

Формально-лингвистическая модель семематической структуры информации*

П. Н. Афонин¹, Н. С. Яснова^{2✉}, П. А. Фалеев²

¹ Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В. И. Ульянова (Ленина), Санкт-Петербург, Россия

² Военно-космическая академия имени А. Ф. Можайского, Санкт-Петербург, Россия

✉ Kiya113@mail.ru

Аннотация. Представлена формализация семематических компонентов допредметной структуры информации, базирующаяся на работе «Философия имени» А. Ф. Лосева [1], продолжающая формализацию фонетических компонентов [2]. Целью статьи является уточнение концептуальных основ построения моделей представления знаний в системах обработки естественного языка в соответствии с многоуровневой структурой информации. В данной статье и в [2] представлены лишь 12 нижних уровней от фонетических компонентов до «символона II», который завершает рассмотрение и обеспечивает единство символических компонентов допредметной структуры информации. Представленная модель может быть использована в компьютерной лингвистике и разработке систем обработки естественного языка.

Ключевые слова: информация, искусственный интеллект, представление знаний, обработка естественного языка

Для цитирования: Афонин П. Н., Яснова Н. С., Фалеев П. А. Формально-лингвистическая модель семематической структуры вербальной информации // Изв. СПбГЭТУ «ЛЭТИ». 2023. Т. 16, № 4. С. 54–60. doi: 10.32603/2071-8985-2023-16-4-54-60.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Original article

A Formal Linguistics Model of Sememetic Structure of Information

P. N. Afonin¹, N. S. Jasnova^{2✉}, P. A. Faleev²

¹ Saint Petersburg Electrotechnical University, Saint Petersburg, Russia

² Mozhaisky Military Space Academy, Saint Petersburg, Russia

✉ Kiya113@mail.ru

Abstract. The paper is devoted to the formalization of sememetic components of presubject structure of information. It is based on the monograph «Philosophy of the name» by A. F. Losev [1] and presents the continuation of the phonetic components formalization [2]. The purpose of the article is to verify the conceptual basis of constructing the models of knowledge presentation in the systems of a natural language in accordance with the multi-level structure of information. Both in the present and in the previous [2] articles only 12 lower levels beginning from the phonetic components up to symbolone II are presented. Symbolone II completes the consideration and provides the unity of the symbolic components of the presubject information structure. The given model can be used in the computer linguistics and in the development of the processing systems of a natural language.

Keywords: information, artificial intelligence, knowledge representation, natural language processing

* Продолжение, начало в Изв. СПбГЭТУ «ЛЭТИ». 2022. Т. 15, № 8.

For citation: Afonin P. N., Jasnova N. S., Faleev P. A. A formal linguistics model of sememetic structure of information // LETI Transactions on Electrical Engineering & Computer Science. 2023. Vol. 16, no. 4. P. 54–60. doi: 10.32603/2071-8985-2023-16-4-54-60.

Conflict of interest. The authors declare no conflicts of interest.

Введение. Данная публикация продолжает статью [2], в которой было представлено формализованное описание компонентов фонетической структуры вербальной информации. Повторим кратко элементы фонетической структуры: 1) звук (воспринимаемое звучание), 2) голос (звук человеческого голоса), 3) членораздельность (членораздельный звук), 4) совокупность членораздельных звуков, 5) особенности произносящего. Нумерация компонентов соответствует нумерации, введенной А. Ф. Лосевым в [1]. В настоящей статье мы рассмотрим формализацию следующих 7 компонентов, соответствующих семематической структуре вербальной информации (текста).

Мы представили фонетическое слово как совокупность членораздельных звуков. Однако не всякая такая совокупность есть слово. Бесконечное число последовательностей звуков является лишь последовательностью значений самих звуков, и ничем большим. То же самое можно сказать и о последовательностях букв, не образующих слов, о последовательностях слов, не образующих никаких значений в целом.

Будем рассматривать лишь те последовательности информационных признаков (фонетических $Y_{\text{фн}}$, символических $Y_{\text{с}}$, мыслительных $Y_{\text{м}}$) имени, которые имеют какое-либо самостоятельное значение, т. е. значимость. Все они объединяются одним понятием – семема. В семеме значимым (т. е. имеющим какое-либо значение) элементам $x_j, x_k, \dots, x_l \in X$ множества информационных признаков $X \in \{Y_{\text{фн}}, Y_{\text{с}}, Y_{\text{м}}\}$ соответствует значимая последовательность этих элементов $\langle x_j, x_k, \dots, x_l \rangle \in R^n$, т. е. также имеющая свое значение:

$$m: X^n \rightarrow R^n.$$

Это значение еще не содержит атрибутивных признаков сущности объектов, но постулирует лишь значимость: есть такое слово, имеющее определенное звучание, написание, воспринимаемый образ, и это слово отличается от других своими элементами.

К уровню семемы следует отнести предмет следующего высказывания Фердинанда де Сос-

юра: «...язык есть система чистых значимостей, определяемая исключительно наличным состоянием входящих в нее элементов» [3, с. 82].

1. Фонематическая семема. Компонент б. Фонематическая семема ($X = Y_{\text{фн}}$).

Из первичных информационных фонетических признаков $y_i \in Y_{\text{фн}}$ строится признаковое описание слов речи конкретного человека, используемое в машинном обучении при распознавании речи:

$$\langle y_1, y_2 \rangle \in R_{\text{ф}}^2, \langle y_1, y_2, y_3 \rangle \in R_{\text{ф}}^3, \dots, \\ \langle y_1, y_2, \dots, y_n \rangle \in R_{\text{ф}}^n,$$

где $R_{\text{ф}}$ – множество отношений на множестве фонетических признаков, характеризующие фонемный состав (упорядоченные последовательности переходов от фонемы к фонеме), информацию об эмоциональном и физическом состоянии говорящего (отношения на более длинных последовательностях фонем), индивидуальных особенностях голоса и характеристиках среды (отношения на наибольших последовательностях фонем).

Если акустические признаки фонетического слова описывают его физические особенности, то совокупности информационных фонетических признаков слова:

$$\langle y_1, y_2, \dots, y_k \rangle \in R_{\text{ф}}^k, \quad k \leq n,$$

где n – максимальная длина последовательности фонем, $R_{\text{ф}}^k$ – множество отношений на конечных последовательностях фонем, составляют его фонематическое значение – фонематическую семему $s_{\text{ф}} \in S_{\text{ф}}$, которая дает возможность распознать из всего множества последовательностей фонем лишь те, которые составляют слова речи, позволяет сопоставлять слова по звучанию:

$$S_{\text{ф}} = \left\{ \langle y_1, y_2, \dots, y_k \rangle \mid y_i \in Y_{\text{фн}}, \right. \\ \left. \langle y_1, y_2, \dots, y_k \rangle \in R_{\text{ф}}^k, \quad i, k \in I \right\},$$

где I – множество индексов.

Произнесенное различными людьми одно слово имеет одну и ту же фонематическую семему.

2. Символическая семема ($X = Y_c$). Компонент 7. Этимон.

Оставим теперь за пределами рассмотрения фонетические компоненты и перейдем к символическому представлению информации – символической семеме слова, понимая под символом фонему при условии, когда звук носит незвуковое значение. «Каждый языковой элемент представляет собой *arriculus* – вычлененный сегмент, в котором понятие закрепляется определенными звуками, а звуки становятся знаком понятия» [3, с. 113].

От рассмотрения семемы как значения простой последовательности символов (по аналогии с фонематической семемой) ее отличает появление надсимвольных (производных от символов) значений: этимона, морфемы, синтагмы, пойемы. В них пока еще не отражается смысл слов, сущность информации, а закрепляются связи морфологического R_{mf} , морфемного R_M , лексического R_L и синтаксического R_C уровней языка:

$$\varphi_3 : X^k \rightarrow R^k, X = Y_c, R = \{R_{mf} \cup R_M \cup R_L \cup R_C\},$$

где φ_3 – отображение последовательности k информационных признаков $\langle x_1, x_2, \dots, x_k \rangle \in X^k$ (X^k – множество отношений, закрепляющих лишь порядок в последовательности k информационных признаков) в надсимвольную значимость последовательности $\langle y_1, y_2, \dots, y_k \rangle \in R^k$ (R^k – множество лексических и словообразовательных отношений, заданное на множестве информационных признаков) морфологических, морфемных, лексических и синтаксических признаков, выделяющее слово из толпы незначимых последовательностей символов – символического фона.

При этом различают лексическое значение слова (несущее его основное содержание), присущее слову в целом, и словообразовательное значение (несущее дополнительное значение), присущее части его словоформы (слова в определенной форме). Значение слова закрепляется в морфе.

Морф $s_{mp} \in S_{mp}$ – минимальная значимая часть словоформы. Лексическое значение включает корневой морф, а словообразовательное – аффиксальные морфы.

Наиболее простейшим компонентом, образующим символическую семему, является этимон $s_3 \in S_3 \subset S_{mp}$ (корневой морф, корень слова). Этимон – центральная минимальная значимая часть слова, заключающая в себе лексическое значение слова:

$$S_3 = \left\{ \begin{aligned} &\langle y_i, y_{i+1}, \dots, y_{i+k} \rangle | r_{л1} (\langle y_i, y_{i+1}, \dots, y_{i+k} \rangle) = \\ &= r_{л2} (\langle y_{i-l}, \dots, y_i, y_{i+1}, \dots, y_{i+k}, \dots, y_{i+m} \rangle), \\ &\quad r_{л1} : \langle y_i, y_{i+1}, \dots, y_{i+k} \rangle \rightarrow R_L, \\ &\quad r_{л2} : \langle y_{i-l}, \dots, y_i, y_{i+1}, \dots, y_{i+k}, \dots, y_{i+m} \rangle \rightarrow R_L, \\ &\quad y_{i-l}, \dots, y_{i+m} \in Y_c, \\ &\quad k = \arg \min_j \left\{ \left| \langle y_i, y_{i+1}, \dots, y_{i+j} \rangle \right|, i, l, m \in N, i > l \right\}, \end{aligned} \right.$$

где $r_{л1}, r_{л2}$ – лексические функции, определяющие для соответствующих наборов фонем их лексические значения.

Другими словами, этимон – это минимальная последовательность фонем, с которыми определен смысл, однако не сам смысл, а лишь постулирование его существования, т. е. лексическая значимость, отличающая его от других этимонов. Постулирование этого смысла принадлежит любой словоформе, образованной от этимона.

Меньше этимона существует лишь последовательность значений символов звуков сама по себе, не образующая общего для них значения, т. е. фонематическая семема. Мы не используем понятие лексем, так как на данном уровне не предполагаются ее различные лексико-семантические варианты, относящиеся уже к сфере значения слова, но постулируется лишь лексическая значимость.

Компонент 8. Морфема. В различных словоформах последовательность символов – этимон – может меняться, например при склонении или спряжении. Возникают другие корневые морфы, которые по смыслу представляют тот же морф, отличающийся от других лишь одной фонемой (отсутствием фонемы). Совокупности таких корневых морфов, тождественных по лексическому значению и близких по составу фонем, образуют корневые морфемы $s_{к.м} \in S_{к.м} \subset S_M$ – подмножество множества морфем, фактор-множество множества корневых морфов по отношению эквивалентности морфем ρ_M :

$$S_{к.м} = S_3 / \rho_M.$$

В зависимости от соседних морфов морфема $s_{mi} = \langle y_1, y_2, \dots, y_k \rangle, s_{mi} \in S_M, i \in I$ может приобретать различные оттенки написания, что порождает множество алломорфов S_{ai} i -й морфемы.

Таким образом, множество алломорфов S_a представляет собой реализацию морфемы в языке (например, при проявлении в различных словоформах) – отображение φ_4 :

$$\varphi_4 : S_M \rightarrow S_a.$$

Также множество морфов $S_{M,p}$ представляет собой реализацию морфемы в речи (отображение φ_5):

$$\varphi_5 : S_M \rightarrow S_{M,p}.$$

Между фонетическим уровнем языка и морфологическим уровнем существует промежуточный уровень языка – морфонологический, образующийся чередованием фонем в составе морфем или морфов. Единицей такого уровня служит морфонема $s_{Mф} \in S_{Mф} \subset S_a$, также представляющая собой алломорф морфемы.

Из этимона (корневого морфа) $s_{zi} \in S_3$, $s_{zi} = s_{M,pj}$, $i, j \in I$, аффиксальных морфов $s_{ai} \in S_{M,p}$, $s_{ai} = s_{M,pk}$, $i, k \in I$ строятся последовательности морфем – словоформы слова $s_{сф} = \langle s_{M,p1}, \dots, s_{M,pk} \rangle \in S_{сф}$, $i, k \in I$:

$$S_{сф} = \left\{ \left\langle s_{M,p1}, \dots, s_{M,pk} \right\rangle \middle| s_{Mpi} \in S_{M,p}, \right. \\ \left. \left\langle s_{M,p1}, \dots, s_{M,pk} \right\rangle \in R_{л}, \right. \\ \left. k = \left| \left\{ s_{M,p1}, \dots, s_{M,pk} \right\} \right| \right\}.$$

Таким образом, каждая корневая морфема и словоформа содержат в себе одно и то же лексическое значение, соответствующее своему этимону.

Компонент 9. Синтагма. Помимо связанности морфем, определяющей словоформы как минимальные значимые совокупности морфов, в символической семеме следует выделить связи, формирующие контекст смысловых отрезков речи. Минимальные (интонационно нечленимые) смысловые отрезки речи – синтагмы $s_c \in S_c$ – включают простое или сложное предложение, главную или придаточную часть сложного предложения, словосочетание, отдельную словоформу, служебное слово. В данном случае – в контексте других слов речевого отрезка – каждое слово несет на себе значение всего речевого отрезка, выражающееся не только в его лексическом составе, но и в его синтаксическом строении и даже звуковом составе (например, при выражении интонации).

Данные связи отражают не самый смысл речевых отрезков, а лишь постулирование его существования, смысловую значимость речевых отрезков, т. е. смысловозначительную возможность, проявляющуюся главным образом при сопоставлении соотношений синтаксического строения R_c речевого отрезка и его лексического со-

става $R_{л}$ или синтаксического строения R_c и звукового состава $R_{ф}^l$ словоформ. Другими словами, цепочки $\langle s_{сф1}, s_{сф2}, \dots, s_{сфm} \rangle$ словоформ $s_{сфj} \in S_{сф}$ синтагмы s_{ci} должны быть синтаксически связаны и иметь собственное лексическое значение $r_{ли}$, общее для всех словоформ цепочки, либо синтаксически связаны и иметь собственный фонематический состав $R_{ф}^l$:

$$S_c = \left\{ \begin{array}{l} s_{ci} | s_{ci} = \langle s_{сф1}, s_{сф2}, \dots, s_{сфm} \rangle, \\ s_{ci} \in R_c^m \times R_{л} \cup R_c^m \times R_{ф}^l, \\ m = \arg \min_j \left| \left\{ s_{сф1}, s_{сф2}, \dots, s_{сфj}, \dots \right\} \right| \end{array} \right\}$$

Наборы одних и тех же словоформ, имеющих различные синтаксические связи, образуют разные синтагмы так же, как и наборы словоформ, отличающиеся лишь интонационно (например, вопрос и его повторение не эквивалентны по смыслу).

Смысловоразличительная функция φ_6 синтагмы s_c как интонационно нечленимого смыслового отрезка речи реализована соотношением тона, тембра, интенсивности и длительности звучащей речи:

$$\varphi_6 : S_c \rightarrow K_I, \quad K_I = F^4 \times A^4 \times T,$$

где $K_I = \{k_{Ij}\}$ – интонационные конструкции на речевых отрезках, $j \in \overline{1, |K_I|}$; $F^4 = \{f_i\}$ – частоты четырех первых формант i -х фонем речевых отрезков, определяющие их тон и тембр; $A^4 = \{a_i\}$ – амплитуды четырех первых формант i -х фонем речевых отрезков, определяющие их интенсивность; $T = \{\tau_i\}$ – длительности i -х фонем речевых отрезков.

Синтагма в современной русской грамматике как совокупность словоформ не является синтагмой слова, введенной А. Ф. Лосевым.

Под синтагмой слова будем понимать контекстно-зависимую (в границах синтагмы как речевого отрезка) значимость слова, выражающуюся лексико-грамматическими категориями слова и его местом в интонационной конструкции определенного типа, определяемых совокупностью лексических и синтаксических или синтаксических и фонетических связей всего речевого отрезка:

$$S_{c,c} = \{s_{c.ci} | \langle s_{c.c1}, \dots, s_{c.cm} \rangle \in S_c, i = \overline{1, m}\}.$$

Компонент 10. Пойема. Расширяя границы рассмотрения допредметной структуры слова от минимальной значимой части слова до значимости слова как элемента речевого отрезка, переходим к рассмотрению значимости слова в контексте всего предложения на примере русского языка.

Пойемой слова s_n будем называть контекстно-зависимую (в рамках предложения, иногда нескольких предложений), смысловозначительную значимость слова, выражающуюся в сочетании его места в предложении с компонентами определенной интонации в слове и субъектно-модальными значениями всего предложения, представленных помимо компонентов интонации и порядка слов элементами специальных синтаксических конструкций, повторением слов, сочетаниями с модальными частицами, междометиями, вводными словами и словосочетаниями, комбинациями этих средств:

$$S_n = \left\{ s_{ni} \mid \forall s_{cfi} : s_{ci} = \langle s_{cf1}, s_{cf2}, \dots, s_{cfm} \rangle \exists s_{ni} \in \in (P_k \cup S_{c.k} \cup S_{cn} \cup S_{cmn}) \times R_c^m \times R_I, i = \overline{1, S_n} \right\},$$

где P_k – коммуникативная парадигма предложения; $S_{c.k}$ – специальные синтаксические конструкции и повторения слов; S_{cn} – субъективно-модальные значения, выражаемые средствами словоупотребления; S_{cmn} – субъективно-модальные значения, выражаемые средствами субъективно-модальных построений; P_I – множество нейтральных и модальных реализаций интонационных конструкций.

Смысловозначительная значимость слова характеризуется, в первую очередь, коммуникативной функцией предложения (выделением наиболее важного аспекта содержания), выражаемой соотношением места слова в предложении (порядком слов) и интонацией, разнообразными оценками и экспрессивными характеристиками, а также другими внешними приемами, используемыми для выразительности (например, стилистическими особенностями языка и особенностями жанра речи).

Интонация выражает субъективно-модальное значение направлением движения и уровнем основного тона, усилением словесного ударения, увеличением длительности ударного гласного – интонационными средствами, образующими множество P_I нейтральных и модальных реализаций r_{Iv} интонационных конструкций типа k_I :

$$R_I = \left\{ r_{Iv} \mid r_{Iv} \in K_I \times C_I \times S_I \times P_I \times Y_a \right\}, v \in \overline{1, |R_I|},$$

где K_I – множество типов k_I интонационных конструкций, $i \in \overline{1, |K_I|}$; C_I – множество возможных мест центра интонационной конструкции; S_I – множество вариантов синтагматического членения; P_I – множество вариантов реализации пауз; Y_a – множество аллофонов языка.

Сочетания порядка слов предложения с определенной интонацией образуют абстрактные шаблоны – коммуникативные варианты (варианты одного предложения, имеющие разное актуальное членение – разделение на тему и рему), образующие коммуникативную парадигму:

$$P_k = \left\{ p_{ij} \mid p_{ij} = \left\langle \left\langle \langle \dots, s_{cfi}, \dots \rangle \right\rangle, r_{Iv} \right\rangle \in R_c^m \times R_I \right\},$$

где p_{ij} – коммуникативные варианты предложения, $\forall i, l \in \overline{1, m-1}, j, t \in \overline{1, m}, i \neq l, j \neq t, p_{ij} \neq p_{lt}$; $\langle \dots, s_{cfi}, \dots \rangle$ – тема сообщения варианта p_{ij} , включающая словоформу $s_{cfi} \in S_{cf} \cup \emptyset$; $\langle \dots, s_{cfj}, \dots \rangle$ – рема сообщения варианта p_{ij} , включающая словоформу $s_{cfj} \in S_{cf}$.

Специальные синтаксические конструкции и повторения слов $S_{c.k}$ представляют собой совокупность синтаксических фразеологизмов F_c и соединений словоформ U_{cf} :

$$S_{c.k} = F_c \cup U_{cf},$$

где $F_c = \langle s_{cfi}, \dots, s_{cfj} \rangle$ – синтаксические фразеологизмы; U_{cf} – соединения словоформ s_{cfi} и s_{cfj} , построенные по определенным образцам, основанные на повторении этимона $s_{\Delta i}$ в словоформах s_{cfi} и s_{cfj} в составе одной синтагмы s_c (либо двух синтагм с растяжением гласных в центре интонационной конструкции второго типа r_{I2}):

$$U_{cf} = \left\{ \begin{array}{l} u_{cfi} \mid u_{cfi} = \langle s_{cfi}, \dots, s_{cfj} \rangle, \\ s_{cfi} = \langle \dots, s_{\Delta i}, \dots \rangle, s_{\Delta i} \in S_{\Delta}, \\ s_{cfj} = \langle \dots, s_{\Delta i}, \dots \rangle, \langle s_{cfi}, \dots, s_{cfj} \rangle \in S_c \vee \\ \vee \langle \langle s_{cfi}, \dots, s_{cfj} \rangle, r_{I2} \rangle \in S_c^2 \times R_I \end{array} \right\}.$$

Средствами словоупотребления выражается не только значение актуализации, но и другие субъективно-модальные значения s_{cni} . Для этого предикативно значимый член предложения s_{cfj} вы-

носится в начальную позицию и выделяется центром интонационной конструкции второго или третьего типа:

$$S_{cn} = \left\{ s_{cni} \mid s_{cni} = \langle \langle s_{c\phi j}, \dots, s_{c\phi l}, \dots \rangle \rangle, \right. \\ \left. k_{I2} \vee k_{I3} \in R_c^m \times K_I \right\}.$$

Функцию выражения субъективно-модальных значений выполняют также построения с модальными частицами, глагольными формами, наречиями в роли модальных частиц, согласуемыми словами, междометиями, вводными словами или вводными сочетаниями слов, вводными предложениями, образующими множество субъективно-модальных построений:

$$S_{cmn} = \left\{ s_{cmni} \mid s_{cmni} = \langle \dots, s_{c\phi j}, \dots \rangle \in R_c^m, \right. \\ \left. i = 1, \overline{S_{cmn}}, j = \overline{1, m} \right\}.$$

Компонент 11. Символон I (символическая система $(X = Y_c)$). Обобщим рассмотренные компоненты символической семемы (этимон, морфему, синтагму, пойма), в которых значимым последовательностям $\langle x_1, x_2, \dots, x_n \rangle \in X$ информационных фонетических признаков $\langle x_1, x_2, \dots, x_n \rangle$ соответствует значимость этой последовательности как символ, выраженный в корне слова $s_3 \in S_3$, его словоформах $s_{mj} \in S_M$, лексико-грамматических категориях слова и его месте в интонационной конструкции определенного типа $s_{c.c.k} \in S_{c.c}$, а также в сочетании его места в предложении с компонентами определенной интонации в слове и субъектно-модальными значениями $s_{nl} \in S_n$. Такая значимость последовательности информационных фонетических признаков, взятая в единстве всех ее компонентов как символ, обозначается у А. Ф. Лосева первым символоном *SI* или первым символическим единством слова (символическим единством семемы, символическим слоем семемы, символической семемой):

$$SI = \langle s_3, s_{mj}, s_{c.c.k}, s_{nl} \rangle,$$

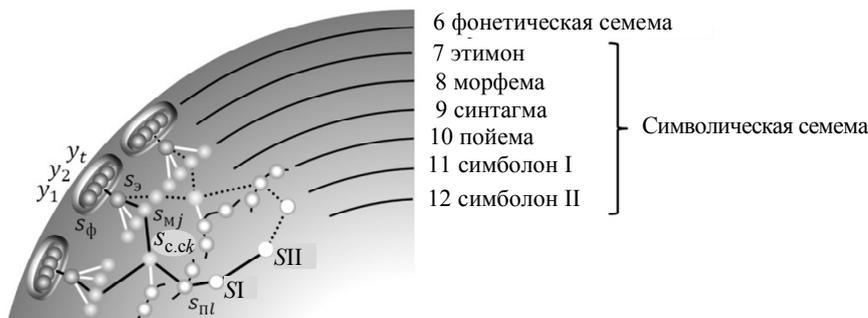
где s_3 – этимон слова; s_{mj} – j -я морфема слова; $s_{c.c.k}$ – k -я синтагма слова; s_{nl} – l -я пойма слова.

Компонент 12. Символон II (полная символическая семема). Завершением рассмотрения единства символических компонентов слова как значения слова относительно его звуковой стороны, в которых звук несет незвуковое значение, является объединение всех возможных вариантов выражения этимона слова s_3 – изменения форм слова (словоформ слова S_M) и контекстов его реализации, образованных синтагматическим членением (синтагм слова $S_{c.c}$), актуальным членением (мест в предложении с компонентами определенной интонации) и субъектно-модальным значением предложения (пойма слова S_n) в виде второго символона *SII* или полной символической семемы (см. рисунок).

$$SII = \langle s_3, S_M, S_{c.c}, S_n \rangle,$$

где s_3 – этимон слова; S_M – множество всех морфем слова; $S_{c.c}$ – множество всех синтагм слова; S_n – множество всех пойм слова.

Заключение. В данной статье мы продолжили рассматривать начатую в [2] формализацию допредметной структуры информации в части, касающейся ее семи символических компонентов. Это, на наш взгляд, позволит составить более качественное описание классов в задаче множественной классификации словоформ при обработке естественного языка, в частности в машинном переводе. Аналогичная задача в отдельных областях, например морфологического или лексико-грамматического анализа, решалась и ранее [4] эвристическими, нейросетевыми и другими методами. Однако описание допредметной структуры – фонетических и символических компонентов – лишь видимая часть айсберга с названием «информация» («вербальная информация»). Невидимая часть касается сферы мышления, смысла.



Семематическая структура вербальной информации по А. Ф. Лосеву
Sememetic structure of verbal information according to A. F. Losev

За границами рассмотрения семематической структуры остались поэтические компоненты допредметной структуры, которые описывают сферу значения слова (но не смысла слова), и идея слова, занимающая пограничное положение в сфере значения, максимально приближенное к

сфере смысла [5], что мы надеемся обсудить в третьей статье, посвященной допредметной структуре информации.

Таким образом, циклом из трех статей мы предполагаем formalizovat допредметную структуру вербальной информации.

Список литературы

1. Лосев А. Ф. *Философия имени*. М.: МГУ, 1989. 81 с.
2. Афонин П. Н., Яснова Н. С., Фалеев П. А. Формально-лингвистическая модель допредметной структуры вербальной информации // *Изв. СПбГЭТУ «ЛЭТИ»*. 2022. Т. 15, № 8. С. 43–49. doi: 10.32603/2071-8985-2022-15-8-43-49.

3. Соссюр Ф. *Курс общей лингвистики* / под ред. Ш. Балли, А. Сеше; пер. с фр. А. Сухотина. Екатеринбург: Изд-во Урал. ун-та, 1999. 432 с.
4. Национальный корпус русского языка. URL: <https://ruscorpora.ru> (дата обращения 08.01.2023).
5. Фалеев П. А. Представление знаний с позиций диалектики числа А. Ф. Лосева // *Вестн. Воронежского гос. ун-та. Сер.: Философия*. 2017. № 3 (25). С. 212–217.

Информация об авторах

Афонин Петр Николаевич – д-р техн. наук, проректор по стратегическому развитию, зав. кафедрой прикладной механики и инженерной графики СПбГЭТУ «ЛЭТИ».

E-mail: pnafonin@etu.ru

Яснова Нина Семеновна – канд. пед. наук, доцент, зав. кафедрой иностранных языков ВКА имени А. Ф. Можайского. Почетный работник высшего профессионального образования РФ.

E-mail: Kiya113@mail.ru

Фалеев Петр Анатольевич – канд. воен. наук, старший преподаватель кафедры информационно-аналитической работы ВКА имени А. Ф. Можайского.

E-mail: fpah@mail.ru

References

1. Losev A. F. *Filosofija imeni*. M.: MGU, 1989. 81 s.
2. Afonin P. N., Jasnova N. S., Faleev P. A. Formal'no-lingvisticheskaja model' dopredmetnoj struktury verbal'noj informacii // *Izv. SPbGjeTU «LETI»*. 2022. T. 15, № 8. S. 43–49. doi: 10.32603/2071-8985-2022-15-8-43-49.
3. Sossjur F. *Kurs obshhej lingvistiki* / pod red. Sh. Balli, A. Seshe; per. s fr. A. Suhotina. Ekaterinburg: Izd-vo Ural. un-ta, 1999. 432 s.

4. Nacional'nyj korpus russkogo jazyka. URL: <https://ruscorpora.ru> (data obrashhenija 08.01.2023).
5. Faleev P. A. Predstavlenie znaniy s pozicij dialektiki chisla A. F. Loseva // *Vestn. Voronezhskogo gos. un-ta. Ser.: Filosofija*. 2017. № 3 (25). S. 212–217. (In Russ.).

Information about the authors

Petr N. Afonin – Dr. Sci. (Eng.), Vice-Rector for Strategic Development, Head of the Department of Applied Mechanics and Engineering Graphics, Saint Petersburg Electrotechnical University.

E-mail: pnafonin@etu.ru

Nina S. Jasnova – Cand. Sci. (Pedagogy), assistant professor, Head of the Foreign Languages Department of Mozhaisky Military Space Academy. Honorary scholar of higher professional education of Russian Federation.

E-mail: Kiya113@mail.ru

Petr A. Faleev – Cand. Sci. (Military), senior teacher of the Information-Analitical Department of Mozhaisky Military Space Academy.

E-mail: fpah@mail.ru

Статья поступила в редакцию 17.12.2022; принята к публикации после рецензирования 13.01.2023; опубликована онлайн 25.04.2023.

Submitted 17.12.2022; accepted 13.01.2023; published online 25.04.2023.
