



In the Case of Protosemiosis: Indexicality vs. Iconicity of Proteins

Dan Faltýnek¹ · Ludmila Lacková¹ 

Received: 17 April 2020 / Accepted: 2 November 2020 / Published online: 18 November 2020
© The Author(s) 2020

Abstract

The concept of protosemiosis or semiosis at the lower levels of the living goes back to Giorgio Prodi, Thomas A. Sebeok and others. More recently, a typology of proto-signs was introduced by Sharov and Vehkavaara. Kull uses the term of vegetative semiosis, defined by iconicity, when referring to plants and lower organism semiosis. The criteria for the typology of proto-signs by Sharov and Vehkavaara are mostly based on two important presuppositions: agency and a lack of representation in low-level semiosis. We would like to focus on an alternative approach to protosign classification. In particular, we aim to provide a sign-typological characteristic of proteins (in analogy to Maran's classification of environmental signs). Our approach is focused on representation, that is, we only consider the relation between a sign and its object. We are considering representation independently from the role of interpretant and interpretation (which is an epiphenomenon of agency). Two hypotheses are investigated and accordingly evaluated in this paper: (I) Proteins are indexical protosigns. (II) Proteins are iconic protosigns. The conclusion our argumentation leads to supports the hypothesis (II).

Keywords Genetic code · Iconicity · Indexicality · Protein · Protosemiosis · Protosigns · Sign typology

Introduction

For a long time, life has been thought of as a sign phenomenon. This means that among natural languages and other sign systems, such as orientation marks or artefacts, rules for behaviour within a specific culture, the very essence of the existence of an organism, even its behaviours, might be classified. This is valid even if we think of life as a sign phenomenon in the context of culture only in the metaphorical sense – that is, not as a sign behaviour of higher organisms or their communication, but as the

✉ Ludmila Lacková
ludmila.lackova@upol.cz

¹ Palacký University, Olomouc, Czech Republic

essence of life as a form of existence (Kull et al. 2011: 15; Sebeok 2001: 228; Markoš 2002). The convergence of the living and the cultural is also in an disadvantageous position for description, as Rodríguez says:

If the task of biosemiotics is the naturalisation of semiotic phenomena and the investigation of the said phenomena in the biological world, and if we are right to consider these as foundational for other semiotic activity, including the realm of culture, the way we present our biosemiotic theories should have an impact on how we develop our theories of culture. (Rodríguez 2018: 109)

The way of thinking about life as a sign phenomenon has a long tradition despite the fact that mainstream semiotics was constantly trying to avoid talking about biological phenomena in the context of the signs and semiosis:

Probably it would be prudent to say that neurophysiological and genetic phenomena are not a matter for semioticians [...] The phenomena on the lower threshold should rather be isolated as indicating the point where semiotic phenomena arise from something non semiotic, as a sort of 'missing link' between the universe of signals and the universe of signs. (Eco 1976: 21)

Modern semiotics in this context refers, for example, to Jakob von Uexküll (1982), who in his ethological approach emphasized the correspondence between the habitats of organisms and subsequently described the biosphere as a coherent and optimizing system of mutual manifestations of their life forms. He thus disregarded the material nature of the system of living organisms and based his understanding of what happens in nature on the relations of life forms. Put another way, this resembles the formulation of Ferdinand de Saussure, who simultaneously began to speak of language as a form – a system of mutual oppositions – contradictory to the way of describing language on the basis of its material nature, such as articulatory or acoustic forms. Later on, the nature of life in relation to its material background was formulated by H. Pattee (Pattee 2001; Pattee 2008). He expressed the epistemic contradiction between the material and the living, which Sharov and Vehkavaara (2015: 110) call “a normative semiotic criterion to delineate life from non-life”.

The sign character of life was most strongly underlined by the discovery of the genetic code. Due to that, the functional correlation between DNA bases and amino acids has been attributed a biological nature of life. The description of the genetic code and related processes in the body, for example, in terms of gene-driven processes, splicing, etc., has helped to anchor the widespread use of the linguistic metaphor of the living (Raible 2001; Rodríguez 2018; Jakobson 1971; Markoš and Faltýnek 2011; Lacková et al. 2017). Returning to the initial discovery of the DNA structure (Crick 1962, 1964, 1967, 1968), the basic relation between matter and function in the case of the biosphere of Uexküll or the language in the case of de Saussure lies between the medium representing DNA entry on one hand and amino acids as the building blocks of proteins on the other hand. Proteins are referred to as builders and drivers of life. Even though viewing various manifestations of life as signs is common today, there is no clear answer the question: What kind of sign is protein? Apart from this one, there are more questions to be answered, such as: Can proteins be classified among sign

categories? And: From the point of view of contemporary semiotics, are all types of proteins one type of sign?

In connection with these questions, we would like to formulate the fundamental question that we want to address in this text: Even though proteins are the cornerstone of life, its very material basis, and at the same time the bearers of the biochemical processes that establish life as a dynamic and evolving phenomenon, why do we think of them as of functional signs and not material ones? Furthermore, if we keep considering proteins to be signs, what does it imply for the understanding of life in terms of matter-function relations? In our opinion, the question *What is life as a sign?* can only be answered if we first resolve the sign classification of its components: proteins. The answer will be unsurprising, given the sign nature of life formulated in the field of biosemiotics, but it will include the role of the primitive bearer of the sign function in a living form: protein.

Proteins and Types of Signs

Let us start with a quotation by Sharov and Vehkavaara: “Because of the enormous variability of sign processes in living systems, one of the important tasks of biosemiotics is to develop criteria that help to classify sign processing various living systems” (Sharov and Vehkavaara 2015: 104).

The concern to determine the sign category of proteins responds in some way to one of the basic tasks of the biosemiotic project. Proteins are not “living systems” by themselves; yet they are the basic constituents of almost all living forms. It therefore seems important to find the criteria necessary to classify proteins as signs.

One remark should be done before we move to our argumentation. When it comes to a sign model or sign classification in the field of biosemiotics, it is Peirce’s terminology and theory which is referred to as the primary source (for an application of Peirce’s terminology in biosemiotics see for instance Favareau 2008). We will not make an exception by contributing to an already existing discourse. Nowadays, most sign typologies provided within biosemiotics deal exclusively with sign-object relations. In other words, they neglect Peirce’s original classification into 10 classes of signs based on a matrix composed by 3 trichotomies and 3 types of relations (as formulated by Peirce in the Syllabus, published as MS, 540). J. Queiroz made an important statement in this regard evidencing that “there is no way to describe these processes (semiotic processes of biological systems) with some accuracy examining only the sign-object relationship” (Queiroz 2012, 60). The omission of other types of sign relations in the field of biosemiotics might be posited as a consequence of intrinsic properties of biological systems which are in their nature less complex semiotic systems (in comparison with cultural semiosis or natural language). Another explanation of the reduction of Peirce’s model in biosemiotics, regardless of the semiotic complexity of the biological sign systems, might be the very complexity of Peirce’s sign classification (see Priscila and Queiroz 2017) and the state of art of the biosemiotics project. We are only trying to comprehend the semiotic nature of life forms, and at this state of art it might be convenient to limit the description of one exclusive part of the whole Peirce’s typology, reduced to sign-object relationship. This does not mean that there are no such relations as there are with interpretants or that there are no other aspects of Peirce’s sign

matrix. This only means that such an approach is convenient for the sake of a formal description of sign typologies in organisms.

In this paper we will describe the type of sign relation between object and representamen in proteins, thus, we are supporting the standpoint that for the sake of clarity in the very beginning of semiotic descriptions of the lowest semiotic thresholds it is convenient to focus on primarily to sign-object relation.

If life at its lowest level is treated as a sign process and if sign manifestations of life are in any way classified (according to an existing sign typology by Peirce), they are often referred to as indexes. This means that they are based on a relation where the sign refers to its object on the basis of a material or causal link (for definition of an indexical sign see Peirce CP 2.248). This does not however apply to degenerate indexes such as demonstrative pronouns since they miss the actual causal relation with the object (Goudge 1965: 56). Note that indexes as the sign types are opposed to icons and symbols. This classification is found on the relation between representamen and object, in other words, it reflects in what manner the representamen represents its object. The definition of an icon is based on the likeness of the sign and its object, the definition of a symbol is based on the conventionally accepted assignment of sign to the object, i.e. without imitation or immediate contact (Peirce CP 1.372, CP 2.92, CP 2.247–2.249). The icon is further defined by the quality of Firstness, that is, “An Icon is a Representamen whose Representative Quality is a Firstness of it as a First. That is, a quality that it has *qua* thing renders it fit to be a representamen. Thus, anything is fit to be a Substitute for anything that it is like” (Peirce CP 2.276).

Similarly, the index is associated with Secondness (CP 2. 283) and the symbol with Thirdness, respectively a rule that determines the interpretant (CP 2. 293).

First of all, let us propose an example of the indexical understanding of the living: Thomas Sebeok describes the behaviour of bacteria as indexical (Sebeok 2001: 91) – for an illustrative description of a sign description of the bacteria and its behaviour see Sharov and Vehkavaara’s protosemiotic project (Sharov and Vehkavaara 2015: 111). The behaviour is based on the direct contact of the direct contact of the bacteria, of its receptive system and the chemical attractant, chemical field in the immediate vicinity of the organism. In the sense of the aforementioned definition of an index, or in other words, based on the immediate material and causal context. Sebeok also evaluates plant indexical behaviour, discussing the vegetative level of indexicality (Sebeok 2001: 92). Finally, indexicality (DNA fingerprinting) is also referred to in the context of the internal states of an organism (as opposed to its behaviour and interaction with the external environment) – e.g. in relation to the immune system (p. 97) or to the level of genetic notation (p. 96).

Besides Sebeok, we can mention other examples of attribution of indexicality to basic living forms. For instance, plant sign behaviour studied by phytosemiotics is thought to be indexical mostly by Krampen (1981), see also Kull (2000). In addition, and similarly to Sebeok, indexical sign nature has also been attributed to the genetic code and the genetic script (Tchertov 2019). Finally, indexical sign character is also the case of the biomarkers (Boenink 2016).

In the field of ecosemiotics, Maran (2017) creates a classification of indexical environmental signs. Environmental signs are “signs understood here in a loose sense and signs that we and other animals perceive and interpret in the natural environment” (Maran 2017: 356). By environmental signs we mean bird singing, animal footprints,

visual stimuli from the environment, etc. According to Maran, all environmental signs are of indexical character. Maran concludes by introducing eight basic types of environmental signs: (1) Inaccessible sign, (2) Accessible sign, (3) Inaccessible detached sign with particular representamen, (4) Inaccessible detached sign with manifold / vague representamen, (5) Accessible detached sign with particular representamen and object, (6) Accessible detached sign with manifold / vague representamen and object, (7) Accessible detached sign with particular representamen and (8) Accessible detached sign with both general / manifold object and object.

Maran's typology of environmental signs is based on a combination of various forms of sign (representamen) and object. The interpretant is discarded in Maran's approach for the sake of simplicity of description. We can go back to Queiroz and his call for the importance of including Peirce's sign typology in biosemiotics. None of the aforementioned pioneer works in bio-sign (or environmental sign) typology really considered Queiroz's appeal, some of them rationalized this position, as it is the case of Maran. In the interest of transparency of our approach, we would like to include a standpoint on the role of the interpreter and interpretant in the proposal of sign classification within biosemiotics.

The most described signs by semiotics are cultural in nature (Nöth, Boussiac, Eco and others). Subject to cultural signs, we always assume the role of a mental interpreter, a sign user in the form of a member of society, who uses a given sign system, and also a role of mental sign interpretation. It is this mental processing that assigns the relationship between the representamen and its object, creating a sign. In this sense, we talk about the differential operation of limiting / assigning sign and object entering the sign function. In the case of life or environmental signs, we cannot speak of a mental differential operation – not even in the case of animal communication, when the mental processing of the sign is methodologically unavailable (nonassessable). In our case, we are discussing proteins, intracellular processes, and internal states of the organism, where the question of mental processing is completely excluded – unless we are considering some extreme mentalism. Of course, this does not apply for the concept of *minimal mind* by A. Sharov (2013), which is not really related to mental phenomena in the strict sense but rather to a capacity of perceiving and classifying objects.

Mental interpreter is only a medium for carrying the role of an interpretant, which is another problem to comment on in this context. We would like to take a standpoint analogic to that one by Maran, who excludes the interpretant from the description of environmental signs and their classification – because of the formal approach to classification, i.e. without the description of the relation of an interpretant to other sign constituents. We can also exclude both interpretant and interpreter from our description of the protein sign nature in order to avoid any discussion of interpretation, which has had a problematic development in the field of biosemiotics due to a different conception of interpretation as such. In any event, we want to state the following: describing a sign without including explicitly both interpretant and interpreter is possible, because interpretant does not establish the functional nature of a sign. This is the assignment of expression (sign) and content (object) that is not purely physical, see the classification of indexes, icons and symbols above, as well as the difference between physical and semiotic (Pattee 2001; Pattee 2008).

The above summarized basic (and simplified) classification of signs into icons, indexes and symbols is based on Peirce's classification of signs, and for Peirce, every

sign, also an iconic sign, is of a mental character. However, in the case of proteins and lower life forms, it is not possible to speak of sign systems or sign behaviour from a mental perspective, in regard to a mental interpretation of signs (Barbieri 2009).

At this point, we would like to mention that exclusion of interpretant and interpreter from the description of the sign of the living does not mean that we do not assume the use of the sign by an individual. The same applies to interpretation, which we understand as individual differential operation; using of sign by an individual. However, we do not work with this aspect for classification of signs. Analogically to Maran, because of the formal approach to classification in which the individual the use of sign obscures its general structural nature. In biosemiotics we find a plethora of approaches undermining the interpretative nature of the living, e.g. in the hermeneutical conception of life by Markoš (2002, 2011). Sharov expresses the role of an individual in his concept of *agency* (Sharov 2010). The character of the living as agency or interpretation is unquestionable (but see Barbieri 2009), but the very nature of this phenomenon cannot be formalized, which must be done in order to describe the structural characteristics of the signs of interest, be it environmental signs in case of Maran or proteins in our case. If we allow for individual usage of signs and interpretation in the description of sign typology, we have to operate with indexes, in the sense of degenerate index as individual reference with some own characters

An Index or Seme is a Representamen whose Representative character consists in its being an individual second. If the Secondness is an existential relation, the Index is genuine. If the Secondness is a reference, the Index is degenerate. A genuine Index and its Object must be existent individuals (whether things or facts), and its immediate Interpretant must be of the same character. But since every individual must have characters, it follows that a genuine Index may contain a Firstness, and so an Icon as a constituent part of it. Any individual is a degenerate Index of its own characters. (Peirce 1998, Syllabus: Syllabus of a course of Lectures at the Lowell Institute beginning, EP 2:274)

The previously described problem of mental phenomena involved (or not) in semiosis of the most basic life forms was affronted by the discipline of “protosemiotics”. G. Prodi (1988) was the first to come up with the term “protosemiotics”, and the concept was later elaborated by Sharov and Vehkavaara (2014). Sharov and Vehkavaara, in response to the concept of Kull’s vegetative semiosis, define protosemiosis as a sign interaction where signs interact immediately and without entering the sign function between representamen and the object (a kind of semiosis without representation). This unilateral concept of sign interaction, since it does not work with sign relation, also excludes the role of the agent’s (mental) associative functional effect in allocating sign function. In the case of protosemiotics, agency is based on the allocation of the function of physical properties in the body (Sharov and Vehkavaara 2014: 107). Therefore, in the case of protosemiotics, the division into proto-icons, proto-indexes and proto-symbols no longer bears the mental characteristics:

Though the classification is inspired by Peirce’s division of iconic, indexical, and symbolic signs, the similarity is somewhat illusory, because here we consider the type of mechanism according to which the functional effect of sign is produced,

whereas Peirce considered the relation of signs to their objects. This classification of proto-signs is an example of a meta-agent approach to protosemiosis, because proto-semiotic agents have no capacity for any kind of classification. Categories of proto-signs are discussed here in the order of increasing complexity, where the simpler signs appear to be components of more complex signs. (Sharov and Vehkavaara 2014: 116)

Complexity in relation to the sign typology of Sharov and Vehkavaara plays the following role. The authors talk about proto-icons in the case of interaction of two physical objects in one and the same action. In the case of interaction between two and more objects (functional subunits) in different actions, they speak about proto-indexes. The more complex interactions are of symbolic nature.

Arguments for Iconicity

Despite Sebeok's and others' suggestions to consider the lowest forms of life (bacteria, DNA fingerprinting, etc.) to be indexical signs, there is a possibility – and we would like to give arguments to support this line of thinking – that these life forms would be iconic in nature. Our reasoning is based on the classification of proto-signs proposed by Sharov and Vehkavaara (2014) in terms of the classification into proto-icons, proto-indexes and proto-symbols based on the degree of complexity and thus on the basis of the categories of Firstness, Secondness and Thirdness. It is obvious why Sebeok proposed to define the behaviour of bacteria (and other low life forms or components) indexically: the definition of indexical relation between a representamen and its object on the basis of a direct physical relationship, or contiguity applied to the lowest forms of life and stereochemical bonds within DNA molecules is appealing. However, we must realize that this definition of an index refers to signs that are mental in nature, and this allows the relationship between sign and its object to be attributed in the interpreter's mind. In the words of T. Vehkavaara,

a mere causal connection between two things is not a sufficient condition to make the one a potential indexical sign of the other, but the causal relation must to be somehow beforehand cognizable by the interpretive mind. (Vehkavaara 2006: 294)

Naturally, the condition of an interpretative mind applies to both indexical and iconic signs. Therefore, in the case of *proto-signs*, it is not appropriate to consider the trichotomy *icon-index-symbol* of sign classification, which already seems to be predisposed to the mental interpretation and to higher levels of semiosis. It is therefore necessary to turn to the classification of proto-signs, which does not require mental representation and is contrarily based on the degree of complexity and this leads us to the concept of icons / proto-icons (and not hypoicons, CP 2.277). U. Eco came to similar conclusions more than 20 years ago and solved the problem of the sign character of lower life forms (where it is not possible to consider mental representation / interpretation) by introducing the concept of “primary iconism”. (Eco 1999: 106–121) claims that primary iconism is an immediate link between an object and a representamen without claiming a mental or cognitive interpretation; natural primary iconism is a term with which Eco refers to Prodi's protosemiotics.

Eco similarly solved the Firstness problem within primary iconism. According to Peirce, the quality of Firstness cannot exist independently of Secondness, in other words, the monad arises only from being separated from the dyad. This creates a problem with the primary iconism that assumes the quality of Firstness as the forerunner of Secondness: lower life forms exist without Secondness, without the actual need for a causal relationship from which they would be deducible. Eco solved this problem as follows (and thus “heretically” deviates from Peirce’s doctrine, more about this divergence in e. g. Paolucci 2015): for Eco, a primary icon is something like a “predisposition to correspond (Eco 1999: 110). In the case of life forms, Eco speaks about *natural primary iconism*. Eco’s favourite example of the natural primary iconism, complementarity between triplets and amino acids, works similarly to a key fitting into a keyhole: a key and a keyhole have a complementary shape fitting each other, as the concave space fits into the convex one. In the case of primary iconism, this fit is based on a mere predisposition of correspondence, hence Firstness: it is a concave space that is ready to fit into the convex space. Eco’s characteristic of primary iconism goes in the following spirit of Peirce’s definition of the icon as a sign that does not *act as* a sign.

An Icon is a sign which refers to the Object that it denotes merely by virtue of characters of its own, and which it possesses, just the same, whether any such Object actually exists or not. It is true that unless there really is such an Object, the Icon does not act as a sign; but this has nothing to do with its character as a sign. (Peirce CP 2.247)

Eco explains natural primary iconism with an example of the transfer of genetic information. Of course, in the case of stereochemical correspondence between triplets, or between an enzyme and its substrate, or between an immune system cell and the molecule with which it interacts, the “object” is physically present there, but this object is not recognized as an Object of sign – hence Eco’s “mere predisposition of correspondence” is a predisposition of correspondence in terms of a sign relation, chemically and physically the correspondence is of course real. Kull also came to a similar conclusion, proposing the term *vegetative semiosis*:

Calling vegetative semiosis iconic may seem, at first glance, weird, because no spatial relations can yet be recognized in this sphere. However, and this is exactly the point of the argument, in all cases when, for instance, a (two-dimensional) picture or just a spatial relation serves as a sign (i.e., a synchronic relation between two or more separate points in the space has to be recognized – not stereochemically, of course, but on the basis of code-processes), an indexical relation will be necessary. Vegetative semiosis being solely iconic, means that it just recognizes (and transforms) something; it cannot at the same time relate the object in question to the one that is left or right of it. Nevertheless, this is a true relation, because it is code-based (thus memory-based that allows similarity to be established) and not a pure stereochemical interaction. Recognition of correlations or linkages (of ordered relationships) would already mean a truly indexical relation. This will require an ability to associate what has not been associated earlier, which is yet a feature of a more complex – animal – semiosis. (Kull 2009: 15)

In this manner, Kull differentiates between vegetative semiosis (related to the concept of *recognition* and based on iconic relations); animal semiosis (related to *association* and based on indexical relations; and cultural semiosis (related to *combination* and based on symbolic relations). Elsewhere, Kull clearly distinguishes between natural and cultural semiosis (Kull 1998: 350). Kull speaks about *secondary threshold zones* being an elaboration of Eco's notion of a semiotic threshold. An indexical threshold zone is one that separates vegetative semiosis from animal semiosis. The indexical threshold zone lies in the capacity of *association* which is lacking in the vegetative semiosis.

Most organisms are supplied with many functional cycles that enable vegetative relations. These are generally responsible for categorization and speciation, and for the simple forms of search, finding, selection, swarming, spreading, etc. Vegetative relations are just correspondences, or relations of pure recognition only, however memory-based (because any semiosis assumes some sort of memory), which should mean that these are exclusively iconic. (Kull 2009: 20)

The previously commented arguments by Eco and Kull and the concept of protosemiotics lead us from Sebeok's initial characterization of proteins as indexes to their iconic nature. This is in correspondence with the last concept we want to mention. Kull expresses the differences of vegetative, animal and cultural semiosis (Kull 2009). Kull classifies all types of semiosis which do not involve the participation of the central nervous system, as iconic. In his point of view, signs do not act indexically until they are cognitively processed. Therefore, in summary, based on Eco and Kull, we can conclude a definition of an icon in the context of lower life forms. This definition would have characteristics of recognition of something as an object without association. In addition, the recognition is in most cases accompanied by a transformation (cellular processes).

We consider the aforementioned arguments to be essential for sign definition of proteins. However, we think it is necessary to base our definition on other arguments that would clearly express the relationship between the sign and the object that determines such signs. The above-mentioned approach by Sharov and Vehkavaara classifies signs on the basis of their relation to agency and ignores the internal sign structure. Since we do not consider the role of interpretation and agency to be decisive for the *formal* description of the sign types (although we state that the role of agency is very important for a semiotic definition of the living), we will continue to look for reasons for the formal description of the sign type in the relation between representamen and object. In this manner, we intend to provide a different approach to classification of protosigns.

Next, we shall take a look at what a protein is, what form it takes. Based on the following description, we can also consider how proteins correspond to the index classification conditions.

Description of Protein Structure

Proteins are chemical compounds belonging to the group of biopolymers, substances composed of one or more repeating parts, which are structurally simpler. Compositional

parts of proteins are amino acids, molecules consisting of a carboxyl group and an amino group. Through these groups, the amino acids bind in the peptide chain. When describing a protein, the amino acid chain is referred to as the primary protein structure. In terms of the primary structure of a protein, we consider its composition as a sequence of amino acids in a linear chain.

Thanks to bonds between amino acids that are not contiguous in the primary structure of the protein, that is, which are distant from one another in the linear chain, spatial formations are created. The spatial formations formed from the amino acid chain are of two basic kinds: alpha-helices are in the form of a spiral and beta-sheets take a laminar zigzag shape (zigzag shape is due to the linkage between amino acids having the opposite order in the primary structure). The parts of the protein located between these formations are called turns. Turns are in the form of chains that link the spatially ordered peptide chains in the form of alpha-helices and beta-sheets. When describing a protein, its composition of alpha-helices and beta-sheets is referred to as a secondary structure. The tertiary structure of a protein is usually mentioned in connection with the spatial formations that interact with other chemical compounds in the surrounding of the protein, such as other proteins, transported substances, chemicals whose reaction catalyses proteins, etc. Domains are further compositional protein parts. A domain is a cohesive part of the protein that is independent from the function of tertiary structure. This means that the domain participates, together with other parts of the protein, in its interaction with the environment, or provides the interaction by itself. These domains are found in a more or less altered form, in different proteins, with different functions. It is domains that are crucial in studying the structure and function of protein in the body. If the protein composition is more complex in terms of its interaction with the environment, it is referred to in the context of quaternary structure.

The Semiotic Nature of Proteins

Let us note a few moments that must be emphasized in relation to proteins, both in terms of the physical and the semiotic.

Firstly, we would like to observe the relation between the linear and spatial nature of the protein structure: the peptide chain has the nature of sequential units, amino acids, which also form a spatial arrangement. The function of a protein in the form of a linear sequence cannot be considered. In general, we think of proteins as of physical objects, not as of signs.

Secondly, the linear amino acid chain is directly related to the DNA (mRNA) transcription, meaning that the linear peptide chain is a circumscription of genetic information and carries its syntactic features. In the respect of a correspondence between DNA transcription and chains from amino acids, the term genetic code is used, genetic code being not the only code in the living, within biosemiotics a plurality of codes is referred to (Barbieri 2002; Trifonov 1988). The crucial message of this point is that the primary structure of the protein is semiotic in its nature, it is a text, or better its transcription/translation, in which the basic role is carried by relations of units in a chain (or in a text if we use the transcriptions from protein or genetic banks). The relationship between the linear nature of the

peptide chain and the DNA structure is the primary component of life and all the relations between the two structures at this level fall within the level of sign foundations of relations in the organism. At this level we are talking about the way life associates its sign (semiotic) level with its physical nature.

Thirdly, a linear chain of amino acids forms elementary spatial relations – secondary structures. It is interesting to note that the transformation of a linear chain into a spatial formation is subject to the classification of basic spatial formations. In other words, it is interesting to note that there is a small number of certain types of such spatial arrangements. The reducibility to a very small number of basic structures is an essential feature of the secondary structures (alpha-helices and beta-sheets). Another essential feature of alpha-helices and beta-sheets resides in that they are mutually composed into hierarchically higher structures. Only with these composed structures can we identify the function of a protein (transport, constructive, receptive, etc.). Keep in mind that proteins consist of lower spatial formations encoded by the primary structure or genetic notation in form of a chain, which has no direct physical interactivity in relation to protein function. Only a combination of the basic shapes provides for processes that we call functions of protein or functions of organ systems of an organism. Initial spatial formations, which are formed from linear physical form of the chain, are compositional variables and play the role of components in the protein structure, providing the desired form by combining the protein, their primary function is not based on their direct physical interactivity.

Fourthly, proteins provide function by combining elementary spatial conformations into a higher conformation that is ready for physical interactions in order to provide the necessary function. That is, proteins interact by shape. Let us recall that these shapes can be combined from more basic units, that different conformations provide the same function. Talking about shape means that proteins interact with the environment in a physically specific way. By shape, proteins mediate a certain type of interaction that could also occur spontaneously, but autonomously from protein shape these interactions and would not have a relation to the rest of the processes in the organism, would not have a (semiotic) function.

Fifthly, the relation between the shape of the protein and the interaction it performs in the organism has several levels of representation in relation to the DNA transcript. Lower units of secondary structures are compositionally derived from an amino acid chain which in its turn is composed of amino acids, these being an indirect translation of the script of the genetic information in the DNA (RNA). Both levels of representation of the genetic information (secondary structures and amino acids) represent different levels of arbitrariness. The shape of secondary structures is arbitrary in relation to protein function, and the string of amino acid is arbitrary in relation to the DNA script (this arbitrariness is guaranteed by the genetic code).

The five aforementioned points about the semiotic nature of proteins can be comprehended as criteria for a semiotic classification of proteins, which can be summarized accordingly:

1. We differentiate between physical and semiotic, which equals in the case of proteins the distinction between spatial (object like) and linear (sequence like). In the case of proteins, the first important criterion for a semiotic nature lies in the moment when physical becomes semiotic.

2. Linearity is related to syntacticity, physical conditions in linearity correspond to syntax of a code. The linearity of the organic strings guarantees their code-like nature.
3. Physical compositionality is another important criterion. The transition from physical to semiotic is based on the construction of elements (objects) that compose the physical form performing the function.
4. Shape mediating function is responsible for semiotic behaviour of proteins. The function is not directly related to the physical interactivity of protein components, but rather to the final interacting shape.
5. The fifth point refers to the concept of double arbitrariness that is reflected in physical compositionality of proteins in smaller units of double level (secondary structures and amino acids).

The characterization of proteins described above has some divergences from the concept of the proto-signs of Sharov and Vehkavaara (2015), to which we refer primarily, but it also differs from the other characterizations of the sign nature of the internal states of the organism. The prerequisite for Sharov-Vehkavaara's model is agency. With the condition of agency, in theory, all physical phenomena in the body can be described as signs, because they are somehow part of the processes in the living forms. On the basis of the normativity condition, it is possible to assign a sign role to all the physical components of a cell or an organism. This step of Sharov and Vehkavaara allows signs to be characterized in living forms as signs without representation – which is the basic design feature of the proto-sign by Sharov and Vehkavaara. Accordingly, they classify proto-signs as icons, indexes, and symbols based on the simple or multiple physical interactions of objects entering the sign. The decision to define proto-signs on the basis of increasing level of complexity of physical interactions could probably be analogized to the increasing arbitrariness and conventionality of the trichotomy “icon, index, symbol” in natural languages. However, we believe that the basic epistemic condition of representation is violated. We are referring to the von Neumann argument to which Pattee (Pattee 2001, Pattee 2008; see de Beule 2012; Waters 2012) responds and in which case representation is an epistemic condition of separation of semiotic from physical. We can reformulate von Neumann's argument as follows: the construction of von Neumann's probe must contain a symbolic description of the probe. Involvement of the physical body of the probe and its description parts is defined by semiotics as representation.

Representation in protosigns was a rather necessary condition also for Prodi to which Cimatti (2000) refers and then formulates the definition of representation at the level of protosemios as physical interaction in which *one physical object signifies something for another physical object* simply by interacting with it, yet it is the interaction itself related to a scope, to a function of the considered object to guarantee representation (sign-object relation). Similarly, Kull speaks about the difference between physical and semiotic in connection to an enzyme and its target molecule within the frame of vegetative (iconic) semiosis: “A is not only a molecule with its chemical relationships, but A becomes a sign-vehicle, a signal, due to its inclusion in a code-relation” (Kull 2009: 19). Elsewhere, Kull differentiates between the concept of recognition and reaction in enzymes, while the recognition is based on sign property, to signification (Kull 2005: 22).

Findings

Our approach to the problem of sign representation in living forms is linguistic. It means that expression and content are involved in the description of the notion of sign of the natural language (e.g. Hjelmslev 1943; de Saussure 1916) and the interpretative aspect is neglected due to the social abstraction of language norms. In other words, the interpretive nature of the language sign does not affect the direct relation between the language form (words, morphemes and so on) and its meaning (lexical meaning, grammatical meaning). Nevertheless, this does not mean that the linguistic method implies studying biological phenomena only at the level where general linguistics and ethology intersect and lead to nothing more than “just one more area where a semiotic description can be applied, as to chemistry, or geology, or geography, or, for that matter, hydrology” (Kull 2005: 16). Our method of sign determination is based on code design. We consider the concept of code, together with representation, as the preliminary condition for establishing the description of signs and their systems. The linguistic method encourages clarification of the conditions of representation and fulfilment of the code function. Differently from Sharov and Vehkavaara we do not find sign descriptions in organisms without representation satisfactory enough since there is a risk of assigning a sign typology according to purely physical qualities (complexity level of physical interactions). We consider the cell-level signs to be representational in the form of an iconic nature – shaping the reactants that ensure the processes in living forms. These shapes are proteins. In the sense of Queiroz (according to Peirce’s cross-relational sign classes), we can classify them as icons, due to their hierarchical composition.

Additionally, to support our argument, we can refer to the concepts of involvement and instantiation of sign qualities. The instantiation rule, according to Liszka (1996) says that a sign in order to be a sign must be instantiated in some *sinsign*. In other words, a *qualisign* cannot act as sign on its own. Higher phenomenological statuses always include lower ones: *sinsigns* include *qualisigns*, *legisigns* always includes *sinsigns*, *indices* involve *icons* and *symbols* include *indices*. We have stated in the first part of this paper that we won’t consider the relation between sign and its interpretant, following Maran, therefore we limit ourselves, in our classification, to the first and second trichotomy from the 1903 *Syllabus* matrix, that is, the trichotomy of the nature of sign (*qualisign*, *sinsign*, *legisign*) and the trichotomy of sign in relation to its object (*icon*, *index*, *symbol*). Now, if we exclude the association from our definition of protein proto-icon, we are referring to the missing connection to interpretant, but at the same time, by refusing the association we do not exclude the Thirdness from our protosemiotic classification matrix. The Thirdness is guaranteed by transformation (recognition of something accompanied by a transformation) which necessarily refers to some future state of the cell/organism, therefore, it is related to Thirdness. Consider the arbitrary nature of the genetic code and the historical and conventional aspect of evolution. Of course, arbitrariness, conventionality and historicity of proteins are qualities supporting the semiotic nature of the living and are apparently characterised by some level of Thirdness. The understanding of these semiotic qualities relies upon a third part, that is, a human mind which understands these qualities as Thirdness. Yet from the low-level-semiosis viewpoint the arbitrariness and conventionality are gratuitous (Markoš 2008), meaning that the agents of protosemiosis are not aware about

these qualities and are unimportant for the metabolic functions of proto-agents. Therefore we decided to exclude them from our protosemiotic protein sign classification. Consequently, we have at our disposition a reduced matrix of trichotomies to classify proto-signs. It could be schematized graphically as follows:

qualisign	sinsign	legisign
icon	index	symbol

According to what has been said above, proteins would be, by means of our reduced matrix, iconic sinsigns (sinsigns because of the rule of instantiation of qualisigns in sinsigns, in the case of proteins it is the instantiation into material embodiment in folded strings).

Last but not least, we should also consider the quality of icons to resemble the objects they refer to. Proteins interact with the target molecules by shape and this shape matching is given thanks to a kind of resemblance which might be defined as a kind of complementarity of shapes. The relation of resemblance lies somewhere between genuine and degenerate Secondness (Kruse 1991) and this is how we also comprehend proteins.

Resemblance, then, is either a form of genuine Secondness or else it is less degenerate than the relation of identity. In either case, we must place it at the threshold between degeneracy and genuineness [...] Resemblances are borderline relations that cannot be classified unequivocally as either genuine or degenerate Secondness. They are more genuine than identities, but they are more degenerate than reactions. (Kruse 1991: 276-277)

As we have pointed out, the necessary condition for the construction of a complex code, such as natural language, is double arbitrariness (point 5 in semiotic criteria). The function of proteins can be described metaphorically as blueprinting/modelling of the amorphous unsegmented physical world. Proteins handle their specific physical conditions by shaping them: the physical interaction of the protein is mediated by its shape (and not by stereochemical interactions) which means that it is not completely physically directed. This is the reason we treat proteins as signs. Proteins should be classified by standard sign classification. First of all, we should assign proteins the quality of symbolicity, indexicality or iconicity. If we base the definition of an icon either on physical resemblance or on a lack of conventionality, we find that both of these definitions are suitable in the case of proteins (as signs they fulfil the role of physical resemblance and have a weaker role of conventionality in the form of normativity condition and agency) (Sharov and Verkhavara). These features of proteins stress our decision to refer to proteins as iconic signs. Eco says that we need to define “the point where semiotic phenomena arise from something non-semiotic” (Eco 1976). In the case of proteins, we are talking about crossing the boundary between physical and semiotic, which is the point of the epistemic cut between these two spheres. Proteins come from a linear string which is coded by a genetic code and stored in DNA – this is the informatic and semiotic part of their role in organisms. The linear string is shaped into an active physical object which constructs and drives the organism – this is the physical role of proteins. From a semiotic point of view, the most interesting is the transition between linear/semiotic and physical. In order to transcend

from physical to semiotic, some passage must be done, which is close enough to physical yet is already starting to be semiotic. It is very likely that this passage is done similarly to physical conditions, that is, by their imitation. And imitation is nothing else than an icon.

Discussion

Throughout the history of linguistics, there has been an ongoing discussion on the functions of language, which were, in the basic way of division, divided into a communicative function and a reflective (modelling) function – representing the external world. The communication function was based on the transmission of information between communicants (Shannon 1948) and, above all, on sharing interpersonal experiences (Itkonen 2003). Language is ultimately understood by structural, generative and other branches of linguistics as an intersubjective norm. The distinction between *langue* and *parole*, linguistic competence and linguistic performance reflects the assumption of linguistic normativity based on the communicative function of language (de Saussure 1916; Chomsky 1957: 38). When expressive, poetic or persuasive functions were later added to language functions list, the conception of language as a means of contact between members of the community was strengthened. For example, Jakobson (1960) distinguishes a number of other functions of a communicative nature contrasting the reflective (modelling) function.

When we flip through the book by Berwick and Chomsky (2016), *Why only us*, we can formulate the basic conclusions that the authors draw from the synthesis of palaeontology, neuroscience, cognitive sciences, computation sciences and linguistics. The authors evaluate the conclusions of contemporary knowledge of language faculty as follows: natural language is primarily a language for thought, it does not primarily serve communication, its primary function is to reflect the states of the outside world, to model it. From a cognitive point of view, language is associated with the ability to orientate in the external environment, its representation serving the beneficial behaviour of the organism (we naturally speak of higher organisms – see the study on the language gene FOX P2). Berwick and Chomsky therefore prioritize the reflective (modelling) function of language.

But what is the position of an organic semiosis with regard to the two types of functions? To answer this question, we use the semiotic concept of *umwelt* (von Uexküll 1982). In terms of the opposition between communication and modelling, *umwelt* is a means for both, but its basic building prerequisite is the representation of the outside world by an organism. *Umwelt* works as the receptive predisposition of an organism. Its interactivity with the selected stimuli from the environment is a means of representation in terms of the semiotic instrumentation of the organism. In our case, we are discussing the semiotic nature of intracellular processes. The foregoing description of proteins has shown that they are based on mediating linear nature, the construction of spatial building components and higher spatial units having a function in the body. We consider the relationship of the shape representation of the outer world as an iconic sign function by which the organism represents the amorphous physical reality. It anticipates it, assumes it, and in some cases, we can talk about deception (but probably not lying). The semiosis of life forms is primarily modelling, secondarily communicative. We

would like to express it in such a way that the living uses its sign character to touch the outside world in the form of a shape representation. And we cannot conceive of it in any other way than iconically. Proteins are icons.

Acknowledgments The author Dan Faltýnek was supported by a funding from the project *Sinophone Borderlands: Interaction at the Edges*, n. CZ.02.1.01/0.0/0.0/16_019/0000791. The author Ludmila Lacková was supported by a funding from the The Ministry of Education, Youth and Sports of Czech Republic (IGA_FF_2020_006), name of the financed project: *Aplikace metod Digital Humanities v interdisciplinárním výzkumu*.

Compliance with Ethical Standards

Conflict of Interest The authors declare that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Barbieri, M. (2002). *The organic codes: An introduction to semantic biology*. Cambridge: Cambridge University Press.
- Barbieri, M. (2009). Three types of Semiosis. *Biosemiotics*, 2(2), 19–30.
- Berwick, R. C., & Chomsky, N. (2016). *Why only us? Language and evolution*. Cambridge: MIT Press.
- Boeninck, M. (2016). Disease in the era of genomic and molecular medicine. In J. A. Marcum (Ed.), *The Bloomsbury companion to contemporary philosophy of medicine* (pp. 65–91). London: Bloomsbury Publishing.
- Beule, J. de (2012). Von Neumann's legacy for a scientific biosemiotics. *Biosemiotics*, 5, 1 1–4.
- Chomsky, N. (1957). *Syntactic structures*. The Hague/Paris: Mouton.
- Cimatti, F. (2000). *Nel segno del cerchio. L'ontologia semiotica di Giorgio Prodi*. Roma: Manifestolibri.
- Crick, F. (1962). Towards the genetic code. *Scientific American*, 207(3), 8–16.
- Crick, F. H. C. (1964). On the genetic code, Nobel lecture, December 11, 1962. Nobel lectures: Physiology or medicine 1942–1962, Elsevier: Amsterdam.
- Crick, F. H. C. (1967). The Croonian Lecture, 1966. The genetic code. *Proceedings of the Royal Society B*, 167(1009), 331–347.
- Crick, F. H. C. (1968). The origin of the genetic code. *Journal of Molecular Biology*, 38, 367–379.
- de Saussure, F. 2011 (1916): *Course in general linguistics*, trans. By W. Baskin, P. Meisel, and H. Saussy, Columbia University press.
- Eco, U. (1976). *A theory of semiotics*. Bloomington and London: Indiana University Press.
- Eco, U. (1999). *Kant and the Platypus: Essays on language and cognition*. New York: Harcourt Brace and Company.
- Favareau, D. (2008). Iconic, indexical, and symbolic understanding. *Journal of the American Psychoanalytic Association*, 56(3), 789–801.
- Goudge, T. A. (1965). Peirce's index. *Transactions of the Charles S. Peirce Society*, 1(2), 52–70.
- Hjelmslev, L. (1969 [1943]). *Prolegomena to a theory of language*, Madison, University of Wisconsin.
- Itkonen, E. (2003). What is language? A study in the philosophy of linguistics. In *Publications in general linguistics*, 8. Åbo/Turku: Åbo Akademi tryckeri.

- Jakobson, R. (1960). Linguistics and poetics. In: T. Sebeok (Ed.), *Style in language* (pp. 350–377). Cambridge: Massachusetts Institute of Technology Press.
- Jakobson, R. (1971). Linguistics in relation to other sciences. In R. Jakobson (Ed.), *Selected writings, volume 2: Word and language* (pp. 678–687). Mouton: The Hague.
- Krampen, M. (1981). Phytosemiotics. *Semiotica*, 36(3–4), 187–209.
- Kruse, F. E. (1991). Genuineness and degeneracy in Peirce's categories. *Transactions of the Charles S. Peirce Society*, 27(3), 267–298.
- Kull, K. (1998). Semiotic ecology: Different natures in the semiosphere. *Sign Systems Studies*, 26(1), 344–371.
- Kull, K. (2000). An introduction to phytosemiotics: Semiotic botany and vegetative sign systems. *Sign Systems Studies*, 28, 326–350.
- Kull, K. (2005). Semiotics is a theory of life. In: Williamson, Rodney; Sbrocchi, Leonard G.; Deely, John (Eds.), *Semiotics 2003: Semiotics and National Identity*. New York, Ottawa, Toronto: Legas, 15–31.
- Kull, K. (2009). Vegetative, animal, and cultural semiosis: The semiotic threshold zones. *Cognitive Semiotics*, 4, 8–27.
- Kull, K., Emmeche, C., Hoffmeyer, J., (2011). Why biosemiotics? An introduction to our view on the biology of life itself. In: Emmeche, Claus; Kull, Kalevi (Eds.), *towards a semiotic biology: Life is the action of signs*. London: Imperial college press, 1–21.
- Lacková, L., Faltýnek, D., & Matlach, V. (2017). Arbitrariness is not enough. *Theory in biosciences*, 136(3–4), 187–191 springer. <https://doi.org/10.1007/s12064-017-0246-1>.
- Liszka, J. J. (1996). *A general introduction to the semiotic of Charles Sanders Peirce*. Bloomington and Indianapolis: Indiana University Press.
- Markoš, A. (2002). *Readers of the book of life: Contextualizing developmental evolutionary biology*. Oxford and New York: Oxford University Press.
- Markoš, A. (2008). *Náhoda a nutnost: Jacques Monod v zrcadle naší doby*. Cervený Kostelec: Pavel Mervart.
- Markoš, A. (2011). Hermeneutics by the living. *Biosemiotics*, 4, 119–125. <https://doi.org/10.1007/s12304-010-9096-7>.
- Markoš, A., & Faltýnek, D. (2011). Language metaphors of life. *Biosemiotics*, 4(2), 171–200.
- Maran, T. (2017). On the diversity of environmental signs: A typological approach. *Biosemiotics*, 10(3), 355–368.
- Paolucci, C. (2015). Iconismo primario e gnosologia semiotica. *VS*, 120, 135–150.
- Pattee, H. H. (2001). The physics of symbols: Bridging the epistemic cut. *Biosystems*, 60, 5–21.
- Pattee, H. H. (2008). Physical and functional conditions for symbols, codes, and languages. *Biosemiotics*, 1, 147–168.
- Peirce, C. S. (1998). The essential peirce. In: Peirce Edition Project (Eds.), *Selected philosophical writings, Volume 2 (1893–1913)*. Bloomington: Indiana University Press.
- Priscila, F., & Queiroz, J. (2017). Visualizing triadic relations: Diagrams for Charles S. Peirce's classifications of signs: A way of thinking & approaching information design. *Information Design Journal*, 23(2), 127–147. <https://doi.org/10.1075/idj.23.2.03far>.
- Prodi, G. (1988). Material bases of signification. *Semiotica*, 69(3/4), 191–241.
- Raible, W. (2001). Linguistics and genetics: Systematic parallels. In M. Haspelmath, E. König, W. Oesterreicher, & W. Raible (Eds.), *language typology and language universals: An international handbook* (pp. 103–123). Berlin — New York: Walter De Gruyter.
- Rodríguez, H. J. C. (2018). Productive perils: On metaphor as a theorybuilding device. *Linguistic Frontiers*, 1(2), 102–111. <https://doi.org/10.2478/lf-2018-0010>.
- Queiroz, J. (2012). Peirce's ten classes of signs: Modeling biosemiotic processes and systems. In Timo Maran, Kati Lindström, Riin Magnus and Morten Tønnessen (Eds.) *biosemiotics turning wild. Essays in honour of Kalevi Kull on the occasion of his 60th birthday*, 55–62.
- Saussure, F. de (2005[1916]). *Cours de linguistique Générale*. Edition critique de Tullio de Mauro, Payot, Paris.
- Sebeok, T. A. (2001). *Signs: An introduction to semiotics* (2nd. ed.). Toronto: University of Toronto Press.
- Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, 27(379–423), 623–656.
- Sharov, A. A. (2010). Functional information: Towards synthesis of biosemiotics and cybernetics. *Entropy*, 12(5), 1050–1070. <https://doi.org/10.3390/e12051050>.
- Sharov, A. A. (2013). Minimal mind. In L. Swan (Ed.), *Origins of mind* (pp. 343–360). Dordrecht: Springer.
- Sharov, A. A., Vehkavaara, T. (2014). Protosemiosis: agency with reduced representation capacity. *Biosemiotics*, 8(1), 103–123.
- Sharov, A. A., & Vehkavaara, T. (2015). Protosemiosis: Agency with reduced representation capacity. *Biosemiotics*, 8(1), 103–123. <https://doi.org/10.1007/s12304-014-9219-7>.

- Tchertov, L. (2019). Signs, codes, spaces, and arts: Papers on general and spatial semiotics. Cambridge University Press.
- Trifonov, E. N. (1988). Codes of nucleotide sequences. *Mathematical Biosciences*, 90(1–2), 507–517.
- von Uexküll, J. (1982). The theory of meaning, Semiotica. *The Hague: Mouton*, 42(1), 25–82.
- Vehkavaara, T. (2006). Limitations on applying Peircean semeiotic biosemiotics as applied objective ethics and esthetics rather than semeiotic. *Journal of biosemiotics* 1(1), 269–308. Nova Science Publishers.
- Waters, P. D. (2012). Von Neumann's theory of self-reproducing automata: A useful framework for biosemiotics? *Biosemiotics*, 5(1), 5–15. <https://doi.org/10.1007/s12304-011-9127-z>.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.