UDK: *811.111'42*

SOME REVIEWS ON MATHEMATIC LINGUISTICS

Laylo Davletnazarova,
davletnazarova@uzswlu.uz
EFL Instructor
Uzbekistan State University of World Languages

Abstract: Mathematics entered linguistics in the mid-20th century to solve practical problems, but it later became an integral part of linguistic research. It was formed as a special branch of linguistics and became a concept of quantitative linguistics. At the same time, other practical and theoretical sciences have been formed that allow to find solutions to practical problems related to the use of certain languages, and at the same time to reveal the laws inherent in the whole natural language.

Keywords: *Mathematics, applied linguistics, quantitative linguistics, natural languages and individual languages, morphemes.*

Here, we briefly review some specific areas of applied linguistics based on the methodology of quantitative study of linguistic structures of texts in natural languages and individual languages. In this regard, the methods of mathematical statistics are used here, in some cases they are very complex, in others they are very simple. We have decided not to enrich this chapter with various mathematical formulas, as all the details are given in the recommended literature to the students. Here we will try to describe some important aspects of quantitative linguistics only in general terms.

Today, linguists consider language to be a structural structure made up of elements that are connected to each other by certain relationships. Semantic units of language - primarily words and parts of speech (morphemes) - are signs consisting of exponential (indicator) and denotation, the exponent is considered as a material

carrier of meaning (sound image of the word) and at the same time o. denotation is a concept associated with a given indicator (meaning). The exponent and the denotation are linked, but this connection is conditional and optional. This means that in Russian the denotation "tree" is associated with the exponent of the tree, and in English the exponent is associated with the tree.

The elements that make up the exponent of words and morphemes are called phonemes. Phonemes are one-sided objects that have only a degree, not a character. Such objects are called numbers. Each language has a limited number of phonemes, but these limited phonemes can produce an infinite number of words.

The quantitative method is based on calculating the exponents of different units of language and their combinations.

Let's look at phonemes with minimal elements of language. Obviously, there are certain rules for combining phonemes into words. In addition, just as letters are used to write words, natural language symbols and transcripts can be used to write a sound image of a word: /d'er'eva/ и /tri:/.

Some analogues of word transcription, vague spelling and orthoepy rules, are used to create phonetic spelling rules and word spelling.

Phoneme combinations in different languages differ, so the compatibility of letters in different languages differs according to the rules. This principle is applied in **decoding** theory and methodology for identifying written language.

By studying the frequency of letters in different languages and calculating the frequency of letters in this text, we can determine in which language the text was written. If the language of the text is known and the text is simply encrypted, each letter of the text is replaced by another character (letter, number, picture), we can encrypt this text by calculating and comparing the frequency of each character. If a

Alisher Navoiy nomidagi Toshkent
davlat o'zbek tili va adabiyoti
universiteti

"KOMPYUTER LINGVISTIKASI:
MUAMMOLAR, YECHIM, ISTIQBOLLAR"
Xalqaro ilmiy-amaliy konferensiya

Vol. 1
№. 01 (2022)

language is unknown, we can first try to determine which language it is in the same way. You can calculate not only the frequency of individual letters and characters, but also the frequency of combinations of these letters or characters (using two, three, four, five, and other combinations). In this case, the speed and accuracy of decryption increases. With the help of computers, this method allows you to determine in a few seconds which language the text will be written. We can then use more complex text research algorithms in a language we already know.

On a modern computer, deciphering any alphabet ciphers is not so difficult for the user. Therefore, modern encryption systems (cryptography) use not complex exchange methods, but only complex mathematical methods that can decrypt using more complex mathematical methods (cryptanalysis).

Despite all this, the problem of encryption in linguistics remains relevant in relation to texts in languages unknown to us, in which there are only written (pictorial) monuments written in a script unknown to us. Simple statistics are not enough to uncover such monuments (written images). In-depth knowledge of how language structures can emerge and how to organize writing systems is required. But this information is not always enough. To uncover these monuments (written images), we need to have historical information about the people who created them, what they did, who they traded with, who fought. Studying the written monuments of neighboring nations can also be of great help.

In many cases, the names of rulers and the names of cities can serve as a key to uncovering inscriptions (images) that are unknown to us, as they are found in historical documents of the language of neighboring peoples. Hence, we can assume that they also encounter encrypted written monuments. In this way the French orientalist J. F. Champollion revealed the mystery of the Egyptian hieroglyphs. In the 20th century, the Crete of Mycenae and the Mayan hieroglyphics were encrypted.

But other written monuments, such as these, are still awaiting encryption, for example, inscriptions found on Easter Island.

Detailed information on decryption can be found in [1], [2].

In the 1950s and 1960s, American scientist Joseph Greenberg tried to measure the complexity of languages by quantitative methods by analyzing the classification of Edward Sapir species, and he proposed 10 typological indicators based on counting morphemes and number of words. calculated for languages [3]. Let's look at some of them for five different types of languages.

1. The first of these is called the synthesis index (or synthetic). It indicates how many average morphemes a word of a given language has and is calculated according to the following formula:

Syn = M / W, where M is the number of morphemes in the text and W is the number of words in the same text.

2. The next is the yield index. This shows how widely word formation is used using morphemes in a particular language:

Der = D / W, where D is the number of derivative morphemes in the text, W is the number of words in the same text.

3. The prefix index indicates how often prefixes are used in a particular language:

Pref = P / W, where P is the number of prefixes in the text, W is the number of words in the same text.

4. Accordingly, the suffix index indicates how often suffixes are used in a particular language:

Alisher Navoiy nomidagi Toshkent
davlat oʻzbek tili va adabiyoti
universiteti

**"KOMPYUTER LINGVISTIKASI:
MUAMMOLAR, YECHIM, ISTIQBOLLAR"**
Xalqaro ilmiy-amaliy konferensiya

**Vol. 1
№. 01 (2022)**

Suf = S / W, where S is the number of suffixes in the text, W is the number of words in the same text.

In table, you can see how different they are in all of these parameters in Russian and English. Russian is a more synthetic language, while English is more analytical. It is possible to see how similar the Russian and Yakut languages are, although Russian is inflected (fusional) - it mainly uses polysemantic inflected morphemes, and Yakut uses mostly unambiguous inflected morphemes, even though it is agglutinative. However, many morphemes are used in both languages for word formation and inflection, but there are no prefixes in Yakut. Vietnamese and Eskimo languages are the opposite of Russian, English and Yakut in all indices. Vietnamese is a highly analytical language and also does not use isolating, word-forming and inflective morphemes at all. The Eskimo language is agglutinative and polysynthetic, we observe that it has a high degree of word complexity, and we can see that both Russian and English are contradictory as inflected and isolated.

Table. Some typological indices on Greenberg

	Russian language	English language	Yakut language	Vietnamese	Eskimo language
Synthesis	2,33	1,68	2,17	1,06	3,72
Derevation	0,37	0,15	0,35	0,00	1,25
Prefix	0,17	0,04	0,00	0,00	0,00
Suffix	1,15	0,64	1,15	0,00	2,72

All five of these languages are very different systems, as different species or linguists point out. Differences between languages can be accurately described using special indices.

Although not all modern linguists dealing with language typology agree with such a classification, it is still used to describe languages in general.

Alisher Navoiy nomidagi Toshkent
davlat oʻzbek tili va adabiyoti
universiteti

**"KOMPYUTER LINGVISTIKASI:
MUAMMOLAR, YECHIM, ISTIQBOLLAR"**
Xalqaro ilmiy-amaliy konferensiya

**Vol. 1
№. 01 (2022)**

REFERENCES

1. Bredis M.A., Dimoglo M.S., Lomakina O.V. Paremias in Modern Linguistics: Approaches to Study, Text-Forming and Linguocultural Potential // RUDN Journal of Language Studies, Semiotics and Semantics. - 2020. - Vol. 11. - N. 2. - P. 265-284. doi: 10.22363/2313-2299-2020-11-2-265-284

2. Singh, K. (2007) Quantitative Social Research Methods. SAGE, New Delhi. https://doi.org/10.4135/9789351507741

3. Greenberg, J. H. (1963). Some universals of grammar with particular reference to the order of meaningful elements. In J. H. Greenberg (Ed.), Universals of Language (pp. 40-70). Cambridge, MA: MIT Press.

4. Khursanov Nurislom Iskandrovich. "On the theoretical and practical foundations of language corpora." Asian Journal of Multidimensional Research 10.9 (2021): 311-318. DOI: http://dx.doi.org/10.5958/2278-4853.2021.00675.3

5. Khursanov, Nurislom Iskandrovich. "On the theoretical and practical foundations of language corpora." Asian Journal of Multidimensional Research 10.9 (2021): 311-318. DOI: http://dx.doi.org/10.5958/2278-4853.2021.00675.3

6. L.R. Raupova, & Sh, Sultonova. (2021). Scientific Basis of Compiling a Dictionary of Grammatical Terms. Current Research Journal of Philological Sciences (2767-3758), 2(10), 116-124.