Языкознание
Linguistics

Linguistic Data Model for Natural Languages and Artificial Intelligence. Part 2. Identification

Oleg M. Polyakov✉
Saint-Petersburg State University of Aerospace Instrumentation, St Petersburg, Russia

✉road.dust.spb@gmail.com

Introduction. The article continues a series of publications on the linguistics of relations (hereinafter R-linguistics) and is devoted to the origin of signs, their independence and determination of the dimension of linguistic spaces.

Methodology and sources. The article is devoted to the continuation of the axiomatic approach, but for the identification process. Research methods are to develop the necessary mathematical concepts for linguistics in the field of identification.

Results and discussion. The concept of a sign is defined and its interrelation with decomposition of linguistic spaces is established. This radically changes the attitude to signs in linguistics, where the "external" origin of signs is assumed. It is shown that the decomposition of linguistic spaces into signs spaces entails the independence of signs and the possibility of decomposition of objects of identification. It is fundamentally distinguished by the signs on the parameters. On the basis of the independence of signs it is possible to formulate the notion of dimension of linguistic spaces, which is defined as the smallest number of signs describing the linguistic space. In the lattice of linguistic spaces there is a division operation, which allows to simplify the selection of signs.

Conclusion. The main conclusions are as follows. Signs, on the basis of which the identification of objects in the category, are abstract mathematical objects associated with the decomposition of linguistic spaces. Signs are independent from each other and allow not only to make decomposition of spaces, but also to decompose on parts objects of identification. Their origin is not related to their presence in the "outside world", so it, as will be shown later, creates the basis for the emergence of language.

Key words: R-linguistics, identification, signs, dimension of linguistic space, signs independence.

Лингвистическая модель данных для естественных языков и искусственного интеллекта. Часть 2. Идентификация

О. М. Поляков✉

Санкт-Петербургский государственный университет
аэрокосмического приборостроения, Санкт-Петербург, Россия

✉road.dust.spb@gmail.com

Введение. Статья продолжает серию публикаций по лингвистике отношений (далее R-лингвистика) и посвящена происхождению признаков, их независимости и определению размерности лингвистических пространств, аксиоматика которых сформулирована в первой части серии.

Методология и источники. Методы исследования заключаются в разработке необходимых математических понятий для лингвистики в области идентификации.

Результаты и обсуждение. Определено понятие признака и установлена его взаимосвязь с разложением лингвистических пространств. Это коренным образом изменяет отношение к признакам в лингвистике, где предполагается их «внешнее» происхождение. Показано, что разложение лингвистических пространств в признаковые пространства влечет за собой независимость признаков и возможность декомпозиции объектов идентификации. Это принципиально отличает признаки от параметров. На основе независимости признаков можно сформулировать понятие размерности лингвистических пространств, которое определяется как наименьшее число признаков, описывающих лингвистическое пространство. В решетке лингвистических пространств действует операция деления, которая позволяет упростить выделение признаков.

Заключение. Основными выводами являются следующие положения. Признаки, на основе которых производится идентификация объектов в категории, являются абстрактными математическими объектами, связанными с разложением лингвистических пространств. Признаки независимы друг от друга и позволяют не только производить разложение пространств, но и декомпозировать на части сами объекты идентификации. Их происхождение не связано с их присутствием в «окружающем мире», поэтому оно создает основу для возникновения языка.

Ключевые слова: R-лингвистика, идентификация, признаки, разложение лингвистического пространства, независимость признаков.

Introduction. The representation of the model in the form of linguistic spaces connected by verbs does not significantly distinguish it from the relational model. You can show how you can go from one model to another. In fact, the linguistic model is not a model yet, because it has no predictive function. In this article, we will take a cardinal step: define the concept of sign, which in the language usually correspond to adjectives. From this point on, the linguistic model will have a predictive function and the similarity of relational and linguistic models will be a thing of the past.

It is important to understand that a sign is an abstract mathematical object. It can be embodied in the world around us, but it is not necessary. This feature differs from the parameters – the actual

measured (sense organs or devices) characteristics of objects and phenomena. A sign stands between categories and the world in the form of a matching element. Its functions are extensive and on its basis the identification of concepts is carried out. This view of signs is the most important feature of this approach and distinguishes it, for example, from the views of cognitive linguistics.

Methodology and sources. The results of the following mathematical sections are used as a research method: the theory of relations, the theory of lattices, the theory of data dependencies. The research materials are presented in the form of mathematical proofs of various properties of the linguistic model associated with the identification process.

Results and discussion.

The origin of the signs.

1. *Signs and decomposition spaces.*

Definition 1. Under a *sign* with the name A is understood a finite set $D(A)$ of sign meanings (domain), consisting of not less than one element. The set $D(A)$ has a selected element called 0 (zero). The relation S_A 'to have the meaning of sign A' is defined as follows: $(x, 0) \in S_A$ for any x from the universe; if $(x, a) \in S_A$ and $(x, a') \in S_A$, then either $a = 0$ or $a' = 0$ or $a = a'$. If $(x, a) \in S_A$, then it is said that object x has the meaning 'a' of sign A. This can also be written as $x(A) = a$. The relation S_A is called the *systematization* relation.

So, the sign has at least zero meaning for each object, and its scale (set of values) $D(A)$ contains at least one element – 0. If the object has a meaning different than 0 of sign A, then it is the only one. Such a slightly strange definition of a sign can be explained as follows. The meaning 'zero' is actually intended to define another relation 'to have the sign A'. If for some object x there is only the meaning $x(A) = 0$, it means that the object x does not have the sign A. If object x has another meaning of 'a', then, of course, object x has a sign A. Definition 1 actually combines two relations: 'to have a sign' and 'to have a sign meaning'. If you look at the S_A as a certain device to which objects from U are 'connected', and its display shows the meaning of the sign A, then the zero meaning for some x shows that the property A is not characteristic of X. If the device gives a non-zero meaning 'a', then, on the contrary, x has this property (sign) and it has the meaning 'a'. Of course, if x has the sign A, we will never see the meaning 0 on the display, although by definition the relation S_A is always $(x, 0) \in S_A$. These are the costs of combining two relationships into one. In addition, the device S_A for the same object cannot show different non-zero meanings. In general, signs are idealized properties of objects, which are usually the result of processing real data.

In contrast to signs, parameters are the real properties of objects that we perceive using devices or senses. Parameters can have different meanings for the same object. For example, I can have a different pulse, eyes closed or open etc. So for parameters, the meaning 0 is not selected: it is one of the possible meanings. There are other differences that we will look at later.

It should be mentioned in the first part was introduced the operation of mixing spaces '∘'. This operation forms a new space from the original spaces (on one universe), so that the original spaces are included in the mixed space and the least number of new categories is added. It follows from the definition of the mix operation that it can easily be extended to an arbitrary set of spaces.

Definition 2. Let $\mathbb{P}1$ and $\mathbb{P}2$ two spaces of the universe U. We say that $\mathbb{P}2 \leq \mathbb{P}1$ ($\mathbb{P}2$ is less than or equal $\mathbb{P}1$) if and only if $\mathbb{P}2 \circ \mathbb{P}1 = \mathbb{P}1$.

From the definition of the mixing operation should be: if $\mathbb{P}2 \circ \mathbb{P}1 = \mathbb{P}1$, then each category of $\mathbb{P}2$ is the category $\mathbb{P}1$, it means $\mathbb{P}2 \subseteq \mathbb{P}1$. As the inclusion relation is the order relation, the relation

$\leq$ is the same order relation. We denote by $\mathbf{P}_U$ the set of all spaces over U and call it a *systematization lattice*. In $\mathbf{P}_U$ there is the largest element (unit) and the smallest element (zero). The largest space is Boolean U, as it includes any space over U, and the smallest space consisting of a single element U, as this element enters any space over U. On the systematization lattice, the intersection operation of spaces, which was introduced in the first part, is also defined.

Theorem 3. Set spaces $\mathbf{P}_U$ is a complete lattice.

Proof. As noted, $\mathbf{P}_U$ has a unit (Boolean U). Let $\mathbb{P} = \bigcap \mathbb{P}'$ for $\mathbb{P}' \in \mathbf{P}'_U \subseteq \mathbf{P}_U$, where $\mathbf{P}'_U$ – some set of spaces from $\mathbf{P}_U$. Show that $\mathbf{P}'_U$ is the exact lower bound in $\mathbf{P}_U$. Let there be a $\mathbb{P}''$ such that $\mathbb{P} < \mathbb{P}'' \leq \mathbb{P}'$ for all $\mathbb{P}'$ from the family $\mathbf{P}'_U$. Then $\mathbb{P}'' \subseteq \mathbb{P}'$ for all $\mathbb{P}'$ of $\mathbf{P}'_U$, consequently, $\mathbb{P}'' \subseteq \bigcap \mathbb{P}'$ and that means $\mathbb{P}'' \leq \mathbb{P}$, which contradicts the assumption. Thus, $\mathbb{P} = \mathbb{P}''$ and $\mathbb{P}$ is the exact lower bound for the family $\mathbf{P}'_U$. Then, by $\mathbf{P}_U$ is a complete lattice [1].

Fig. 1 shows the lattice $\mathbf{P}$ on a universe of two elements $\{a, b\}$. The ratio of the precedence in $\mathbf{P}$ between the spaces is shown by arrows, the relation of the precedence inside the spaces is shown by solid lines. The lattice $\mathbf{P}$ itself is an interesting formation, since it has many interesting properties that are of practical importance. E. g., for the lattice $\mathbf{P}$, a strengthened version of the Kurosh-Ore theorem on the decomposition of elements $\mathbf{P}$ is fulfilled.

Definition 4. A space $\mathbb{P}$ is called *indecomposable* by the mix operation (hereinafter simply indecomposable) if $\mathbb{P} = \mathbb{P}_1 \circ \mathbb{P}_2$ implies $\mathbb{P} = \mathbb{P}_1$ or $\mathbb{P} = \mathbb{P}_2$. The decomposition of an arbitrary linguistic space $\mathbb{P}$ in the form $\mathbb{P} = \mathbb{P}_1 \circ \ldots \circ \mathbb{P}_n$ is called an *irreducible* representation if $\mathbb{P}_1, \ldots, \mathbb{P}_n$ are indecomposable and for any $i = 1, \ldots, n$ $\mathbb{P}_i \nleq \mathbb{P}_1 \circ \ldots \circ \mathbb{P}_{i-1} \circ \mathbb{P}_{i+1} \circ \ldots \circ \mathbb{P}_n$.
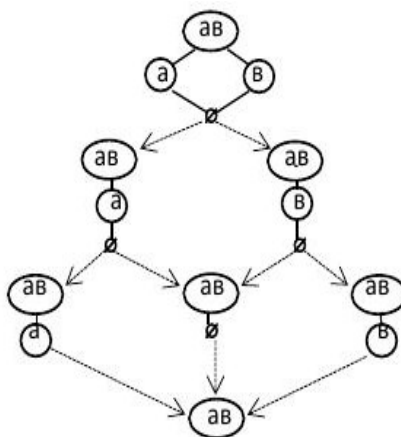

Fig. 1. Lattice of linguistic spaces for two objects

Theorem 5. For each space of $\mathbf{P}$ there exists a unique irreducible representation.

Proof. We first show that in $\mathbf{P}$ zero are indecomposable, as well as all spaces consisting of two categories (one of which is U), and only they. Let $\mathbb{P} = \{U, X\}$, where $X \subseteq U$, and for some $\mathbb{P}_1$, $\mathbb{P}_2$ $\mathbb{P} = \mathbb{P}_1 \circ \mathbb{P}_2$. Then $\mathbb{P}_1 \subseteq \mathbb{P}$ and $\mathbb{P}_2 \subseteq \mathbb{P}$. If both inclusions are strict, then $\mathbb{P}_1 = \mathbb{P}_2 = \{U\}$, which contradicts the assumption. Thus, either $\mathbb{P} = \mathbb{P}_1$ or $\mathbb{P} = \mathbb{P}_2$. Now suppose that $\mathbb{P} = \{U, X_1, \ldots, X_r\}, r > 1$. Then for $\mathbb{P}_1 = \{U, X_1\}, \ldots, \mathbb{P}_r = \{U, X_r\}$ we have $\mathbb{P} = \mathbb{P}_1 \circ \ldots \circ \mathbb{P}_r$ and, therefore, $\mathbb{P}$ is decomposable.

We now show the existence and uniqueness of the irreducible representation. First note that for the space $\mathbb{P} = \{U, X_1, \ldots, X_r\}, r > 1$, its representation can include the space $\{U, X\}$ only if

$X \in \{X_1, \ldots, X_r\}$. This is due to the fact that if $\{U, X\}$ is an element of the decomposition, then $X \in \mathbb{P}$. Let $\{U, Y_1, \ldots, Y_k\}$ be the set of all $\cap$-generators of the space $\mathbb{P}$. By definition, the $\cap$-generator space $\{U, Y_j\}$ ($j = 1, \ldots, k$) must belong to each representation. Moreover, for any $\cap$-generator, by definition, $Y_j \notin \mathbb{P}_1 \circ \ldots \circ \mathbb{P}_{j-1} \circ \mathbb{P}_{j+1} \circ \ldots \circ \mathbb{P}_k$, therefore, $\mathbb{P}_j \nleq \mathbb{P}_1 \circ \ldots \circ \mathbb{P}_{j-1} \circ \mathbb{P}_{j+1} \circ \ldots \circ \mathbb{P}_k$. Finally, we note that $\mathbb{P}_1 \circ \ldots \circ \mathbb{P}_k = \mathbb{P}$, therefore $\mathbb{P}_1 \circ \ldots \circ \mathbb{P}_k$ is an irreducible representation of $\mathbb{P}$ and, moreover, it is unique because the set of $\cap$-generators is unique.

We note several properties of the lattice **P** , which follow from Theorem 5.

Definition 6. A *point (atomically generated)* lattice is such a lattice, all of which elements are the union of points (atoms).

Implication 1. The lattice **P** is a point (atomically generated). Indeed, the points (atoms) in the lattice **P** are spaces consisting of two categories, one of which is U, the operation of mixing replaces the of joining spaces, and any linguistic space is obtained by mixing some points (atoms).

Implication 2. The length of the lattice **P** is equal to the number of atoms (indecomposable spaces). As noted, the length lattice **P** is 2^n, where n is the number of elements in U, so the number of indecomposable elements in **P** is also 2^n.

Definition 7. Let $\mathbb{P}_1$, $\mathbb{P}_2 \in$ **P** and $\mathbb{P}_2 \leq \mathbb{P}_1$. We call $\mathbb{P}_3$ quotient by dividing $\mathbb{P}_1$ by $\mathbb{P}_2$ ($\mathbb{P}_3 = \mathbb{P}_1/\mathbb{P}_2$), if:

a) $\mathbb{P}_2 \circ \mathbb{P}_3 = \mathbb{P}_1$;

b) from $\mathbb{P}_2 \circ \mathbb{P}_4 = \mathbb{P}_1$ it follows that $\mathbb{P}_3 \leq \mathbb{P}_4$.

Proposition 8. For any $\mathbb{P}_1$, $\mathbb{P}_2 \in$ **P** and $\mathbb{P}_2 \leq \mathbb{P}_1$ there exists a quotient $\mathbb{P}_3$.

Proof. If $\mathbb{P}_1$ is indecomposable, then either $\mathbb{P}_1 = \mathbb{P}_2$ or $\mathbb{P}_2 = \{U\}$. In any case, $\mathbb{P}_1 = \mathbb{P}_3$. Let $\mathbb{P}_1$ be decomposable. If $\mathbb{P}_2 = \{U\}$, then $\mathbb{P}_1 = \mathbb{P}_3$. Let $\mathbb{P}_2 \neq \{U\}$ and $\{X_1, \ldots, X_k\}$ be the $\cap$-generators of $\mathbb{P}_1$. Then $\mathbb{P}_{X1} \circ \ldots \circ \mathbb{P}_{Xk}$ is the only irreducible representation of $\mathbb{P}_1$ ($\mathbb{P}_{Xi} = \{U, X_i\}$; $i = 1, \ldots, k$). Let $\mathbb{P}_{Y1} \circ \ldots \circ \mathbb{P}_{Yr}$ be an irreducible representation of $\mathbb{P}_2$. Remove from the irreducible representation for $\mathbb{P}_1$ the elements included in both representations. The remaining mixing is denoted by $\mathbb{P}_3$. Since $\mathbb{P}_2 \leq \mathbb{P}_1$ and $\mathbb{P}_3 \leq \mathbb{P}_1$ we get $\mathbb{P}_2 \circ \mathbb{P}_3 \leq \mathbb{P}_1$. However, by construction $\mathbb{P}_3$ we have $\mathbb{P}_1 \leq \mathbb{P}_2 \circ \mathbb{P}_3$, therefore, $\mathbb{P}_2 \circ \mathbb{P}_3 = \mathbb{P}_1$. If $\mathbb{P}_1 = \mathbb{P}_2 \circ \mathbb{P}_4$, then at least $\mathbb{P}_4$ must contain in its irreducible representation of space, including the generators that are part of the irreducible representation $\mathbb{P}_3$, and therefore, $\mathbb{P}_3 \subseteq \mathbb{P}_4$.

Remark 9. Returning to the discussion on the definition of a sign, it should be noted that a slightly strange definition of a sign, which requires $(x, 0) \in S_A$ for any x from the universe, made it possible to obtain a systematic lattice with a number of remarkable properties. Another definition of a sign leads to a different, poorer lattice [2]. Since we are dealing with an abstract definition of a sign, we have the opportunity to make a choice, especially since it does not affect the final result. Another thing when it comes to real characteristics. For example, for indirect additions, a different, less convenient lattice is obtained, and in this case we cannot act in the same way as with signs.

2. *Dimension of linguistic space.*

We must understand that the main thing is the transition to the forecast. In other words, the main thing is the transformation of somehow structured data into a model. We will return to this issue a little bit later, but for now, we note that from the point of view of the forecast, it does not matter whether the spaces are indecomposable if they participate in an irreducible representation. The main

thing is to find the signs that give such partitions of the universe that lead to the decomposition of the space according to these signs. This means that it is necessary to find such partitions U so that:

a) in each partition, the partition classes do not intersect and coincide with some ∩-generators of the space;

b) the classes of all partitions cover the set of all ∩-generators space.

So, let R be a relation on the set of ∩-generators $\mathbf{O} = \{X_1, \ldots, X_k\}$ of the space $\mathbb{P}$, defined as follows:

a) for any $i = 1, \ldots, k$ $(X_i, X_i) \in R$;

b) if $i \neq j$, $(X_i, X_j) \in R$ if and only if $X_i \cap X_j = \varnothing$.

The ratio R is tolerance. Recall that the tolerance class is the maximum set of elements that are pairwise related to each other by the relation R. In accordance with the foregoing, to select the number and type of multivalued signs, you must select the smallest number of tolerance classes R covering the entire set $\mathbf{O}$. In various aspects of tolerance are considered in detail, the concept of basis is introduced [3]. It should be noted that the concept of covering classes of tolerance and the concept of the basis of tolerance turn out to be significantly different, since the basis is also required to preserve the entire relationship. In general, there may be several coverages with a different number of tolerance classes within them.

<u>Definition 10.</u> *The dimension of the linguistic space* is the smallest number of tolerance classes on the set of ∩-generators covering this set.

In other words, dimensionality is the smallest number of signs that can be used to describe linguistic space. For example, any equivalence generates a one-dimensional space, as for its identification is enough to have one sign with the number of nonzero meanings equal to the number of equivalence classes. Actually, we call the dimension of linguistic space the minimum number of one-dimensional spaces sufficient for its description. Or, for example, the dimension of a human face is equal to the number of signs in the identikit plus deviations from the default position. The coordinates of my face are: 16th nose, shifted 2mm to the right, 45th lips, etc. If you set the meanings of all these coordinates, you will see me.
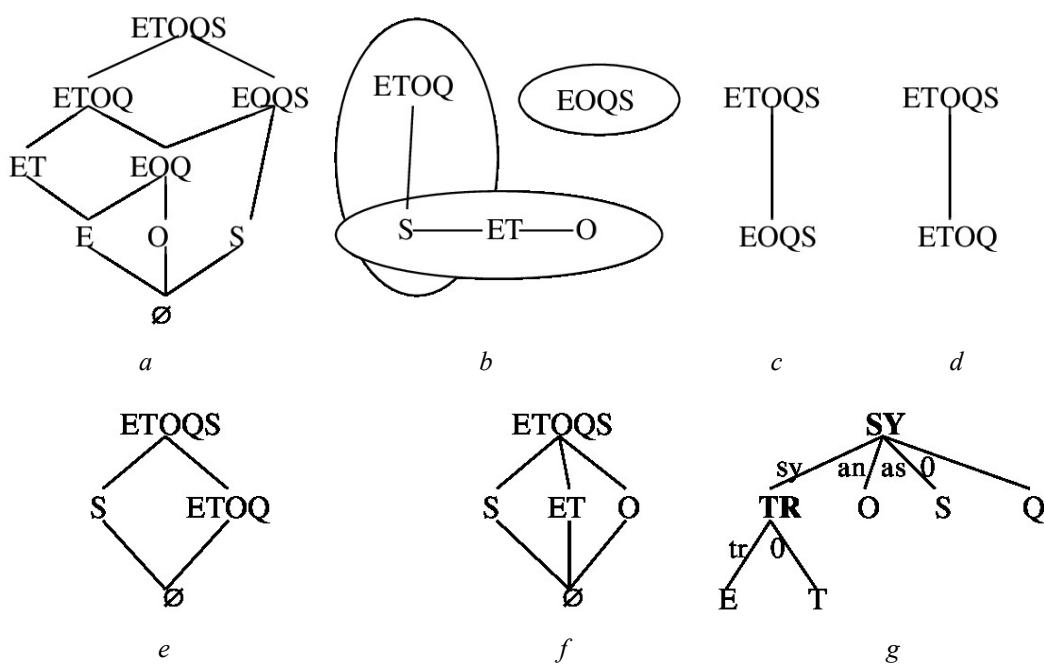

Fig. 2. Signs spaces for the structure of binary relations

Let us turn to the example of the linguistic space from, shown in fig. 2*a* [4]. In this linguistic space there are five ∩-generators: {ETOQ}, {EOQS}, {ET}, {O}, {S}. There is no universe in this list, because we are not interested in signs with one zero meaning. In fig. 2*b*, the lines are connected to tolerant ∩-generators (reflexive arcs are not shown), the ellipses cover the tolerance classes.

From fig. 2*b* it can be seen that there are three classes of tolerance on the set of ∩-generators. All three classes are necessary to cover the set of ∩-generators. Thus, there is only one option to choose the number and type of signs, so that the dimension of this space is three. For further discussion of this example, I will reveal the nature of the objects in this space. The objects of the space are different binary relations: E-equivalence, T – tolerance, O – order, Q – quasi-order, S – strict order. So, one of the signs must have one nonzero meaning, which have objects E, O, Q, S and the object T does not have it. This sign is 'transitivity' with meanings: 0 and 'transitive'. The meaning 'transitive' will not have only the object T. The space for this sign is shown in fig. 2*c*. The following sign corresponds to another tolerance class consisting of two ∩-generators: {ETOQ} and {S}. Obviously, this partition corresponds to the sign of 'reflexivity'. It is important to note that the members of this class ∩-generator cover the whole universe, breaking it into two parts. Therefore, the sign for this tolerance class can have two or one nonzero meaning depending on whether we are going to send one partition class to the zero-value region with the formulation 'the object does not have the reflexivity sign'. So, the first version of the scale of sign 'reflexivity' has a meaning of 0 and the meaning of 'reflexive'. In this case, the object S does not have the sign (fig. 2*d*). In the second variant, the sign has the values 0, 'reflexive' and 'antireflexive'. In this case, the object S matches the meaning of the "anti-reflective" (fig. 2*e*). ∩-generators of the third class of tolerance do not cover the whole universe, so there is only one option: three non-zero values of the sign. This sign obviously corresponds to symmetry and has four values: 0, 'symmetric', 'anti-symmetric', 'asymmetric'. Under anti-symmetric relation here is defined as the non-symmetrical relation in which each element is in this relation with itself (reflexivity). The asymmetric relation is understood as the non-symmetrical relation which does not allow the reflexivity of elements. And again because of redundancy, if we used for the sign of 'reflexivity' the variant fig. 2*e*, which is already presented ∩-generator S, it would be possible in the sign of 'symmetry' to remove the meaning of 'asymmetrically'. The space for the sign 'symmetry' is shown in fig. 2*f*.

So, the linguistic space of fig. 2*a* is described by minimum with three signs with specific meanings, which, due to redundancy, can vary slightly. Fig. 2*g* depicts the recognition tree of binary relations. The bold typeface identifies the abbreviated names of the signs being analyzed, in small letters – the meanings of these signs; the leaves of the tree are marked by relationships from the universe. At first glance, it seems strange that the problem of recognizing relations here is solved only with the use of two signs, while the dimension of space is three. This is the difference between the task of recognizing objects and the task of identifying categories of a certain linguistic space. The identification of space categories is aimed to identify all categories, and not just recognizing species. However, the task of recognition is an essential part of identification. We will return to this issue a little later.

Up to this point, all reasonings, starting with the axioms of categorization, were of a purely formal nature. Even the signs, their number, the number of their meanings and the correspondence of the meanings to the objects of the universe are the consequence of the decomposition of spaces and do not require the presence of an 'external world'. But there comes a time when the calculated signs should

get real content. In the considered example, everything happened was just fine, but only because I initially moved from the signs when forming the linguistic space. But how to find in a real object the signs with calculated properties? Here the division operation can help.

Suppose we learned (guessed, spied) that there is a sign of 'symmetry', which has four meanings: 0, "symmetrical", "antisymmetric", "asymmetrical". Thus, on this sign, we get the space shown in fig. 3a.
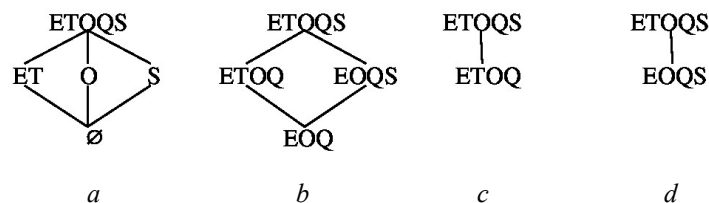

Fig. 3. Division in the decomposition of linguistic spaces

This sign splits the universe into four non-intersecting classes (the object Q has no sign 'symmetry'). All categories of this space are included in the space in fig. 2a, so it is less than the original. In accordance with proposition 8, there exists a quotient of dividing the original space into the space of fig. 3a. The quotient is presented in fig. 3b. This quotient is decomposed into two spaces, presented in fig. 3c and 3d. The space in fig. 3d corresponds to the "transitivity" sign, and the space in fig. 3c – the sign of "reflexivity". In this case, the sign "reflexivity" has only two meanings: 0 and "reflexively". So, the division operation allows us to simplify the decomposition of spaces when some information about a part of (real) signs is already available.

<u>Example 11.</u> Consider the use of the Theorem 5 with the example of the trophic ratio of freshwater fish. Suppose a certain observer, sitting on the bank of a pond, observed the hunting of some fish for others, and recorded these observations as a relation in fig. 4a. Here in the left and right columns are fishes from the pond: pike, perch, crucian, and roach. Two fishes are connected with the line, if the fish on the left ate the fish on the right. According to these observations, one can construct space and co-space (fig. 4b) as it was described in [4]. Since space and co-space are defined on the same universe, they can be mixed into one space (fig. 4c). Moreover, as noted in, dual isomorphism transforms into dual epimorphism [4]. In fig. 4c, the action of the verb 'eat' is shown by arrows, so now, for example, two categories PPe and Pe are mapped into one category CRPe. Note also that one of the types of CR in this space includes two fishes at once: crucian and roach, since from the point of view of 'eating' they are indistinguishable. By the method described above, we can determine the dimension of this space and select two signs, the spaces of which are depicted in fig. 4d and fig. 4e.
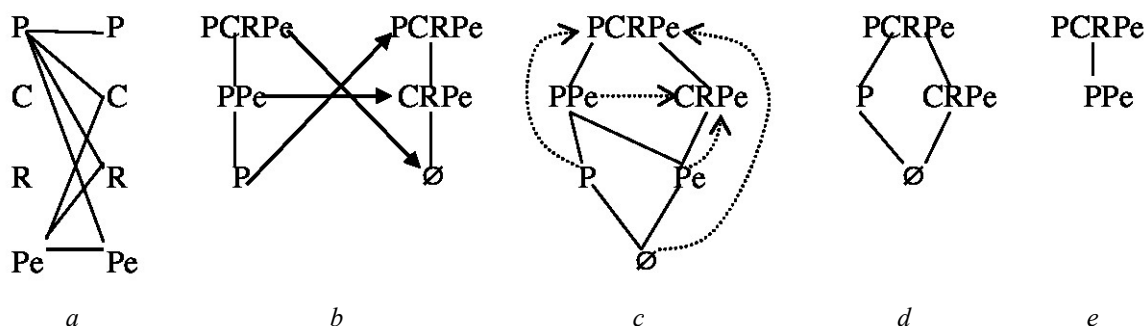

Fig. 4. Decomposition of linguistic space

Now it is necessary to identify the calculated signs with the real properties of these fishes. For the sign in fig. 4*e*, I easily found such a property – this is the presence of teeth. Pike and perch have teeth, and crucian and roach do not have this sign. For another sign, the reader may suggest some suitable property, but nothing good came to my mind except size. In other words, the fish is 'big' (belongs to category P), if its size exceeds 30 cm, and 'small' (belongs to category CRPe), otherwise. Now, for example, if we meet a big and toothy fish, then we will predict its aggressive behavior towards all fishes in the pond. We have a forecast! Moreover, if we are confronted with a large pike perch, we will also predict its aggressive behavior towards all fish, although we have not previously observed the behavior of pike perch! Our structure has become a model. It is clear that this is an extremely important evolutionary conquest. It allows you to predict and act even in a collision with new objects. So, systematization turns interconnected spaces into a model due to a transition consistent with spaces to the definition of categories through signs.

Of course, there are mistakes in modeling. So, probably, it happened when the first Europeans found themselves on the Amazon and were confronted with piranhas. In principle, the forecast for piranhas is similar to perch: a small, toothed fish. Probably, some pioneers strongly regretted their prediction. As a result, we had to change the original trophic relation and create new categories with new verbal connections, and then find new signs for these new categories. In the end, the model was adjusted and again began to give correct predictions.

3. *Centaurs*.

This, at first glance, a bit strange section, which seems to stand apart, is essential for understanding thinking in general and language in particular. In the previous section, we looked at what happens to spaces when decomposed into simple components – signs. Now we could give another definition of a sign: signs are the allocated spaces into which linguistic spaces are decomposed. We chose the atoms as signs because the systematization space is an atomically generated lattice (Implication 1 of Theorem 5). Then we added some more gratings to them to reduce the number of decomposition elements. But our choice is only our choice. It is possible that the reader will consider it sensible to make another choice. In this section, we digress from the concrete embodiment of the type of signs. The main thing is that the spaces are decomposed in them. And within the framework of this view, we will be interested in what happens to the signs when they are mixed in restoring space?

To answer this question, we will have to turn to data dependencies, though in some not quite traditional form. For example, mutual dependence, being by its nature the simplest dependence of a compound, in our case acquires a special "layered" character [5]. Later in this section we will use the template definition of dependencies.

<u>Definition 12.</u> Let $A = A_1 \ldots A_n$, $B \subseteq A$, $C = A–B$. A sign B is called potentially independent (in the context of A) if for any $x \in U$ from $(x, b, c) \in S_A$, $(x, b', c') \in S_A$ and $(x', b, c') \in S_A$ follows $(x, b, c') \in S_A$.

The definition of potential independence corresponds exactly to the definition of mutual dependence [6].

<u>Definition 13.</u> Let $A = A_1 \ldots A_n$, $B \subseteq A$, $C = A–B$. A sign B is called *independent* (in the context of A) if for any $x \in U$ from $(x, b, c) \in S_A$ and $(x, b', c') \in S_A$ it follows $(x, b, c') \in S_A$.

The definition of independence corresponds to the definition of a multivalued dependence [6].

For mutual dependence (denoted by $\sim$), which corresponds to potentially independent signs, the following properties are satisfied [5]:

1. If U$\sim$B, then B$\sim$U.

2. If U$\sim$B, then U$\sim$C, so instead of U$\sim$B you can write U$\sim$B/C.

3. All three mutual dependencies are equivalent: U$\sim$B/C, B$\sim$U/C, C$\sim$U/B.

4. If the mutual dependence U$\sim$B/C is fulfilled in S_A, then U$\sim$B'/C' is fulfilled in S_A[UB'C'], where A' $\subseteq$ A, B' $\subseteq$ B. Here S_A[UB'C'] denotes the projection of the SA relation on the UB'C' attributes. The last fifth property defines the following.

<u>Lemma 14.</u> If U$\sim$B/C in S_A[UBC], and U$\sim$D/BC in S_A, then U$\sim$B/CD, where A = BCD and B, C, D are not mutually intersecting.

Proof. Let (x, b, c, d) $\in$ S_A, (x, b', c', d') $\in$ S_A and (x', b, c', d') $\in$ S_A. Since (x, b, c) $\in$ S_A[UBC], (x, b', c') $\in$ S_A [UBC], (x', b, c') $\in$ S_A[UBC] and U$\sim$B/C in S_A[UBC], then (x, b, c') $\in$ S_A[UBC]. Therefore, there is d" such that (x, b, c', d") $\in$ S_A. Since (x, b', c', d') $\in$ S_A and (x', b, c', d') $\in$ S_A, then by virtue of U$\sim$D /BC (x, b, c', d') $\in$ S_A. Thus, in S_A it is performed U$\sim$B/CD.

The meaning of Lemma 14 is that if we connect the systematization relations in such a way that mutual dependence is fulfilled, then all potentially independent signs that are already the part of the resulting relation will still maintain potential independence. Conversely, property 4 allows us to state that if all signs are potentially independent, then the projections of the systematization relation on a part of the signs retain potential independence.

<u>Lemma 15.</u> Let $S_{Ai} = S_A[UA_i]$, where A = $\bigcup_{i=1}^{n} A_i$, $A_i \cap A_j = \varnothing$ for any $i \neq j$; $i, j = 1, \ldots, n$. Signs $A_1, \ldots, A_n$ are potentially independent if and only if for any tuple of signs $(a_1, \ldots, a_n) \in S_A[A]$ $(a_1, \ldots, a_n)^\Delta = \bigcap_{i=1}^{n} a_i^\Delta$, where $(a_1, \ldots, a_n)^\Delta = \{x \in D(U) \& (x, a_1, \ldots, a_n) \in S_A\}$; $a_i^\Delta = \{x \in D(U) \& \&(x, a_i) \in S_{Ai}\}$.

Proof. We prove by induction on the number of signs. For n = 2, from $S_{Ai} = S_A[UA_i]$, it follows that $(a_1, a_2)^\Delta \subseteq a_1^\Delta \cap a_2^\Delta$ for any $(a_1, a_2) \in S_A[A]$. Now let $x \in a_1^\Delta \cap a_2^\Delta$. Then for some a'$_1$, a'$_2$ we have $(x, a_1, a_2') \in S_A$, $(x, a_1', a_2) \in S_A$ and $(a_1, a_2) \in S_A[A]$. From here, by virtue of potential independence, we get $(x, a_1, a_2) \in S_A$ and $x \in (a_1, a_2)^\Delta$, so that, $(a_1, a_2)^\Delta = a_1^\Delta \cap a_2^\Delta$. Conversely, let $(x, a_1, a_2) \in S_A$, $(x, a'_1, a'_2) \in S_A$ and $(a_1, a'_2) \in S_A[A]$. Then $x \in a_1^\Delta$, $x \in a'^\Delta_2$, which means $x \in (a_1, a'_2)^\Delta$ and signs A_1, A_2 are potentially independent.

Let the lemma be true for all $2 \leq n \leq k$ and let n = $k+1$. Denote by A' the set $\bigcup_{i=1}^{k} A_i$. By virtue of the fourth property of mutual dependence, the signs $A_1, \ldots, A_k$ are potentially independent in S_A[UA']. By the induction hypothesis, for every $(a_1, \ldots, a_k) \in S_A[A']$ the equality $(a_1, \ldots, a_k)^\Delta = \bigcap_{i=1}^{k} a_i^\Delta$ is fulfilled. Since the sign A_{k+1} is potentially independent in S_A, then for everyone $(a_1, \ldots, a_{k+1}) \in S_A[A]$ is true $(a_1, \ldots, a_{k+1})^\Delta = (a_1, \ldots, a_k)^\Delta \cap a_{k+1}^\Delta = \bigcap_{i=1}^{k+1} a_i^\Delta$. Conversely, suppose that one of the signs, say, A_{k+1} is not potentially independent, but for every $(a_1, \ldots, a_{k+1}) \in S_A[A]$, $(a_1, \ldots, a_{k+1})^\Delta = \bigcap_{i=1}^{k+1} a_i^\Delta$. Then there are $(x, a_1, \ldots, a_{k+1}) \in S_A$, $(x, a'_1, \ldots, a'_{k+1}) \in S_A$, $(a'_1, \ldots, a'_k, a_{k+1}) \in S_A[A]$, but $(x, a'_1, \ldots, a'_k, a_{k+1}) \notin S_A$. Hence $x \in a_{k+1}^\Delta$, $x \in a'^\Delta_1, \ldots, x \in a'^\Delta_k$ and, therefore, $x \in \bigcap_{i=1}^{k} a'^\Delta_i \cap a_{k+1}^\Delta = (a'_1, \ldots, a'_k, a_{k+1})^\Delta$, which contradicts

the assumption that there is no potential independence for A_{k+1}. Thus, signs $A_1, \ldots, A_{k+1}$ are potentially independent.

Lemma 15 clarifies the meaning of the concept of potential independence. It justifies a parallel independent analysis of the signs of the objects of the universe, followed by their general mixing into one sign through conjunction. Thus, the potential independence of signs is equivalent to the possibility of conjunctive assessment of signs. However, from the potential independence it does not follow that with a conjunctive analysis of the meanings of the signs, any category of linguistic space can be identified.

<u>Definition 16.</u> Let S_A be a systematization relation, $S_{Ai} = S_A[UA_i]$ $A = \bigcup_{i=1}^{n} A_i$, $A_i \cap A_j = \varnothing$ for any $i \neq j$; $i, j = 1, \ldots, n$; $\mathbb{P}, \mathbb{P}_1, \ldots, \mathbb{P}_n$ are the spaces according to the corresponding systematization relations. The relation S_A is called *covering* if $\mathbb{P}_1 \circ \ldots \circ \mathbb{P}_n \subseteq \mathbb{P}$.

Obviously, if S_A is covering, then $\mathbb{P}_i \subseteq \mathbb{P}$ for any $i = 1, \ldots, n$. Conversely, $\mathbb{P}_i \subseteq \mathbb{P}$ for any $i = 1, \ldots, n$, then $\mathbb{P}_1 \circ \ldots \circ \mathbb{P}_n \subseteq \mathbb{P}$ and, therefore, S_A is covering.

<u>Theorem 17.</u> Let S_A be a systematization relation, $S_{Ai} = S_A[UA_i]$ $A = \bigcup_{i=1}^{n} A_i$, $A_i \cap A_j = \varnothing$ for any $i \neq j$; $i, j = 1, \ldots, n$. If S_A is a covering and the signs $A_1, \ldots, A_n$ are potentially independent, then $\mathbb{P} = \mathbb{P}_1 \circ \ldots \circ \mathbb{P}_n$.

Proof. Let X be a $\cap$-generator in $\mathbb{P}$. In accordance with there exists a tuple $(a_1, \ldots, a_n) \in S_A[A]$ such that $(a_1, \ldots, a_n)^\Delta = X$. By Lemma 15, $X = \bigcap_{i=1}^{n} a_i^\Delta$ [4]. But $\bigcap_{i=1}^{n} a_i^\Delta \in \mathbb{P}_1 \circ \ldots \circ \mathbb{P}_n$, since $a_i^\Delta \in \mathbb{P}_i$ for each $i = 1, \ldots, n$. Thus, each generator of $\mathbb{P}$ belongs to $\mathbb{P}_1 \circ \ldots \circ \mathbb{P}_n$, which means that $\mathbb{P} \subseteq \mathbb{P}_1 \circ \ldots \circ \mathbb{P}_n$. Since S_A is a covering, $\mathbb{P} = \mathbb{P}_1 \circ \ldots \circ \mathbb{P}_n$.

<u>Implication</u>. If the signs $A_1, \ldots, A_n$ are independent, then the relation S_A is covering.

Proof. So, let the signs be independent, but S_A is not covering. Then in $\mathbb{P}_1 \circ \ldots \circ \mathbb{P}_n$ there is a category X not belonging to $\mathbb{P}$. Of course, in this case, the $\cap$-generators cannot coincide in $\mathbb{P}_1 \circ \ldots \circ \mathbb{P}_n$ and $\mathbb{P}$. Since the proof of Theorem 17 implies that $\mathbb{P} \subseteq \mathbb{P}_1 \circ \ldots \circ \mathbb{P}_n$, then in $\mathbb{P}_1 \circ \ldots \circ \mathbb{P}_n$ there is a $\cap$-generator X' that is absent in $\mathbb{P}$. In accordance with, there are a_i ($i = 1, \ldots, n$), such that $X' = \bigcap_{i=1}^{n} a_i^\Delta$ [4]. Thus, for any $x \in X'$, $(x, a_i) \in S_{Ai}$. Since the signs are independent, in S_A there are n multi-valued dependencies and S_A is equal to the connection of all S_{Ai}. But then for any $x \in X'$ $(x, a_1, \ldots, a_n) \in S_A$, which means, $X' \in \mathbb{P}$, which contradicts the assumption. Thus, S_A is covering.

<u>Remark 18.</u> From the condition of Theorem 17 one cannot exclude the requirement that S_A be a covering one. In fig. 5a shows the relation $S_{PP'}$, where $D(U) = \{a, b, c\}$.
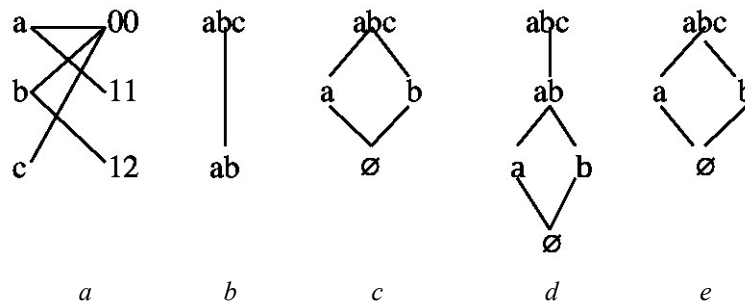

Fig. 5. Illustration of the characteristics of independence and potential for independence of signs

Fig. 5*b*, 5*c*, 5*d*, e shows $\mathbb{P}_1$, $\mathbb{P}_2$, $\mathbb{P}_1 \circ \mathbb{P}_2$ and $\mathbb{P}$ respectively. Although in relation to $S_{PP'}$, the signs of P and P' are potentially independent, but $\mathbb{P} \neq \mathbb{P}_1 \circ \mathbb{P}_2$, because $S_{PP'}$ is not a covering one. The reason for this is a combination meanings of signs (tuple 10), which is absent in the original relation. In principle, this can be avoided if in the process of identifying objects on signs the initial space obtained on the basis of observing the behavior of objects is controlled, and on the basis of this, 'inadmissible' combinations of signs are rejected.

Remark 19. Independence of signs is not equivalent to the equality $\mathbb{P} = \mathbb{P}_1 \circ \ldots \circ \mathbb{P}_n$, since covering can be a relation with potentially independent signs. For example, a relation consisting of two tuples {(x, a1, в1), (x, a2, в2)} has potentially independent signs and is simultaneously covering, but its signs do not have multivalued dependencies. For this reason, if the equality $\mathbb{P} = \mathbb{P}_1 \circ \ldots \circ \mathbb{P}_n$ is satisfied, then it means that either the signs are independent or the signs are potentially independent and the ratio covering. It has already been noted that no matter how long we fly in mathematical abstractions, it is time to return to the world and compare the abstract signs of something real from life, as it was in example 11. From this section it becomes clear that whenever we use combinations of signs we secretly consider them independent (potentially independent), or, what is the same, we believe the existence of a multivalued (mutual) dependence of each sign on the objects of the universe. But are the signs found independent? After all, this is not what guided us in their choice when we looked at the world. Therefore, it is important to understand the boundaries of a safe retreat from the property of independence.

The safest easing of the require for independence is the potential independence of signs in combination with the covering relation. So, when we analyze signs and use their combinations, on the one hand, we mix their spaces to $\mathbb{P}_1 \circ \ldots \circ \mathbb{P}_n$, and on the other hand, we joined private systematization relations S_{Ai} by the attribute U, as if there are multivalued dependencies. Denote the product $\mathbb{P}_1 \circ \ldots \circ \mathbb{P}_n$ by $\mathbb{P}'$. For each $\cap$-generator from $\mathbb{P}'$ according to there exists a certain tuple of signs values $(a_1, \ldots, a_n)$ for which the set $(a_1, \ldots, a_n)^\Delta$ coincides with this generator [4]. Choose for each $\cap$-generator one such tuple and combine products $(a_1, \ldots, a_n)^\Delta \times \{(a_1, \ldots, a_n)\}$ in one new relation S_A'. By definition of this relation $S_A' \subseteq \bowtie S_{Ai}$ (for all $i = 1, \ldots, n$; $\bowtie$ – the operation of joining relations on the general attribute U) and S_A' generates the same space $\mathbb{P}'$. Since for each $(a_1, \ldots, a_n)$ of S_A' [A] by construction is performed $(a_1, \ldots, a_n)^\Delta = \bigcap_{i=1}^{n} a_i^\Delta$, then by virtue of Lemma 15 all signs are potentially independent. Recall that D (U) is also a $\cap$-generator and for it the tuple (0, ..., 0) will enter S_A'. Thus S_A' is covering and for it $\mathbb{P}' = \mathbb{P}_1 \circ \ldots \circ \mathbb{P}_n$. Since for each $\cap$-generator in S_A' is only one tuple of the values of signs that is the smallest relation for which the equality of theorem 17 holds true. Thus, the safe range of deviations from the independence of the signs extends from S_A' to the joining of systematic relations of the signs.

The next weakening is the rejection of the property of cover ability. In this case, it seems that it is necessary to control the permissible combinations of signs in relation to the original linguistic space, as the mixing of spaces can generate invalid combinations. But how do you know if some combination is acceptable? After all, we do not know the 'correct' systematized relation. This is a problem. In any case, it is impossible to retreat further, since the violation of potential independence by virtue of Lemma 15 blocks the possibility of conjunctively combining the results of the analysis of individual features, and hence mixing the corresponding spaces.

So, we always have to assume the independence of the real signs. If we are wrong in this hypothesis, we may be helped by the potential independence of features and the property of

coverability, but we may not know about it. It is because of the independence of a certain way the selected signs determine the dimension of linguistic space like orthogonality in ordinary spatial representations.

Remark 20. So, the signs are always associated with multivalued or mutual dependencies, and thus, with the decomposition of objects or independence. This is a consequence of the nature of signs, arising from the method of decomposition of space. In example 11, the sign 'teeth' is clearly a consequence of the decomposition of the fish, which allocates such a part of it as teeth. But the 'size' sign is not an obvious result of decomposition. However, as a 'size' is a sign, then it is independent of having teeth: toothy fish can be both small and large, just like toothless fish. Here is one of the dividing lines between signs and parameters.

Remark 21. The hypothesis of the independence of signs (and the associated decomposition of objects) is a universal tool of consciousness, although is not always aware of them. Any consciousness uses this hypothesis. In the culture of all peoples of the world, we find unacceptable combinations of features arising from the violation of the hypothesis of independence, or rather from its incorrect usage. All these dragons, centaurs, mermaids, sphinxes, angels and other mythical creatures are the result of unlimited use of the hypothesis of independence of signs. But we know there are no centaurs. How can we be sure of that? Because subconsciously we understand that the independence of signs acts only within the concept of 'people' or 'horses', but does not act on the united universe. This reminds of reminiscent of the futility of the interspecific interbreeding of animals, and only unbridled imagination rises to the surface of the mind these devils with hooves, tails, fur and horns.

Another example of a centaur can be obtained from the example of the space of different binary relations. As signs we have identified 'reflexivity' and 'symmetry'. These signs are independent and therefore there must be a reflexive and asymmetric relationship. But this relationship cannot be, because the asymmetry does not allow the reflexivity. Here again, there is an incorrect use of independence: the independence of the signs of 'reflexivity' and 'symmetry' is performed only in the presence of some relation x, and this relationship in our universe does not exist.

Furthermore, for example, in the described space, the $\cap$-generator {ET} is completely described by the value 'symmetrically'. In other words, any symmetric relation here is tolerance or equivalence. But within the broader universe, there are symmetric relationships outside of this list. Therefore, tolerance is a reflexive and symmetrical relationship at the same time. So, in another universe, the sign of 'symmetry' expands and does not describe any category by itself. By analogy, people often choose such broad signs to create the necessary signs using conjunctions. In it was mentioned the case of the erroneous definition of Linnaeus primate squad, for which he chose the signs of 'five-toed limb' and 'two mammary glands' [4]. Both signs are broad and do not correspond separately to any category of classification of animals. But the conjunction of these signs is a new sign, which according to Linnaeus could well be suitable for the Primate squad. Such conjunctions often replace the calculated signs, especially since according to the results of this section, the signs themselves are always analyzed conjunctively.

Centaurs draw our attention to the fact that universums are chosen for a reason. They are changeable and are themselves categories in some spaces. This leads to the fact that within a certain category another space is constructed from a certain space, in which the initial category is considered as a universe. Such nesting (recursion in depth) can be propagated deeply enough. In addition, there

may be a recursion in depth for a sign that is represented through other particular signs. For example, 'head' is one of the signs in the decomposition of the whole person, but this sign itself is decomposed into signs of the identikit. This can be continued until we have exhausted the resource of independence.

This article took an important next step from the axiomatic approach in categorization to its continuation in identification [4]. Since the definitio is abstract, until now the linguistic data model does not impose any requirements on the world, and vice versa. Identification itself is an intermediate link between the world with its limitations and an abstract model, since signs, being elements of the decomposition of linguistic spaces (Theorem 5), can always be formed, regardless of the actual initial data expressed in parameters. In the future we will see that the very emergence of language would not have been possible without this construction, which is guaranteed to exist outside the limits of the senses.

Signs resemble parameters, but their nature is completely different. Their nature is associated with the decomposition of spaces. In addition, as it turned out, they are also associated with the decomposition of the concepts themselves into parts and with the use of ideas about data independence. Thus, the signs create a movement 'in depth' concepts.

Conclusion. The appearance of signs may seem strange, given that the linguistic model itself is built on real observation data, and movement in the opposite direction for some reason involves the use of some new artificial objects. But if we recall that the task of modeling is to forecast behavior, it becomes clear: without a certain structuring of the input information flow, it is impossible to make a quick and effective forecast.

The concept of a sign is introduced in this article, and it is shown that many such signs always exist. If the concept of a sign is expanded a little, then there is a slight variability in the choice of the composition of signs and the ability to determine the dimension of linguistic space as the smallest number of independent features that fully describe the concepts of space.

Finally, the article shows that signs are inherently independent (potentially independent), which creates the basis for decomposing objects and analyzing data in parts.

REFERENCES

1. Birkhoff, G. (1984), *Lattice Theory*, in Skornyakov, L.A. (ed.), Transl. by Salii V.N., Nauka, Moscow, USSR.

2. Polyakov, O.M. (1988), "About systematization", *NTI,* ser. 2, no. 12, pp. 21–28.

3. Shreider, Yu.A. (1971), *Ravenstvo, skhodstvo, poryadok* [Equality, similarity, order], Nauka, Moscow, USSR.

4. Polyakov, O.M. (2019), "Linguistic Data Model for Natural Languages and Artificial Intelligence. Part 1. Categorization", *DISCOURSE*, vol. 5, no. 4, pp. 102–114. DOI: 10.32603/2412-8562-2019-5-4-102-114.

5. Nicolas, J.M. (978), "First Order Logic Formalization for Functional, Multivalued and Mutual Dependencies", *SIGMOD '78: International Conf. on Management of Data*, Austin, Texas, 31 may – 02 june, 1978, pp. 40–46. DOI: 10.1145/509252.509269.

6. Fagin, R. (1977), "Multivalued Dependencies and a New Normal Form for Relational Databases", *ACM TODS*, vol. 2, iss. 3, pp. 262–278. DOI: 10.1145/320557.320571.

Information about the authors.

Oleg M. Polyakov – Can. Sci. (Engineering) (1982), Associate Professor at the Department of Information Technology of Entrepreneurship, Saint-Petersburg State University of Aerospace Instrumentation, 67 lit. A Bol'shaya Morskaya str., St Petersburg 190000, Russia. The author of 30 scientific publications. Areas of expertise: linguistics, artificial intelligence, mathematics, database design theory, philosophy. E-mail: road.dust.spb@gmail.com

СПИСОК ЛИТЕРАТУРЫ

1. Биркгоф Г. Теория решеток / пер. В. Н. Салий; под ред. Л. А. Скорнякова. М.: Наука, 1984.

2. Поляков О. М. О систематизации // НТИ. Сер. 2. 1988. № 12. С. 21–28.

3. Шрейдер Ю. А. Равенство, сходство, порядок. М.: Наука, 1971.

4. Polyakov O. M. Linguistic Data Model for Natural Languages and Artificial Intelligence. Part 1. Categorization // DISCOURSE. 2019. Vol. 5, No. 4. P. 102–114. DOI: 10.32603/2412-8562-2019-5-4-102-114.

5. Nicolas J. M. First Order Logic Formalization for Functional, Multivalued and Mutual Dependencies // SIGMOD '78: International Conf. on Management of Data, Austin, Texas, 31 may – 02 june, 1978 / Univ. of Texas at Austin. Austin, Texas, 1978. P. 40–46. DOI: 10.1145/509252.509269.

6. Fagin R. Multivalued Dependencies and a New Normal Form for Relational Databases // ACM TODS.1977. Т. 2, No. 3. P. 262–278.

Информация об авторах.

Поляков Олег Маратович – кандидат технических наук (1982), доцент кафедры информационных технологий предпринимательства Санкт-Петербургского государственного университета аэрокосмического приборостроения, ул. Большая Морская, д. 67, лит. А, Санкт-Петербург, 190000, Россия. Автор более 30 научных публикаций. Сфера научных интересов: лингвистика, искусственный интеллект, математика, теория проектирования баз данных, философия. E-mail: road.dust.spb@gmail.com