Yenok Grishkyan
Lecturer at Yerevan State Medical University,
Ph.D. in Philology, Assistant
Email: yenokgrishkyan@gmail.com

DOI: 10.53614/18294952-2024.2-189

UNIVERSAL LANGUAGE (METALANGUAGE WE CAN'T LIVE WITHOUT)

Nowadays, the modeling of the knowledge has become one of the problems in the fields of application of various sciences: logics, linguistics, mathematics, psychology, cybernetics, etc. The reason for such interest in this problem is due to the creation of artificial intelligence (AI) systems.

The present article explores metalanguage, a type of universal language present in machine translation that facilitates the encoding of the input language and the decoding of the output language, regardless of the language family or typological traits. The modeling or formalized representation of words plays a vital role in ensuring accurate translation and understanding of meaning.

The article contains sample translations from English to Chinese, Spanish, Latin, Armenian, Russian, and Turkish via corresponding formal description based on the metalanguage used by the Universal Network Language (UNL). One of the vital parts of the formalized description – the component analysis, is also discussed from the point of view of both morphology and semantic meaning.

Key words: Universal language, metalanguage, component analysis, machine translation, formalized description.

Introduction

Knowledge is given to us directly by language and only through it: knowledge exists in daily use (monologue, dialogue, phrases, etc.), as well as in written texts

(literary monuments, artistic, scientific and technical literature, press, various linguistic styles, etc.), and finally, through language, we transmit knowledge from generation to generation. Thus, language is both a means of acquiring and preserving knowledge, and a means of its transmission. In science and technology, knowledge outside of the texts generally does not exist, and the modeling of the meaning of scientific and technical texts implies the modeling of the knowledge system of the given field.

Natural language knowledge, in general, can be divided into three main levels: linguistic – semantic information of words for understanding sentences in a natural language; concept definition – a sort of connection with other concepts that in turn specify the concepts; and other – including any other knowledge, e.g. encyclopedic. When aligned with metalanguage, this knowledge modelling plays a crucial part in forming the corrected meaning of the word specifying each meaning to singularity, which enhances the correct translation of the word within the sentence.

Simple, yet hard to translate

“Добрый день!” (Russian)
“नमस्ते.” (Hindi)
“Is there a museum in the city?” (English)
“当然!” (Chinese)
“¿Cómo puedo encontrarlo?” (Spanish)
“Գնացեք ուղիղ, ապա թեքվեք դեպի աջ:” (Armenian)
“ありがとう” (Japanese)

As simple as the conversation above seems, it is quite difficult to formulate them in other languages, especially when a person is a carrier of only one language and has no understanding of the others. It might be easier for him to comprehend and converse in a foreign language, other than his own. But what if he doesn’t?

Not long ago people used to help foreigners by translating for them either orally or in written form: a tradition that has come down to us since Ancient Egypt [Salevsky, 2018, pp. 184–198]. Many countries of that period used to establish translation service agencies to be able to communicate with the slaves or the people of the neighboring territories they conquered. Some were talented and achieved success. However, with the discovery of the informational era, these services are being used less and less due to the use of the fast growing internet.

Thanks to the worldwide trade, the web gave rise to computational linguistics which, in turn, facilitated the communication of people all over the world with the aid of machine translation. Attempts to make communication between languages easier by means of different translation devices haven’t always been successful. The route wasn’t easy, but the continuous changes have made the system more or less perfect because of an intermediate language that is used in translating from one language into another. This tertiary language creates a sphere

filled with linguistic models that should, in a common sense, be able to translate from one language into another and vice versa. This so called sphere is filled with phonetical, lexical, grammatical (morphological and syntactic) and even stylistic categories of all available languages, no matter alive or dead.

A metalanguage is an artificial language that helps in doing automated translations by means of online translating services. In difference to artificially constructed languages, a metalanguage is general for artificial and natural languages and is used as an auxiliary way to transfer one language data to another.

Alternatively said, the metalanguage considers the language outside its usually ambient. The reality of one language is encoded through it inside itself by means of grammatical expressions it is using (words written in italics, punctuation marks, dashes, slashes, etc. with each expression being on a separate line) (Karaulov, Molchanov, Afanasyev & Mikhalev, 1982).

Three types of metalanguages are – embedded, ordered and nested. Each of them has its own function: embedded metalanguage is fixed to a certain language. An ordered metalanguage creates its own metalanguage to discuss the one prior to it on the same level, whilst a nested one discuss the prior on a higher level (hence hierarchy). The latter two coexist in a natural language, however the former is mostly used by the translating system to move one language forms and expressions into the other. One of the first models of a metalanguage is Backus–Naur Form which was developed by John Backus and Peter Naur in the 1960s (Backus, 1959, pp. 125–132).

Hereafter we shall be discussing phonetic, lexical, grammatical and syntactical aspects of a metalanguage, as well as give examples of the formalized description of the metalanguage in different languages.

Since 1998 the creation of speech recognition system in Microsoft office has been a vital part of having a good translation output. Even though it is done prior to using the metalanguage, correct speech recognition helps to improve the correct choice of modeling for encoding. The speech recognition is still in development for many languages that have been less studied.

Lexical modeling of a language depends on the type of language whether it is synthetic or analytic. Due to these two main differences between world languages, the words should be carefully sorted into models that give the understanding of the word incomplete details: logical and analytical details present not only in the input language, but also in the output one.

For example: the word «հավք» in Armenian is mostly used in stylistic sense denoting birds in general, however when used without the suffix ք it will denote only a hen. A detailed modeling showing the type of the bird, being a neutral or generalized name, ability of laying eggs or building nests, etc. should be used for translating from Armenian (syntactical language) into English (analytical language).

Simple as yet a glance deep into the metalanguage shows us its backbone – an analytical structure of words in such details, that show not only the form and meaning of the word, but also all possible grammatical forms (affixation, endings,

prepositions, postpositions, etc.) that the word may have in all languages having already been input in the software. Thus, the English word "red" can have 1 form each in English and Armenian, 4 forms in most Romance languages (2 genders multiplied by 2 numbers (e.g. in Spanish – *rojo/rojos, roja/rojas*), 24 in German (3 genders, 2 numbers in all 4 cases), 36 in Russian (3 genders, 2 numbers in all 6 cases), etc. This means, that for English, Russian, Spanish, Italian, German and Armenian only the word "red" in a metalanguage will have 70 morphological endings. In difference to syntactical languages like Russian, German, Latin, etc., modern Armenian lacks the agreement between the noun and the adjective and, as the latter has no declension system, only one form is extant.

All these forms are part of a linguistic method, so called component analysis, a unique attempt to describe word meanings universally not only semantically, but also formally, creating a model with all possible combinations for both input and output languages.

Formalizing the meaning

Turning to the internal structure of the word, the meaning, the logical characterization method of the semantic form of the word describes component analysis in a better way. Here a word class is divided into types, which in turn are divided into subtypes required for lexical explanation. Therefore, it is essential that lexicological explanations provide enough suitable material for the examination of word groups as a limited system of semantic components (Hyvärinen, Karhunen, Oja, 2001, pp. 147–161).

Some lexical segments can be characterized by formalizing and reducing these explanations to a certain number guided by certain rules. Component analysis can also be achieved through transformative analysis. It is well known that the transformational identity or difference is directly proportional to the identity or difference of the constituent structure of the word unit. It allows monosemantic and words with binary contrast express short meaning more efficiently. However, binary relations are not always the best option for semantic analysis, e.g., it is possible to distinguish some species of animals, e.g. some types of cattle, by a number of features, taking into account their binary connections, but one of the important conditions – the age, will be lost.

As such, the component analysis was developed and is used mainly the meaning of the words, thus its main task is to formulize the semantic field of the word in its relation with the words that surround it in particular and the text in general. It can also be used in other areas of linguistics: e.g. phonemics, where the binary relations are considered essential. E. Gulyga and E. Shendels concluded that component analysis plays an essential role not only in extracting the exact meaning of the word, but assists in formalizing morphology and syntax as well (Gulyga & Shendels 1969).

If in component analysis the meaning of a word depends on its components, in linguistics, it is called semantic components, differential semantic signs, semantic

dimensions, etc. The semantic component is also called a *sema* – a meaning that is presented more concisely and clearly and proves the component's affiliation to the meaning – the content field and has lexical flexibility, i.e. *semantic field, hypersema, classema,* etc.

The term "*sema"* or meaning can be used to mean the smallest component meaning of a word – an element that reflects the signs of the signified meaning. In other words, its meaning entirety constitutes the main sense of the lexical unit. However, it can be further segmented: e.g. the word denoting gender can be separated as male and female, which in turn can be divided into ages or other kinds of certain conditions – a lady is of feminine gender, usually of a young age and commonly not married, while a man is of masculine gender, where any age past adulthood is possible and no certain conditions are required.

Thus, what's common to all languages formalized here, is the pronoun description, the other parts of speech have characteristics somewhat different from each other. Note, that due to the syntax structure, some of these characteristics are put in the middle or in the end of the formalized metalanguage. When translated in an output language, this placement is done by the general syntax rules already written in that language. During the decodification process the syntax structure is written taking these rules into account.

To give a more accurate formal description of the word, three conditions have to be observed:

- Describe the words in the most detailed form possible, especially in case of homonyms;
- Write the descriptions as short as possible to provide a faster and easier work for the translating system.
- Determine the lexical–semantic description of the word mostly according to the text, rather than the word alone.

Quite often the word should be viewed within the context for a better outcome of the translation. This is due to the fact, that the formalization of the word meaning in machine translation implies not only an intra–lexical correlation, but also an inter–lexical correlation, as the intra–lexical correlation is the internal form of the word, while the inter–lexical one is the semantic dependence of the word on other lexical meanings present in the vocabulary and attached to the given word. The transformation of the latter, however, is more complicated because it passes from lexicological level to syntax.

Imagine a polyglot understating all the 7 languages in the conversation in the beginning of the article trying to translate them to the carriers of these languages simultaneously. The brain neurons in the Broca's area in the left hemisphere move 0.5–2.0 m/s to and fro the Wernicke's area to comprehend (encode), translate and interpret (decode) the sentences. This passage between them assists in translation, becoming kind of metalanguage. In difference to the artificial translation that uses formalized description, natural one just creates the images pictured during encodification and interprets them during decodification using the words and

grammatical rules memorized previously.

A metalanguage functions much faster and much more correct due to technical compatibilities and rules previously added to it. This includes not only the meaning, but also the formalized description of each word separately and, at the same time, as a whole in the ambient of the context. What connects them is the component analysis thanks to which the translation is carried out accurately.

In short, a metalanguage is crucial for translating and it becomes a form of universal language – a language that is one for the whole humanity, like Esperanto, Volpuk (constructive languages), or English, Arabic, Russian (natural languages) that serve as Lingua Franca for different nationalities – carriers of different languages that might not have a lot in common with each other.

Universal language theory

A theory of one language for everyone, universal language theory (ULT) (Cook, 1996), has been proposed by Armenian academician, prof. G. Jahukyan, who has described it as a universal language model (ULM) – a generalized system for describing real objects and their correlations, and that reality includes physical as well as spiritual and emotional realities.

The ULM is built with a harmonious combination of introductory and descriptive methods and is based on five main principles: materiality, generality, integrity, unity and economy. Based on the principle of generality, the ULM can be applied to both natural and artificial languages. When building it, the concepts of the given language in general are taken into account as the perfect reality of the language. The purpose of ULM is to understand and reproduce language as a means of comprehension and communication between carriers of various languages.

According to the principle of universality, the ULM is applicable to all aspects of language, describing not only the expression and content of the word, but also the units formed by combinations in different languages (Roberts, 2016).

The main feature and the advantage of ULM is suggesting a minimum number (only 8) of universal concept categories that characterize objects materially (composition, hanging, number and class) and spatially (shape, direction, size and location). Apart from these 8, other paradigmatic categories are also described characterizing not only the objects themselves, but also their properties and relations (comparison, causality, frequency, manner, process, change, duration and time). Finally, 6 additional categories are added to characterize the objects of the spiritual–emotional sphere: perception, awareness, material expression of a language, modality, thinking, communication. Each of these 22 orders is characterized by hexadecimal edges: 1. main, 2. opposite, 3. middle (neither/nor), 4. joint (and/or), 5. reflexive, 6. absolute or neutral (Jahukyan, 2003).

Any machine translation program must index, label and number the words imported into it, present them according to the grammatical rules included in its system. All words corresponding to 22 categories in ULM should be indexed (See table 1 below).

Table 1.

	1	2	3	4	5	6
1	1-1	1-2	1-3	1-4	1-5	1-6
2	2-1	2-2	2-3	2-4	2-5	2-6
3	3-1	3-2	3-3	3-4	3-5	3-6
4	4-1	4-2	4-3	4-4	4-5	4-6
5	5-1	5-2	5-3	5-4	5-5	5-6
6	6-1	6-2	6-3	6-4	6-5	6-6
7	7-1	7-2	7-3	7-4	7-5	7-6
8	8-1	8-2	8-3	8-4	8-5	8-6
9	9-1	9-2	9-3	9-4	9-5	9-6
10	10-1	10-2	10-3	10-4	10-5	10-6
11	11-1	11-2	11-3	11-4	11-5	11-6
12	12-1	12-2	12-3	12-4	12-5	12-6
13	13-1	13-2	13-3	13-4	13-5	13-6
14	14-1	14-2	14-3	14-4	14-5	14-6
15	15-1	15-2	15-3	15-4	15-5	15-6
16	16-1	16-2	16-3	16-4	16-5	16-6
17	17-1	17-2	17-3	17-4	17-5	17-6
18	18-1	18-2	18-3	18-4	18-5	18-6
19	19-1	19-2	19-3	19-4	19-5	19-6
20	20-1	20-2	20-3	20-4	20-5	20-6
21	21-1	21-2	21-3	21-4	21-5	21-6
22	22-1	22-2	22-3	22-4	22-5	22-6

In addition, the ULM can also be used in other translation systems to form description and choose the foreign language words correctly. Whatever the choice, ULM is the most complete system of meanings presented to date capable of acting as a metalanguage in translation devices and performing full semantic inter–lingual translations.

Working metalanguage

It's not only the meaning that the component analysis differs. A more common approach is using different components of the metalanguage to specify the grammatical category of the word within the sentence. Main problems arise when there is a need to translate from a syntactic language into an analytic one. When the translation is carried out in this manner, formal modeling of the output language becomes much more complex and therefore long. This is due to the fact, that analytical languages use articles, prepositions, different particles, etc. that

belong to different parts of speech which, in turn, are added to the word when creating a formal description. Translation from an analytical language into a syntactical one is much simpler, because formalized description of a syntactical language commonly uses only one word already including various grammatical forms expressing the correct meaning of the analytical language (less parts of speech are in use).

Let's observe a sample on a Universal Networking Language (UNL) based translation from English into Chinese (A), Spanish (B), Latin (C), Armenian (D), Russian (E) and Turkish (F).

Input language formalized description for the metalanguage:
Initial input: I am reading a book.
I(PPronPrs1Sg)–personal pronoun, 1st person, singular;
am(AuxVPresContPrs1Sg)–auxiliary verb, present continuous tense, 1st person, singular;
reading(VGer)–verb, gerund;
a(IndfSg)–indefinite article, singular;
book(CNSg)–common noun, singular.
Metalanguage input description for: I am reading a book.
SNT*(PPronPrs1SgAuxVPresContPrs1SgVGerIndfSgCNSg)

A. Output language (Chinese) formalized description for the metalanguage:
我 **(wo)** (PPronPrs1Sg)–personal pronoun, 1st person, singular;
正在 **(zheng zai)** (ConjT)–conjunction of time;
读 **(du)** (V)–verb;
一 **(yi)** (CNum1)–cardinal number 1;
本 **(yi ben)** (ClBk)–classifier for books;
书 **(shu)** (CNSg)–common noun, singular;
Metalanguage output description for: 我正在读(一本)**书 (wo zheng zai du (yi ben) shu).
SNT(PPronPrs1SgConjTVCNum1ClBkCNSg)

B. Output language (Spanish) formalized description for the metalanguage:
Yo(PPronPrs1Sg)–personal pronoun, 1st person, singular;
estoy(AuxVPrContPrs1Sg)–auxiliary verb, present continuous tense, 1st person, singular;
leyendo(VGer)–verb, gerund;
un(IndfMSg)–indefinite article, masculine, singular;
libro(CNMSg)–common noun, masculine, singular.
Metalanguage output description for: (Yo)*** estoy leyendo un libro.
SNT(PPronPrs1SgAuxVPresContPrs1SgVGerIndfMSgCNMSg)

C. Output language (Latin) formalized description for the metalanguage:
Ego(PPronPrs1Sg)–personal pronoun, 1st person, singular;
librum(ND2nAccSg)–noun, declension 2, neutral gender, accusative, singular;
lego(VC3PresPrs1Sg)–verb, conjucation 3, present tense, 1st person, singular.
Metalanguage output description for: (Ego)** librum lego.
SNT(PPronPrs1SgND2nAccSgVC3PresPrs1Sg)

D. Output language (Armenian) formalized description for the metalanguage:
Ես (yes) (PPronPrs1Sg)–personal pronoun, 1st person, singular;
գիրք (girk') (NEInfիAccSg)–noun, external Inflection ի, accusative, singular;
եմ (em) (AuxVPresContPrs1Sg)–verb, present continuous tense, 1st person, singular;
կարդում (kardum) (VPres)– verb, present continuous tense;
Metalanguage output description for: Ես գիրք եմ կարդում: (Yes girk' em kardum)
SNT(PPronPrs1SgNEInfիAccSgAuxVPresContPrs1SgVPres)

E. Output language (Russian) formalized description for the metalanguage:
Я (ya) (PPronPrs1Sg)–personal pronoun, 1st person, singular;
читаю (chitayu) (VPresPrs1Sg)–verb, present tense, 1st person, singular;
книгу (knigu) (ND1fAccSg)–noun, declension 1, feminine gender, accusative, singular.
Metalanguage output description for: Я читаю книгу (Ya chitayu knigu).
SNT(PPronPrs1SgVPresPrs1SgND1fAccSg)

F. Output language (Turkish) formalized description for the metalanguage:
Ben(PpronPrs1Sg)–personal pronoun, 1st person, singular;
kitap(NNomSg)–noun, Accusative, singular;
okuyorum(VPrContPrs1Sg)–verb, present continuous tense, 1st person, singular.
Metalanguage output description for: Ben kitap**** okuyorum.
SNT(PPronPrs1SgND1fNomSgVPrContPrs1Sg)

NB
*–sentence (SNT)
**–in formal speech;
***–in informal speech;
****–indefinite noun in Turkish doesn't get an inflection ending in accusative case.

Some languages, however, require less grammatical rules to model sentence structures of output languages, thus metalanguage here is needed either for mainly input or output. Due to the capabilities of these languages a detailed modelled dictionary might be present prior to the translation.

As it can be observed here, many languages have the same grammatical

endings that are crucial for the encodification and decodification process – a unified, generalized action of splitting the words of one language into detailed segments, tagging them according their grammatical features, then matching these tags with the ones of the output language no matter the language family or linguistic typology.

This suggests usage of specific field indexation, extra information within modelling of the words to assist in enhance the effectiveness of the translation. This accounts for the specific text domain integrates into the formal definitions of words based on a hierarchical structure, thus clarifying the meaning of each word.

Conclusion

The development of the principles of component analysis are still ongoing. This method is considered promising from the point of view of communicative linguistics, which is combined to create communicative lexicography and to study the semantic components of the word. Derived from the lexical data, it can, in turn, assist in clarifying and refining explanations in encyclopedic dictionaries. Improving the method of extraction and description of meanings will help to re-discover the semantic circle of the word that expresses knowledge from the centuries-old human life experience.

Linguistic models can be both parsing and complying. The former is the extraction of the necessary information from the complete text by analyzing and identifying its elements, words, auxiliary particles and recognizing the correlations between them, whilst the latter is completing text information created by certain derivation rules from elements given prior. Every natural language has a complex and heterogeneous structure and is not suitable for modeling without the metalanguage.

Whatever the case, many systems are created every year around the globe, of which only a few gain general recognition. Cognitive systems are introduced into them to be essentially able to express the full meaning of the word that yet again confront the peculiarities of linguistic thinking. Whatever these systems are, the work of the specialists who strive to generalize and unify the knowledge of mankind by making it available to everyone should be properly evaluated.

References

Salevsky, H. (2018). The Origins of Interpreting in the Old Testament and the Meturgeman in the Synagogue.*The Bible Translator, 69*(2). (In Eng.)

Karaulov, YU. N., Molchanov, V. I., Afanasyev, V. A., Mikhalev, N. V. (1982). *Analiz metayazyka slovarya s pomoshch'yu EVM* [Analysis of the dictionary metalanguage using a computer]. M. (In Rus.)

Backus, J. W. (1959). *The Syntax and Semantics of the Proposed International Algebraic Language of the Zurich ACM-GAMM Conference.* Proceedings of the International

Conference on Information Processing. UNESCO. (In Eng.)
Hyvärinen, A., Karhunen, J., Oja E. (2001). *Independent Component Analysis. John Wiley & Sons.* (In Eng.)
Gulyga, Ye. V., Shendels, Ye. I. (1969). *Grammatiko-leksicheskiye polya v sovremennom nemetskom yazyke* [Grammatical and lexical fields in modern German.]. M., Prosveshcheniye. (In Rus.)
Cook, V. J. (1996). *Chomsky's Universal Grammar and Second Language Learning.* (In Eng.)
Roberts, I. G. (2016). *The Oxford Handbook of Universal Grammar.* Oxford University Press. (In Eng.)
Jahukyan, G. B. (2003). *"A Universal Theory of Language".* Caravan Books. (In Eng.)

Ենոք Գրիշկյան
Երևանի Պետական Բժշկական Համալսարանի դասախոս, բանասիրական գիտությունների թեկնածու, ասիստենտ
Էլ. հասցե՝ yenokgrishkyan@gmail.com

ՀԱՄԸՆԴՀԱՆՈՒՐ ԼԵԶՈՒ (ՄԵՏԱԼԵԶՈՒ, ԱՌԱՆՑ ՈՐԻ ՀՆԱՐԱՎՈՐ ՉԷ ԱՊՐԵԼ)

Մեր օրերում գիտելիքի մոդելավորումը դարձել է տարբեր գիտությունների կիրառման բնագավառների հիմնախնդիրներից մեկը՝ տրամաբանություն, լեզվաբանություն, մաթեմատիկա, հոգեբանություն, կիբեռնետիկա և այլն: Այս խնդրի նկատմամբ նման հետաքրքրության պատճառը արհեստական բանականության ստեղծումն է: (AI) համակարգեր. Հոդվածը քննարկում է մետալեզուն՝ որպես մեքենայական թարգմանության մեջ առկա ունիվերսալ՝ համընդհանուր լեզու, որը կոդավորում է մուտքային լեզուն, թարգմանում ու վերծանում ելքային ցանկացած լեզվի՝ անկախ լեզվական ընտանիքից կամ տիպաբանական բնութագրից:

Բազմաթիվ անվանի մտածողներ դեռևս միջնադարում փորձել են ստեղծել արհեստական լեզուներ, որոնք պետք է աջակցեին լեզուների միջև ճիշտ թարգմանությանը: Սրանք ենթադրում էին նաև լեզվի միասնական ձևային կաղապար՝ մետալեզու: Արդեն 20-րդ դարում լեզվական նյութը զետեղելով ալգորիթմների և մաթեմատիկական բանաձևերի մեջ, գիտնականները կարողացան ստանալ հնարավորինս ճշգրիտ թարգմանություններ, որոնց հիմնական նախապայմանը կարգաբերված և գործող մետալեզվի գոյությունն էր: Ինքնին մետալեզուն արհեստական լեզվական երևույթ է, որի օգնությամբ կատարվում են համակարգչային թարգմանությունները: Ի տարբերություն արհեստական կառուցվածքային լեզուների, մետալեզուն ընդհանուր է բոլոր արհեստական և բնական լեզուների համար և գործածվում է որպես օժանդակ միջոց՝ լեզվական տվյալները մեկ լեզվից մյուսին փոխանցելու համար:

Մետալեզվի օգտագործումը պահանջում է բառի մանրամասն կաղապարավորում կամ նկարագրություն ոչ միայն որպես առանձին, այլ որպես նախադասության և համատեքստի մեջ իր միջավայրում գտնվող այլ բառերից

կախված միավոր: Բառի մանրազնին նկարագրությունը՝ կաղապարավորումը կարևոր նշանակություն ունի իմաստի ճիշտ մեկնաբանման, հետևաբար՝ բառի, նախադասության և ենթատեքստի ճիշտ թարգմանության համար: Որպես օրինակ ներկայացված են անգլերենից չինարեն, իսպաներեն, լատիներեն, հայերեն, ռուսերեն և թուրքերեն թարգմանություններ՝ համապատասխան ձևային նկարագրություններով՝ հիմնված Համընդհանուր Ցանցային Լեզվի (ՀՑԼ) կիրառած մետալեզվի վրա: Հոդվածը քննության է առնում նաև ձևային նկարագրության կենսական մասերից մեկը՝ բաղադրիչի վերլուծությունը, քննարկվում է նաև ինչպես ձևաբանության, այնպես էլ իմաստային իմաստի տեսանկյունից:

Հիմնաբառեր. համընդհանուր լեզու, մետալեզու, բաղադրիչային վերլուծություն, մեքենայական թարգմանություն, ձևային նկարագրություն:

Енок Гришкян
Преподаватель Ереванского Государственного Медицинского Университета,
кандидат филологических наук, **ассистент**
Эл. адрес: yenokgrishkyan@gmail.com

УНИВЕРСАЛЬНЫЙ ЯЗЫК (МЕТАЯЗЫК, БЕЗ ЧЕГО ЖИЗНЬ НЕВОЗМОЖНА)

В настоящее время моделирование знаний стало одной из проблем в различных научных областях: логике, лингвистике, математике, психологии, кибернетике и др. Причина такого интереса к этой проблеме связана с созданием искусственного интеллекта (ИИ).

В этой статье обсуждается метаязык, форма универсального языка, которая существует в машинном переводе для кодирования входного языка, перевода и создания декодификации выходного языка любого живого языка, независимо от языковой семьи или типологических характеристик. Многие выдающиеся учёные ещё в средние века пытались создать искусственные лингвистические системы, чтобы поддержать языки для осуществления правильных переводов, главным условием которых было существование формализованного и функционирующего метаязыка, который сам по себе является искусственным лингвистическим явлением, с помощью которого осуществляется синхронный перевод. Моделирование или формальное описание слова имеет решающее значение для правильного перевода и интерпретации значения.

Использование метаязыка требует детального формализованного описания слова не только как отдельной единицы, но и как слова, зависящего от других слов в его среде внутри предложения и контекста. В следующей статье приведены примеры переводов с английского на китайский, испанский, латинский, армянский, русский и турецкий языки с помощью соответствующего формального описания, основанного на метаязыке,

используемом универсальным сетевым языком (UNL). Одна из важнейших частей формализованного описания – компонентный анализ – также обсуждается с точки зрения как морфологии, так и семантического значения.

Ключевые слова: Универсальный язык, метаязык, компонентный анализ, машинный перевод, формальное описание.

Հոդվածը խմբագրություն է ներկայացվել՝ 2024թ. հոկտեմբերի 04-ին:
Հոդվածը հանձնվել է գրախոսման՝ 2024թ. հոկտեմբերի 31-ին:
Հոդվածն ընդունվել է տպագրության՝ 2024թ. դեկտեմբերի 26-ին: