure creating new domains. We may speak of a necessary freedom, referring back to Hans Jonas "in speaking of "necessity" and "freedom" as the basic features (and paradoxes) of organic life" (Weber and Varela, 2002, p. 102).

The organism, then, is getting to know the uncertainties of the environment through its own certainties (and vice versa) and this does not imply complete rigidity or complete flexibility on either part, but points at flexibility through rigidity. Let us refer back to our previous terminology of identity and coupling. The rigidity is the necessity of identity maintenance and the flexibility is the domain this necessity for maintenance opens up: the coupling with the environment from the perspective of the organism. And a whole new domain with the organism as a behavioral entity whose world is laid down by perception and action loops: the organism who experiments in its environment. This is a crucial point for us as our idea rests on the necessity for the organism to experiment with its surroundings and itself. We claim that this is the unifying thread between experience and sense of bodily self.

To sum up: when we refer to a (sense of) bodily self, we will situate ourselves at the level of organisms characterized by the closure of the nervous system, or the sensori-motor closure. We are interested in, in Varela's words:

“situated agents, continually coming up with what to do faced with ongoing parallel activities in their various perceptuo-motor systems. This continual re-definition of what to do is not at all like a plan, stored in a repertoire of potential alternatives, but enormously dependent on contingency, improvisation, and more flexible than planning.” (Varela, 1991, p.96)

It is in the third layer of the Varelian (1991) meshwork of organismic selves, i.e., cognitive perceptuo-motor self, that we encounter this layer, that of nervous system closure. Cognitive self is no longer a physically bounded system, as the case with the minimal self, but rather a “behavioral” one:

“The operational closure of the nervous system then brings forth a specific mode of coherence, which is embedded in the organism. This coherence is a cognitive self: a unit of perception/motion in space, sensory-motor invariances mediated through the interneuron network.” (Varela, 1991, p.94)

The crucial point is that the aspect of identity has changed as the global term shifts from a topological boundary in space, i.e., the cell membrane, to a virtual boundary created by the closure of the nervous system, i.e., the sensorium-motorium connection of the organism, from the domain of which it cannot break out. Accordingly, at this level, we have the significance term, so the surplus created by the organism’s identity, shifting away from cellular semantics and self-nonsense assertion towards behavioral perception and action (Varela, 1991). The organism’s surplus, then, is no longer merely due to its presence but due to its directed activity. The organism perceives to act and acts to perceive and the closure of this circle is its surplus of signification in the extra-organismic environment. In nuce, the cognitive self is “a distinct coherent self which, by the very same process of constituting itself, configures an external world of perception and action” (Varela, 1991, p.94).

Let us now wrap up the whole discussion and conclude with a comprehensive summary. The organism, by virtue of its identity creation only partially separates from the environment, differentiates itself from a background by linking itself to it (Varela, 1991). This creation of identity, by virtue of the partial separation, is what lends the organism a perspective. As the organism does not self-separate into a vacuum, i.e., it does not cut itself off completely from its surroundings, it needs to, so to speak, “extend itself” into the environment to be able to maintain its identity. In other words, the organism has to maintain its identity by always preserving a coherent (dynamically stable) relation with its environment that allows for this maintenance. Or put briefly: the need for identity-maintenance is created by partial-separation. The link with the environment needs to be maintained because the identity needs to be maintained: in the movement between the identity and structural coupling, we see the self. With added constraints emerging from the meshing of autonomous systems within a coherent whole, this movement of the self is stabilized with repeated activity and acquires depth as each experience functions as an added constraint. This stabilization gives rise to the sense of bodily self. We can think of this stabilization of movement as a spiral with a dynamical core that acquires depth the more it loops around itself, giving rise to new degrees of freedom. Through repeated experimentation the organism acquires the ability to distinguish between the consequences of its own actions from that of others. But, crucially, being able to experiment and thus experience, being capable of directed activity requires the organism to have a sense of bodily self, as it is this that operates as the frame of reference for directed activity. Each sensory interaction or motor response constrains the organism's view of what belongs to the realm of consequences of its actions and what belongs to the realm of consequences of others actions and environmental factors. This was the core of our discussion in this thesis.

5. CONCLUSION

“I cannot say too strongly that this book is offered in the spirit of synthesis and exploration, not of treatise, dogma, or set opinion. Nothing presented here can be regarded as fixed, and I am well aware of it.” (Varela, 1979a, xviii)

“The problems are solved, not by giving new information, but by arranging what we have always known.” (Wittgenstein, 1986, p.47)

We began our discussions by critiquing the inadequacy of reductionist tendencies in "as" frameworks, which restrict central issues in cognitive science to mere computational or functional analogies. With Varela, we rooted our analysis in the onto-epistemological status of autonomous systems and the fundamental bio-logic of the minimal living system. By following his footsteps, we were able to apply the same logic consistently at both the lowest level (the minimal living system) and the highest level (epistemology). In this framework we identified an overarching structure—a conceptual arch that both shelters our inquiry and grounds it in fundamental principles. By anchoring our analysis in Varela's autonomy perspective, we emphasized the emergent, dynamic nature of the self instead of seeing the self as a static construct.

With Varela, we argued that the self is the constant interplay of what-is-in-here and what-is-out-there, an interplay between lack and surplus, emerging from the organism's ongoing dynamicity in maintaining its identity within an environment.

Each chapter of this thesis built towards the final synthesis by first contrasting modes of inquiry within cognitive science (epistemology), then grounding the self in the logics of minimal living systems (fundamental bio-logic), extending this to the autonomy of multicellular organisms (emergence), and finally bridging these foundations to develop our thesis that the sense of the bodily self arises in a dialectical relationship with experience in multicellular organisms with nervous systems.

The organism that partially separates from the environment, creates the possibility (surplus-horizon) and the necessity (lack-need) of active identity maintenance. Identity creation, i.e., *dialectics of identity*, has logical but not ontological priority over world-creation, i.e., *dialectics of knowledge*⁶⁰. This organism, which creates itself as a part *a-part* from the environment while remaining an integral piece of and within it, is the paradox inherent to the organism—much like the organism that creates itself only to fulfil the condition of life, which is the impossibility of full self-realization (i.e., death) and existing only within the oscillation or movement between lack and surplus. The two dialectics, then, in their interrelation give rise to the self, the organism as a center of concern without needing any external agent pulling, pushing or directing it. This applies to the bio-logic, neuro-logic as well as the socio-linguistic-logic, albeit with different, more or less concrete, manifestations.

We have accounted for the emergent properties and identified their mechanism by invoking the idea of constraints and operational closure. The first order of constraint, the first layer of operational closure, we argued, is the single cell, the single layered autonomous system maintaining its identity and coupling. The second layer of closure is associated with multiple interoperating autonomous systems, which help to maintain the identity and coupling of the whole they are part of, while simultaneously maintaining their own identity and coupling. Finally, the third layer of constraints, as we argued, is the virtual dimension created by a

⁶⁰ The two dialectics are ontologically inseparable but for there to be an interaction of the organism with the environment, the existence of the organism is presupposed.

recorded history of interactions of an autonomous system with its environment, the self created as the dynamicity between two dialectics. The self stabilizes through repeated directed activity to give rise to a sense of bodily self. Our customary analogy: the spiral with a dynamic core that is stabilized and can loop around itself with each experiment to gain depth.

We saw that the cognitive activity of the system is the activity that arises due to breakdowns to its organization, perturbations from the environment that needs to be compensated. This is *behavior* at the level of organisms with nervous systems that change their environment not only by their mere presence and re-activity but through their directed activity. Behavior is the emergent term of the third layer of Varelian self: cognitive self, where neural assemblies give rise to sensorimotor-body-in-space-and-time, with behavioral perception and action as the surplus of signification or the lack that the organism brings into the environment with macroscopic physical properties. The surplus of signification is always the organism's orientedness towards what is missing and should be thought of in terms of an open horizon that "lends itself" to experimentation (Varela, 1991).

It is only with the necessity and possibility of experimentation that we have the notion of sense of bodily self. At the risk of repeating ourselves: As opposed to a basic form of perspectivity, the sense of bodily self requires the functioning of the whole, i.e., the organism, made up of multiple autonomous subsystems working within it by their own modes of existence. A unicellular organism, despite its autonomy and history of interactions, lacks the layered structure needed to support the stabilization of the self (the movement, the oscillation). In multicellular organisms, on the other hand, the sense of bodily self stabilizes over time through repeated (directed) activity, functioning like a spiral where each iteration deepens the organismic coherence and creates new degrees of freedom. This frame of reference is what enables the organism to distinguish between the consequences of its own

actions from those of other agents, and agentless changes in the environment, which has been our definition for a sense of bodily self.

Now, as we approach the conclusion of this thesis, we would like to take a reflexive turn in the genuine Varelian spirit⁶¹. Varela's ideas provide a fundamental framework not only for (cognitive) science and philosophy, but also for everyday life, the ordinary life of a scientist. However, or perhaps because of this, his ideas are often characterized by a certain ambiguity: his extremely prolific output, leaving certain ideas underdeveloped, and his radical interdisciplinary approach, rendering his research (at first sight) somewhat fragmented. Despite his notable recognition within the field of cognitive science, a considerable portion of his key ideas remains either overlooked or insufficiently acknowledged. We believe that one reason for this may be the absence of a comprehensive, “holistic” study of his project characterized, as discussed in the introduction, by a lifelong interest in the emergent selves and a strong focus on epistemology.

We are convinced that attempting to build a framework within cognitive science, be it an enactive one or not, without a proper grounding in the autonomy approach, without a view from within, and without alternative research programmes, is not only insufficient but ultimately undermines the potential and *raison d'être* of cognitive science. Without an understanding at the roots, at the level of bio-logic of the minimal living system, as well as the neuro-logic of the cognitive self, the very core of (enactive) cognition risks being diluted or lost altogether.

Thus, we believe that one of the main strengths of this thesis lies precisely in the fact that we have focused on Varela's (mostly) early work. This is what gives the thesis its true discursive value, a quality we have tried to cultivate from the outset. We started our discussions by arguing that, if we aim to address the notion of the self or the subject directly, we must take

⁶¹ This phrase is borrowed from Sebastjan Vörös.

Varela's contribution seriously: his insistence to build a framework of science by engaging earnestly with epistemology and by not losing sight of the alternative perspective of studying systems, which is the autonomy point of view with its emphasis on the view from within and thus links the knower with the known within a broader framework.

One reason for our insistence on primary literature and (mostly) early Varela is that the considerable interest in Varela and enactivism, while often lacking a comprehensive grounding in his theoretical position or his overarching project, coupled with the lack of conviction to tackle the problems of experience, mind and subjectivity directly represents, in our view, a significant gap in the field of cognitive science. This oversight or neglect is particularly unfortunate given the characteristic of Varela's ideas to provide the necessary ground for understanding not only other systems, other living beings, and other identities and selves from a detached, objective, external perspective, but also being rooted in the necessity and possibility of learning about our own identity through engaging with other identities, and vice versa.

In seeking to build a science of mind, in trying to make any sort of sense of the socio-linguistic animals that are human beings, we need to recognize that we are dealing with multiple autonomous systems and varying levels of constraints acting on and through the individual. A purely external perspective fails to capture the full complexity of the issues that cognitive science seeks to address. What we observe is that the boundaries between the inner and the outer, the self and the world, the observer and the observed, and the organism and its environment are separable—but not separate, interrelated but never fully merged.

“It becomes apparent that if certain facts about our common experience of perception, or what we might call the inside world, can be revealed by an extended study of what we call, in contrast, the outside world, then an equally extended study of this inside world will reveal, in turn, the facts first met with in the world outside: for what we

approach, in either case, from one side or the other, is the common boundary between them.” (Spencer-Brown, 1969, p. xxv)

REFERENCES

Ashby, W. R. (1956). *An Introduction to Cybernetics*. Chapman & Hall, London.

Bortoft, H. (2012) *Taking Appearance Seriously: The Dynamic Way of Seeing in Goethe and European Thought*. Edinburgh: Floris Books.

Brockman, J. (1996). *The Third Culture: Beyond the Scientific Revolution*. Simon & Schuster.

Dupuy, J. P. (2009). *On the Origins of Cognitive Science: The Mechanization of the Mind*. MIT Press.

Etcheberria, A. (2004). *Autopoiesis and Natural Drift: Genetic Information, Reproduction, and Evolution Revisited*. *Artificial Life*, 10(3), 347-360.

Friedenberg, J., & Silverman, G. (2006). *Cognitive science: An introduction to the study of the mind*. Sage Publications Ltd.

Glaserfeld, E. von & Varela, F. J. (1985). *Problems of Knowledge and Cognizing Organisms*. *Aprendizagem/Desenvolvimento (Learning and Development)*, 2(6), 17-23.

Goguen, J. A., & Varela, F. J. (1979). *Systems and distinctions; duality and complementarity*. *International Journal of General Systems*, 5(1), 31-43.

Goldstein, J. (2012). *Polanyi's Anti-Reductionist Philosophy of Biology: Introduction to a Classic Paper in ECO*. *Emergence: Complexity & Organization*, 14, 139-153.

Jakob, F. (1977). *Evolution and tinkering*. *Science*, 196(4295), 1161-1166.

Jonas, H. (2001). *The Phenomenon of Life: Toward a Philosophical Biology*. Northwestern University Press.

Levins, R., & Lewontin, R. (1985). *The Dialectical Biologist*. Harvard University Press.

Varela, F., & Maturana, H. (1972). Mechanism and biological explanation. *Philosophy of Science*, 39(3), 378-382.

Maturana, H. R., & Varela, F. J. (1980). *Autopoiesis and Cognition: The Realization of the Living*. D. Reidel Publishing Company.

Maturana, H. R., & Varela, F. J. (1987). *The Tree of Knowledge: The Biological Roots of Human Understanding*. New Science Library/Shambhala Publications.

Pattee, H. H. (1973). *Hierarchy Theory: The Challenge of Complex Systems*. Edited by Howard H. Pattee. G. Braziller.

Pattee, H. H. (1977). Dynamic and linguistic modes of complex systems. *International Journal of General Systems*, 3(4), 259-266.

Pattee, H. H., Rączaszek-Leonardi, J., & Pattee, H. H. (2012). Laws and constraints, symbols and languages. *LAWS, LANGUAGE and LIFE: Howard Pattee's Classic Papers on the Physics of Symbols with Contemporary Commentary*, 81-89.

Pattee, H. H. (2013). Epistemic, evolutionary, and physical conditions for biological information. *Biosemiotics*, 6(1), 9-31.

Polanyi, M. (1968). *Life's Irreducible Structure: Live mechanisms and information in DNA are boundary conditions with a sequence of boundaries above them*. *Science*, 160(3834), 1308-1312.

Polanyi, M. (2009). The tacit dimension. In *Knowledge in Organisations* (pp. 135-146). Routledge.

Raimondi, V. (2021). Autopoiesis and evolution: The role of organisms in natural drift. *Adaptive Behavior*, 29(5), 511-522.

Rosenblueth, A., Wiener, N., & Bigelow, J. (1943). Behavior, purpose and teleology. *Philosophy of science*, 10(1), 18-24.

Rudrauf, D., Lutz, A., Cosmelli, D., Lachaux, J. P., & Le Van Quyen, M. (2003). From autopoiesis to neurophenomenology: Francisco Varela's exploration of the biophysics of being. *Biological research*, 36(1), 27-65.

Spencer Brown, G. (1969). Laws of for. *Allen and Unvvin*.

Thompson, E. (2007). *Mind in Life: Biology, Phenomenology, and the Sciences of Mind*. Belknap Press/Harvard University Press.

Varela, F. J. (1975). A calculus for self-reference. *International Journal of General Systems*, 2, 5-24.

Varela, F. J. (1976a). On observing natural systems: Francisco Varela in conversation with Donna Johnson. *CoEvolution Quarterly*, 10, 26–31.

Varela, F. J. (1976b). Not one, not two. *Coevolution Quarterly*, 12, 62-67.

Varela, F. J. (1977). The nervous system as a closed network. *Brain Theory Newsletter*, 2(4), 66–67.

Varela, F. J. (1978). On Being Autonomous: The Lessons of Natural History for Systems Theory. In G. J. Klir (Ed.), *Applied General Systems Research* (Vol. 5, pp. 1-19). Springer, Boston, MA.

Varela, F. J., & Goguen, J. A. (1978). *The Arithmetic of Closure*. Cybernetics and System, 8(3-4), 291-324.

Vaz, N. M., & Varela, F. J. (1978). Self and non-sense: An organism-centered approach to immunology. *Medical Hypotheses*, 4(3), 231-267.

Varela, F. J. (1979a). *Principles of Biological Autonomy*. Elsevier, New York & Oxford.

Varela, F. J., Maturana, H. R., & Uribe, R. (1974). Autopoiesis: The organization of living systems, its characterization and a model. *Biosystems*, 5(4), 187-196.

Varela, F. J. (1979b). *Reflections on the Chilean Civil War*. Lindisfarne Letter, 8 (winter), 13-19.

Varela, F. J. (1981). Autonomy and autopoiesis. In *Self-organizing systems: An interdisciplinary approach* (pp. 14-24). Campus Verlag.

Varela, F. J. (1984a). Living ways of sense-making: A middle path for neuroscience. In P. Livingstone (Ed.), *Order and disorder: Proceedings of the Stanford International Symposium* (pp. 208-224). Anma Libri, Stanford.

Varela, F. J. (1984b). The creative circle: Sketches on the natural history of circularity. *The Invented Reality*, 309-324.

Varela, F. (1986a). Steps to a cybernetics of autonomy. In *Power, Autonomy, Utopia: New Approaches Toward Complex Systems* (pp. 117-122). Springer US.

Varela, F. J. (1986b). Experimental epistemology: Background and future. *Cahiers du Centre de Recherche, Epistemologie et Autonomie (CREA)*, 9, 107–121.

Varela, F. J. (1987). Laying down a path in walking: A biologist's look at a new biology. *Cybernetics*, 2, 6–15.

Varela, F. J. (1991). Organism: A meshwork of selfless selves. In *Organism and the Origins of Self* (pp. 79-107). Springer Netherlands.

Varela, F. J. (1997). Patterns of life: Intertwining identity and cognition. *Brain and Cognition*, 34(1), 72-87.

Von Foerster, H. (2013). *The Beginning of Heaven and Earth Has No Name: Seven Days with Second-Order Cybernetics*. Fordham Univ Press.

Vörös, S., & Halák, J. (2024, August 25). *Life as the advent of the problematic* [Video]. The Organic and the Normative conference. <https://www.youtube.com/watch?v=IVXt8K0rY8s>

Vörös, S. (2023). At the Cradle of Things The Act of Distinction and Francisco Varela's Non-Dualist Thought. *Journal of Consciousness Studies*, 30(11-12), 17-46.

Weber, A., & Varela, F. J. (2002). Life after Kant: Natural purposes and the autopoietic foundations of biological individuality. *Phenomenology and the Cognitive Sciences*, 1(2), 97-125.

Wiener, N. (1948) *Cybernetics: Or, Control and Communication in the Animal and the Machine*. MIT University Press, Cambridge, MA, 212.

Wittgenstein, L. (1986). Part I. In G. Anscombe, R. Rhees, & G. Von Wright (Eds.), *Philosophical Investigations* (3rd ed., pp. 1-172). Basil Blackwell.