КОГНИТИВНОЕ ПРОСТРАНСТВО

Философия науки

DOI: 10.30727/0235-1188-2020-63-2-46-62
Original research paper
Оригинальная исследовательская статья

Does a Ribosome Really Read? On the Cognitive Roots and Heuristic Value of Linguistic Metaphors in Molecular Genetics[*]

Part 2[**]

S.T. Zolyan
Immanuel Kant Baltic Federal University, Kaliningrad, Russia
Institute of Scientific Information on Social Sciences of the Russian Academy of Sciences, Moscow, Russia
Institute of Philosophy, Sociology and Law, National Academy of Sciences of the Republic of Armenia, Yerevan, Armenia

Abstract

We discuss the role of linguistic metaphors as a cognitive frame for the understanding of genetic information processing. The essential similarity between language and genetic information processing has been recognized since the very beginning, and many prominent scholars have noted the possibility of considering genes and genomes as texts or languages. Most of the core terms in molecular biology are based on linguistic metaphors. The processing of genetic information is understood as some operations on text – writing, reading and editing and their specification (encoding/decoding, proofreading, transcription, translation, reading frame). The concept of gene reading can be traced from the archaic idea of the equation of Life and Nature with the Book. Thus, the genetics itself can be metaphorically represented as some operations on text (deciphering, understanding, code-

---

[*] This research is supported by the grant from the Russian Science Foundation (RSF), project no. 17-18-01536 "Knowledge transfer and convergence of methodological practices: Cases of interdisciplinary integration of political, biological, and linguistic research" at the Institute of Scientific Information on Social Sciences of the Russian Academy of Sciences.

[**] For part I, see: *Russ. J. Philos. Sci.* 2020. Vol. 63, no. 1, pp. 101-115.

breaking, transcribing, editing, etc.), which are performed by scientists. At the same time linguistic metaphors portrayed gene entities also as having the ability of reading. In the case of such "bio-reading" some essential features similar to the processes of human reading can be revealed: this is an ability to identify the biochemical sequences based on their function in an abstract system and distinguish between type and its contextual tokens of the same type. Metaphors seem to be an effective instrument for representation, as they make possible a two-dimensional description: biochemical by its experimental empirical results and textual based on the cognitive models of comprehension. In addition to their heuristic value, linguistic metaphors are based on the essential characteristics of genetic information derived from its dual nature: biochemical by its substance, textual (or quasi-textual) by its formal organization. It can be concluded that linguistic metaphors denoting biochemical objects and processes seem to be a method of description and explanation of these heterogeneous properties.

Keywords: epistemology, philosophy of science, genetic code, genetic reading, genetic translation, genetic information, quasi-mind, cognitive metaphor.

Suren T. Zolyan – D.Sc. in Philology, Professor, Visiting Professor of the Immanuel Kant Baltic Federal University; of the Republic of Armenia; Research Project Member, Institute of Scientific Information on Social Sciences, Russian Academy of Sciences; Leading Research Fellow, Institute of Philosophy, Sociology and Law, National Academy of Sciences of the Republic of Armenia.

surenzolyan@gmail.com

http://orcid.org/0000-0002-4422-5792

For citation: Zolyan S.T. (2020) Does a Ribosome Really Read? On the Cognitive Roots and Heuristic Value of Linguistic Metaphors in Molecular Genetics (Part 2). *Russian Journal of Philosophical Sciences = Filosofskie nauki.* Vol. 63, no. 2, pp. 46–62. DOI: 10.30727/0235-1188-2020-63-2-46-62

Может ли рибосома «читать»? О когнитивной основе и эвристической значимости лингвистических метафор в молекулярной генетике[*]

Часть 2[**]

С.Т. Золян

Балтийский федеральный университет имени Иммануила Канта, Калининград, Россия

---

[*] Исследование выполнено за счет гранта Российского научного фонда. Проект № 17-18-01536 «Трансфер знаний и конвергенция методологических традиций» в Институте научной информации по общественным наукам РАН.

[**] Часть 1 см.: ФН. 2020. Т. 63. № 1. С. 101–115.

Институт научной информации по общественным наукам РАН, Москва, Россия
Институт философии, социологии и права Национальной академии наук Республики Армения, Ереван, Армения

Аннотация

В статье обсуждается роль лингвистических метафор как когнитивной модели концептуализации процесса обработки генетической информации. Сущностное сходство между языком и обработкой генетической информации было осознано уже со времени становления генетики, многие выдающиеся исследователи отмечали возможность рассматривать гены и геномы как тексты или языки. Большинство базовых терминов в молекулярной биологии основаны на лингвистических метафорах. Процессы обработки генетической информации осмысляются как определенные операции с текстом: запись, чтение, редактирование и их разновидности (кодирование/декодирование, корректура, транскрипция, перевод, считывание). Истоки концепта генетического чтения можно проследить начиная с древних представлений о Жизни и Природе как о Книге. Сама генетика может быть метафорически представлена в качестве выполняемых учеными определенных операций над текстом (декодирование, понимание, взлом кода, транскрибирование, редактирование и т.д.). В то же время лингвистические метафоры представляют гены как объекты, обладающие способностью читать. В случае с «биологическим чтением» могут быть выявлены некоторые существенные черты, сходные с процессами чтения человеком: это способность идентифицировать биохимические последовательности на основе их функции в абстрактной системе и опознавать типы и их контекстуальные варианты. Метафоры видятся как эффективный инструмент репрезентации, поскольку они создают возможность двумерного описания: биохимического по своим экспериментальным эмпирическим результатам и текстуального по когнитивным моделям его осмысления. В дополнение к их эвристической значимости, лингвистические метафоры основаны на базовых свойствах генетической информации, двойственной по своей природе: биохимической по материальной субстанции, текстуальной (или квази-текстуальной) по форме организации. Отсюда делается вывод, что лингвистические метафоры, обозначающие биохимические объекты и процессы, представляются адекватным методом описания и объяснения этих гетерогенных качеств.

Ключевые слова: эпистемология, философия науки, генетический код, считывание гена, транскрипция гена, генетическая информация, квазисознание, когнитивная метафора.

Золян Сурен Тигранович – доктор филологических наук, профессор, приглашенный профессор Балтийского федерального университета имени Иммануила Канта, исполнитель проекта в Институте научной информации по общественным наукам РАН. ведущий научный сотрудник Института философии, социологии и права Национальной академии наук Республики Армения.

surenzolyan@gmail.com

http://orcid.org/0000-0002-4422-5792

Для цитирования: *Золян С.Т.* Может ли рибосома «читать»? О когнитивной основе и эвристической значимости лингвистических метафор в молекулярной генетике (Часть 2) // Философские науки. 2020. Т. 63. № 2. С. 46–62. DOI: 10.30727/0235-1188-2020-63-2-46-62

On the semiotics of genetic reading

In this chapter, we address the semiotics of genetic “reading.” Like the other metaphorically used linguistic terms in genetics, it refers to some very complicated biochemical processes. However, the matter will change if one makes a shift from the biochemical substance of genetic translation and concentrates on its processing. Then, considerable similarity between human and genetic reading can be drawn. In general, reading is understood as a complex cognitive process of interaction between a text and a reader, and, apart from some particular cases, its physiological or physical features are not taken into account. Thus, eye-tracking and the font size are necessarily crucial for reading, but they are out of semiotic consideration. As it is generally accepted, reading has different degrees of complexity and presupposes faculties of (a) recognition of letters, (b) combining them into syllables, words, and sentences, (c) assigning them a meaning, and (d) comprehension of a complete text [Aase et al. 2009]. All these operations are based on different types of semiotic correlation between signified and signifier.

It should be clarified whether such relationships exist in genetics? Reading in the most general sense is an operation with signs. Is it possible to present biochemical entities as signs and what are the signifier and the signified in this case?

It can be seen that the relations between triplets and amino acids are not determined by biochemical properties, but are arbitrary, as it is usual in a language. Thus, human genetics is determined by two similar, but non-coincident codes, canonical and mitochondrial [Barrell, Bankier, & Drouin 1979], where tryptophan and methionine are encoded differently.

The arbitrariness of genetic code as its essential property was already mentioned by Francis Crick. Although he referred to an interaction of two languages, the arbitrariness that he indicated is closer to a relationship between the signifier and the signified. In accordance with the Central dogma, this relation is asymmetric: the nucleotides are "translated" into proteins, but the opposite process is impossible, so this is a process of a hierarchical signification, but not a translation between two languages. F. Crick considered this arbitrariness as the main difference between the regularity of the genetic code from the chemical one described in the Mendeleev Periodic Table of Elements:

> The Periodic Table would be the same everywhere in the universe. The genetic code appears rather arbitrary, or at least partially so. Many attempts have been made to deduce the relationship between two languages from chemical principles, but so far none have been successful. The code has a few regular features, but these might be due to chance [Crick 1981, 46–47].
>
> Since that time, this assumption of the arbitrariness of the genetic code has been repeatedly confirmed in spite of the still existing prevalence of the opposite, "stereo-chemical" point of view [Barbieri 2018, 2].

How is it possible to attribute the semiotic characteristics to the gene expression (transcription and translation)? Firstly, one should explicate what a relation between signifier and signified regarding genetic code is. Crick's above mentioned definition of the genetic code as a relationship between word-stocks of two languages, between nucleotides and amino acids can be re-formulated in a linguistically more accurate way (esp. having in mind the irreversibility between them: nucleotides are replaced by amino acids, but not vice versa). Any of triplets can be considered as a signifier related to a signified, i.e., the amino acid which is associated with the given triplet in DNA (codon – in mRNA, anticodon – in tRNA). For example, the triplet/codon AUC can be described as a signifier referring to its signified, isoleucine[1].

---

[1] Cf.: "We can thus interpret the gene as a triadic sign: it has a 'primary side,' the chemical structure or the 'chemical sensuality' of the DNA molecules, the significant of the gene. Furthermore this piece of DNA enters into a relation which mediates its signification as a code for a specific sequence of amino acids. That is, the DNA piece as gene (i.e., as sign) is a relation to another object, the protein, symbolized by the genetic code in this gene. Finally, this relationship between the primary sign (the chemical DNA) and the protein (the object) is mediated by a complex mechanism of transcription, RNA-processing and translation, that interprets the DNA sequence in the cell: ultimately it is the

In the case of genetic "reading," the processes are beyond the trivial sign copying, re-writing, recognition, and non-alternative encoding: as there is no one-to-one correspondence between triplets and amino acids. Besides, during its expression, at each stage a triplet is transformed into its reverse mirror counterpart, as well as each of the nucleotides A, T/U, T, C is replaced by its "complement," or "antipode" (in DNA: A ↔ T, C ↔ G, in RNA: A ↔ U, C ↔ G.) For example, the codon AUC, associated with isoleucine, is transformed into the anticodon GAU.

The ribosome together with mRNA and tRNA is capable of recognizing biochemically completely different nucleotide triplets (the signifiers) and correlating them with the same amino acid (the signified). However, it can also associate the same signifier (triplet) with different signs. Thus, being located in different strands, different complementary nucleotide sequences (triplets) may encode the same amino acid, but if it is located in one and the same strand, the same sequence will encode other amino acids. For instance, methionine in the different strands of DNA and RNA is encoded by these different triplets:

ATG (in the non-transcribed strand of DNA),
TAC (in the transcribed strand of DNA),
AUG (codon in a messenger RNA),
UAC (anticodon on transfer RNA).

At the same time, when AUG is located on tRNA or UAC on mRNA, they are related to another amino acid: tyrosine, or, under certain conditions, are recognized as a signal for termination of translation).

However, this alternating situation will be changed if one uses the other type of notation based on the signified. In our example, if instead of the triplet of letters (A, U, G, C, T) denoting the composition of nucleotides, we use the names of the amino acids corresponding to them, i.e., methionine, then the entire process of transferring genetic information will appear as a repetition of the same sign: Methionine (as it is encoded in the untranscribed DNA strand) – Methionine (as it is encoded on the transcribed DNA strand) – Methionine (as it is coded for mRNA) – Methionine (as it is coded for tRNA).

This form of notation ceases to reflect the biochemical composition of the triplet, but it explicates the identity of its function in the different contexts. The location on different strands can be considered as contexts, so it is possible to reformulate the previous notation as a set of context-dependent varieties:

---

whole cell itself that participates in the network necessary for such an interpretation" [Emmech & Hoffamyer 1991, 34].

Methionine => ATG (in context of non-transcribed strand of DNA),

=> TAC (in the context of transcribed strand of DNA),
=> AUG (in the context of mRNA)
=> UAC (in the context of tRNA).

These changes are similar to alterations of the same linguistic entities under different conditions and contexts, where the identity of this unity is preserved in spite of variations of forms and meanings (cf.: "I – me –mine," "to be – is – are – was – were – been – being," etc.). Of course, the regularity of alteration of triplets/codons resembles the regular varieties of forms (as in a case of regular verbs or declination). As usual, in vocabularies and grammars, only the single "canonical" variant is given, and the same occurs in a case of genetic code – in all the tables and schemes a codon in the mRNA represents the class of all of its contextual variants.

Thus, it is possible to describe the gene expression using three forms of a meta-linguistic notation: (a) a signifier, a biochemical composition of the nucleotide triplet and its canonical changes depending on its location, is to be described; (b) description can be based on the correlation between initial and resulting entities, as it is given in the genetic code: some biochemical composition, as it occurs in mRNA, within one of possible contexts, is taken with the corresponding amino acid in the standard genetic code; (c) finally, a description can be based on a signified, meaning the resulting amino acids, and all the varieties are regarded as context-sensitive signifiers, regardless of their biochemical substance (that is, methionine – as it is encoded in this or that context). All of them have their specific scope of applicability. The first type is more convenient for describing experimental data. For the modeling of information processes, a combination of the second and third types seems to be more appropriate, as it allows taking into account correlations between biochemical forms, functional contexts, and resulting meanings.

Such representation makes it clear that the decisive factor for protein synthesis is not only a biochemical composition of a triplet, but a coding function: the correlation between a nucleotide triplet, its position, and the amino acid to be chosen. Thus, a codon (or anticodon) is not a biochemical substance, but a particular function defined within some positional context. Formally, it can be represented as a function operated under some contextual features. Besides their location on some strand, a position within the sequence also is relevant. The same codon AUG, which is usually related to methionine, when under specific conditions located in the initial position, serves as a marker of the beginning of the text and does not perform an encoding function. The opposite situation occurs with other nonsense codons – UAA, UAG, UGA. Typically, they are used as markers of termination, but under some specific conditions, they are associated with non-canonical amino acids – selenocysteine and pyrrolizine [Turanov et al. 2008].

These basic characteristics of genetic translation demonstrate that there is a distinction that is fundamental for all sign systems, between abstract elements and their concrete manifestations, between types and tokens. Thus, a cell identifies the biochemical sequences based on their function in an abstract system; and it is capable of reading not only "letter by letter," but also – at least – by identifying words and understanding their meanings. This processing presupposes more than the simple recognition of signs and implies the ability to operate with abstract entities. Coming back to the abovementioned abilities, let us consider to what degree they are comparable with the characteristics of gene expression.

(1) Recognition of letters. It is evident that the minimal constituents of genetic sequences, nucleotides are recognized.

(2) Combining them into syllables, words, and sentences. The term "reading frame" properly refers to processes of segmentation – a ribosome deals with three units and identifies them as a "word" standing for one of the amino acids.

(3) Correlating them with a meaning. The relation between some triplet and amino acid associated with it constitutes the genetic code and is manifested through the whole process of gene expression. This presupposes, in addition to the ordinary portrayal of a ribosome and RNA as very complex and fast-running bio-machines, that it is able to operate with abstract entities: types and meanings (signifieds).

(4) Comprehension of a complete text. Four triplets are used as special marks (start-codon and stop-codons) for initializing and finalizing sequences – when the polypeptide chain associated with some gene (operon) is completed; the process of translation should be terminated in order to start a new one. Thus, it is possible to notice some understanding of what a text is like, where it begins and when and where it should be stopped. Besides, contextual sensitivity and distinguishing between homonymic forms presupposes an ability to read a text as a whole (see more in: [Zolyan & Zhdanov 2018]). Generally speaking, this reminds of one of the crucial statements

of the theory of linguistic translation: "to understand is to translate" [Steiner 1975; Ricœur 2006, 28].

Of course, the reading ability is very modest in comparison with Schrödinger's "all-penetrating mind." But it is real and is not a product of scientific imagination like Maxwell's Demon that was destined to measure a temperature of gas and to open doors for hot molecules. It is also evident that the processes of gene expression cannot be exclusively described as characteristic for complex self-regulating systems, but presuppose some relevant quasi-intelligent faculties and features.

Therefore, a question can arise: who is the subject of this semiotic activity? Do we assign these features to some self-reading and self-editing and even self-writing Creature, as can be inferred from some passages mentioned above? Alternatively, this is a typical metaphorical personification when natural processes are represented as deeds of some physical or metaphysical actors – the wind blows, the sun comes out, life punishes, etc. This mode of thinking and speaking about natural processes is inevitably extended on scientific descriptions if they are presented in narratives told in natural language.

However, this case differs from the usual metaphoric or metonymic personification, and it is possible to speak about some quasi-mind, but in a very narrow and peculiar Peircian sense, as derived characteristics of a sign system. Of course, in linguistics, it is possible and, as a rule, productive to consider language as an autonomous system, without regard to speakers, as this was stated by the founding father of structural linguistics [de Saussure 1959, 232].

Nevertheless, this is valid to some degree. The very notion of a sign presupposes the pragmatic dimension. Thus, one should refer to some implicitly constructed subject of semiosis. If a sign and a sign system are taken in their active expression, within communication, agents emerge as necessary constituents of those processes regardless their actual physical presence. However, in the case of linguistic communication one can get away from this problem, as it is possible to depict some correspondences for abstract constructed communicants with their actual counterparts. An abstract reader can be substituted by some actual reader, and vice versa. However, this is not a case in bio-semiosis, or in formal semiotic systems like logic and mathematics. For example, there can be self-describing and even self-predicting formal systems (automata) [Albert 1987] without "actual" communicants, or some social self-referring systems [Luhmann 1990], where it is impossible to explicate actual "senders" and "recipients" of messages of communication between social institutions. At the same time, in all these cases the typical characteristic of a sign or a sign system and semiosis are definitely present.

In his time, having foreseen this situation, Charles Peirce introduced the notion of quasi-mind, as a characteristic inherent to sign systems:

> Thought is not necessarily connected with a brain… Admitting that connected Signs must have a Quasi-mind, it may further be declared that there can be no isolated sign. Moreover, signs require at least two Quasi-minds; a Quasi-utterer and a Quasi-interpreter; and although these two are at one (i.e., are one mind) in the sign itself, they must nevertheless be distinct. In the Sign they are, so to say, welded [Peirce 1906a, 523].
>
> …But a thought, to gain any active mode of being must be embodied in a Sign. A thought is a special variety of sign. All thinking is necessarily a sort of dialogue, an appeal from the momentary self to the better considered self of the immediate and of the general future. Now as all thinking requires a mind, so every sign even if external to all minds must be a determina-tion of a quasi-mind. The quasi-mind is itself a sign, a determinable sign [Peirce 1906b].

This notion will prevent from a quest of cell's and ribosome`s cognitive faculties and its physiological apparatus, as well as from attributing it to the monotheistic God or the pantheistic Nature. A cell can be considered as quasi-intelligence endowed with some cognitive faculties and capable of operating with semiotic entities in the same sense that the Piercian quasi-mind: thus is a determinable sign. However, this is rather a semiotic, but not a biological characteristic. The idea of C. Peirce to consider quasi-mind as the faculty derived from operations with determinable signs seems to be the most appropriate. Of course, this cannot explain who really reads either the whole Book of life, or particular codons of mRNA, (if there were such a Reader), but, perhaps, this does not matter, as it provides a methodological opportunity for investigating genetic information as a semiotic entity.

To conclude: the semiotic/biochemical duality of genetic information and complementarity of its description

In the final section, we would like to discuss the heuristic value of linguistic (or semiotic) approach, if it claims to be more than a cognitive metaphor but a method of description. Is there any reason for such consideration and for describing the biochemical reactions as some semiotic operations?

Different modes of description based on this duality enable different correlations of genetic information with language and linguistics. In general, linguistic terms may be used: (a) in their literal terminological sense that they retain in molecular genetics, (b) metaphorically, as a popular explanation or heuristic pattern, (c) in their literal terminological sense, as they used in semiotics. In accordance with these approaches, genetic reading can be interpreted as (a) a frozen metaphor, where the term has lost its linguistic connotation, (b) proper metaphor, (c) "more than a metaphor" (model, pattern), (d) proper (non-metaphorical) linguistic, or semiotic term. Three of these options relate to the different modes of describing the biochemical substance, and the fourth can be suggested as its semiotic complement.

(1) In the first case, linguistic terms are only frozen metaphors (or "dead metaphors"), the rigid interpretation regulates their semantics in terms of biochemical reactions. "Reading" is understood as a regular operation of correlating between some biochemical entities (codons, anticodons, amino acids), without any textual connotations. From this point of view, DNA is not a language, a genome is not an information system, as it was assumed in [Kay 2000] and then repeated by many.

(2) The proper metaphorical usage related to a text, language and reading one can find in papers of the pioneers of molecular genetics (F. Miesher, G. Gamow, F. Crick, M. Nierenberg), who invented most of its terminology and were directed by semiotic analogies in their vision of processing of genetic information. However, a ribosome who reads out while jumping, which appears in Crick's book, can be interpreted only as a personification: in this collocation the term "ribosome" is used in its proper sense, and both "reading" and "jumping" are used metaphorically.

(3) The usage of linguistic terms can be viewed as "more than a metaphorical usage" or "is not a mere figurative expression" [Jakobson 1970, 437] – this statement was first made by linguists:

> When I [Jakobson] first came across linguistic terms in the biological literature, I said to myself: we need to check whether this is just a manner of speech, a metaphoric usage, or whether there is something deeper here. I must say that what biologists have done is quite legitimate from a linguistic standpoint, and in fact, we can take things even further [Lévi-Strauss, Jakobson, & L'Heritier 1968] (cited in: [Katz 2007, 57]).

Indeed, this approach also was legitimized by new findings in bio-linguistics, bio-semiotics, and bioinformatics. It does not matter, whether the genetic code and genomics form a language or not; at any rate, they can be described as a language, with alphabets and grammar systems. From this point of view, most of existing genetic descriptions are not mere metaphors, but heuristic models [Jacob 1977], where biochemical substances are represented as semiotic sequences of letters and letter sequences (triplets, amino-acids, proteins, genes, genome). Bioinformatics, then biosemiotics, and now biolinguistics do it in an explicit way and then reveal the numerous analogies and similarities between language and genomics, and this can lead even to some hypotheses about the common origin of genetic and natural languages. This makes it possible, to consider DNA-ese as a nucleic sign system: "After exploring the structural symmetries between the genetic and verbal codes, we conclude that the linguistic concepts used in biology are more than just heuristic metaphors. Though tainted by anthropomorphism, they may refer to a sophisticated form of protolanguage whose genetic grammar could have gradually mutated into several stages of expression: nucleic, proteic, physiological, verbal" [Katz 2008, 69] (see also: [Searls 2002]).

At the same time, the deciphering of a genome and its equating with the Book have become a methodological basis for bioinformatics: rather than concentrating on the computer processing of the biochemical data, bioinformatics practitioners claims to develop an instrument of reading and interpretation of genetic texts:

> With the publication of the human genome sequence, we are passing into a new phase in the analysis of what is popularly being called the "Book of Life"… The role of the may also be expected to change, by degrees: one may become less like an archaeologist, discovering and poring over shards of evidence to piece together rudimentary translations, and more like a literary critic, attuned to theme and variation, elucidating ever more subtle nuances of meaning and interrelationship in a well-worn textus receptus… Tools and techniques of a linguistic character have proven useful in biological sequence analysis, especially in the trend toward algorithms that model the syntactic features of the domain with increasing sophistication… While the comparison may seem fanciful, there are clearly instructive analogies to be drawn between genomic and literary texts, and perhaps it is not so great a stretch to contemplate the grammar of genes, the poetics of proteins, and

the essential kinship of philology and phylogenetics [Searls 2001, 579, 580] (see also: [Abel & Trevors 2006]).

(4) However, the approach "more than a metaphor" has some inconsistency. One can see that in some cases linguistic terms cease to be dead or living metaphors and re-acquire their literal meaning, the one they have in linguistics and semiotics. Actually, for some features of the genetic code (its arbitrariness, symbolic nature of triplets, reading frame and its shift, context-dependence) the semiotic terms are more applicable, as they do not have correlates in biochemical substances. The dual – biochemical and informational – nature of the genetic code and genome presupposes that its description one should be based on the principle of complementarity. As in the case of the wave-particle duality of physical entities, it is impossible to observe and describe both the biochemical and informational properties of genetic entities, and only when taken together, they present a fuller picture.

Linguistic metaphors denoting biochemical entities and processes seem to be a method of combined theoretical description and explanation of these heterogeneous properties. The duality of genetic information will be represented through the double theoretical description, as we demonstrated earlier in the case of genetic translation [Zolyan & Zhdanov 2018]. However, since the conceptual apparata of molecular biochemistry and semiotics cannot be combined, one can suggest multi-level explication of bio-semiotic metaphors: the genetic reading may be interpreted as a frozen metaphor for biochemical interaction between codon and anticodon, and at the same time as a linguistic operation with textual entities (reading, proofreading, editing, etc.).

These two aspects are not equal. Since the biochemical substance seems to be evident and certain, descriptions in biochemical terms have never been questioned. On the contrary, suggestions to restrict semiotic connotations and analogies are often raised by biologists, as if these may led away from the essence of the processes, or were redundant and added nothing to the existing biochemical descriptions. This seems to be natural: when semiotic phenomena are reduced to a biochemical substratum, then semiotics cannot say anything new, except some metaphorical reformulations of the known facts. However, some regularities (or irregularities) demonstrate that the biochemical explanation of genetic coding is not sufficient.

> [T]he genetic code has been accepted under the assumption that its rules were determined by chemistry and do not have the arbitrariness that is essential in any real code… This inevitably implies that the genetic code is a metaphorical entity, not a real code. This idea has a long history and let us not forget that for many decades it has been the dominant view in molecular biology… It has taken a long time and much experimental work to overturn this conclusion, but eventually, it has been shown that there is no deterministic link between codons and amino acids because any codon can be associated with any amino acid. This means that the rules of the genetic code do not descend from chemical necessity and in this sense they are arbitrary. Today, in other words, we have the experimental evidence that the genetic code is a real code, a code that is compatible with the laws of physics and chemistry but is not dictated by them [Barbieri 2018, 2].

This is an essential point which exceptionally motivates such an extension and application of linguistic-semiotic metaphors and converting them into models and theories. There are some significant characteristics of a genome which cannot be explained only by biochemical principles. Besides the above-mentioned principle of arbitrariness of relation between triplets and amino acids, the principle of contextual dependence (sensitivity) should also be mentioned, when the same biochemical sequence of nucleotides, depending on their location, acquires a different meaning. However, this is only a manifestation of a more general regularity – it is not only the composition of elements, but also their arrangement that is valid: AUG is not equal to GAU or UGA, etc. (compare this with some other chemical or biochemical entities, where order of components is not essential: $H_2O = OH_2$). This principle is valid at higher levels, too: “A protein is like a paragraph written in a twenty-letter language, the exact nature of the protein being determined by the exact order of the letters” [Crick 1981, 48].

The same comparison can also be extrapolated regarding genes (operons) and genome. In general, the order of symbols is a crucial semiotic principle of formation (for example, the words “done” and “node” are composed of the same letters but are different as they have different signifieds). Based on this, in his last works the pioneer of bioinformatics in the USSR Vadim Ratner suggested considering a genome not only as a biochemical phenomenon, but also as a semiotic one:

> Genes are not the germs of biological structures, but resemble linear texts (DNA sections), written under certain rules and carrying genetic information about molecular structures and functions… Some "meaningless" areas do not encode a protein or encode spoiled proteins. In all cases, both genes and non-coding areas are segments of DNA molecules, i.e., are constructed from the same alphabet of four nucleotides. Therefore, the differences between such texts are not in their physical nature, but exclusively in a succession of symbol-monomers. This is the key to the information-linguistic approach. Hence, genes are not physical, but information units of heredity [Ratner 2000, 24].

The combination of these heterogeneous beginnings can be regarded as an organic metaphor, metaphor in the flesh in blood, as this was suggested by Russian biologist and bio-semiotician Aleksandr Sedov:

> It is not our language that uses this metaphor, the genome itself is such a dual agglomerate, where, in certain aspects of it, we see its semiotic essence, and in others – a biochemical one, and one cannot be separated from the other. Many biological subjects and processes resemble written texts and/or hand-made artifacts, while being billions of years older than human beings… Hence, using the semiotic approach, biologists and humanitarians can cross-share their cognitive models, and then test them empirically for fruitfulness and predictive ability inside living organisms… When comparing organisms with texts, genes and cells are like "fleshy metaphors" that emerged long before human speech, languages, and written texts [Sedov 2000, 532–533].

Thus, there is possible interaction between the duality of a genetic code and duality of a linguistic expression, which can be interpreted in both of its meanings, as either literal or metaphorical. The principle of complementarity assumes that a semiotic description, originally formed from a metaphor, can enter a theory and give a new vision of processing genetic information. From this point of view, the perception of operations involved with processing of genetic information will cease to appear as a sequence of reactions denoted through useful or redundant metaphors; instead, they can be regarded as purposeful and therefore meaningful actions. Accordingly, the genetic code appears not as a frozen accident (F. Crick), but as a regulated semiotic system (language) which generates texts that should be adequately interpreted (read and translated).

REFERENCES

Aase L., Fleming M., Ongstad S., Pieper I., & Samihaian F. (2009) Reading. In: *Platform of Resources and References for Plurilingual and Intercultural Education.* Strasbourg: Language Policy Division, Council of Europe. Retrieved from http://rm.coe.int/CoERMPublicCommonSearchServices/DisplayDCTMContent?documentId=09000016805a234b

Abel D.L. & Trevors J.T. (2006) More than Metaphor: Genomes are Objective Sign Systems. *Journal of BioSemiotics.* Vol. 1, no. 2, pp. 253–267.

Albert D.Z. (1987) Quantum-Mechanical Automation. *Philosophy of Science,* Vol. 54, no. 4, pp. 577–585.

Barbieri M. (2018) What is Code Biology? *Biosystems.* Vol. 164, pp. 1–10.

Barrell B.G., Bankier A.T., & Drouin J. (1979) A Different Genetic Code in Human Mitochondria. *Nature.* Vol. 282, no. 5735, pp. 189–194.

Crick F. (1981) *Life Itself: Its Origin and Nature.* London: Simon and Schuster.

de Saussure F. 1959. *Course in General Linguistics.* (C. Bally, A. Sechehaye, Eds.; W. Baskin, Trans.). New York: Mc-Graw Hill Book.

Emmeche C. & Hoffmeyer J. (1991) From Language to Nature – the Semiotic Metaphor in Biology. *Semiotica.* Vol. 84, no. 1/2, pp. 1–42.

Jacob F. (1977) The L*inguistic Model* in B*iology. In: Armstrong D. & van Schooneveld C.H. (Eds.) Roman Jakobson. Echoes of His Scholarship* (pp. 185–192). Lisse: Peter de Ridder.

Jakobson R.O. (1970) Linguistics and Its Relation to Other Science. In: *Main Trends of Research in Social and Human Sciences. Part 1. Social Sciences* (pp. 419–463). Paris: Mouton.

Katz G. (2008) The Hypothesis of a Genetic Protolanguage: an Epistemological Investigation. *Biosemiotics.* Vol. 1, no. 1, pp. 57–73.

Kay L. (2000) *Who Wrote the Book of Life?: A History of the Genetic Code.* Stanford: Stanford University Press.

Lévi-Strauss C., Jakobson R., & L'Heritier P. (1968) Vivre et parler. *Les Lettres françaises,* No. 1221, pp. 3–7 (in French).

Luhmann N. (1990) *Essays on Self-Reference*. New York: Columbia University Press.

Peirce C. S. (1906a) Prolegomena to an Apology for Pragmaticism. *The Monist.* Vol. 16, pp. 492–546.

Peirce C.S. (1906b) Letters to Lady Welby. In: Bergman M. & Paavola S. (Eds.) *The Commens Dictionary: Peirce's Terms in His Own Words. New Edition.* Retrieved from http://www.commens.org/dictionary/entry/quote-letters-lady-welby-17.

Ratner V.A. (2000) Khronika velikogo otkryiya: idei i litsa [Chronicle of the great discovery: ideas and persons]. *Priroda.* 2000. No. 6, pp. 22–30 (in Russian).

Ricœur P. (2006) *On Translation* (E. Brennan, Trans.). London: Routledge.

Searls D.B. (2001) Reading the book of life. *Bioinformatics.* Vol. 17, no. 7, pp. 579–580.

Searls D.B. (2002) The Language of Genes. *Nature*. Vol. 420, no. 6912, pp. 211–217.

Sedov A.E. (2000) Metafory v genetike [Metaphors in genetics]. *Herald of the Russian Academy of Sciences.* 2000. No. 6, pp. 526–534 (in Russian).

Steiner G. (1975) *After Babel: Aspects of Language and Translation.* Oxford: Oxford University Press.

Turanov A.A., Lobanov A.V., Fomenko D.E., Morrison H.G., Sogin M.L., Klobutcher L.A., Hatfield D.L., & Gladyshev V.N. (2008) Genetic Code Supports Targeted Insertion of Two Amino Acids by One Codon. *Science.* Vol. 323, no. 5911, pp. 259–261.

Zolyan S. & Zhdanov R. (2018) Genome as (Hyper)Text: From Metaphor to Theory. *Semiotica.* No. 225, pp. 1–18.