

*Materials of Conferences***USING SYNERGETIC THEORY OF INFORMATION FOR STRUCTURAL ANALYSIS OF TEXTS IN THE ASPECT OF THEIR RANDOMNESS AND ORDER**

Ospanova B.R., Azimbayeva Z.A.

*Karaganda State Technical University,
Karaganda, e-mail: o.b.r@mail.ru*

The article deals with studying synergetic theory of information for structural analysis of texts from the aspect of their randomness and order. In the work there are reflected the results of studies which purpose was the use of the information quantity measure that permits to analyze common mechanisms of the texts information-and-entropy characteristics lying in the base of all spontaneously running in the surrounding world processes of information accumulation that lead to the system structure self-organization.

The object of this study is a text as a multilevel natural object developing according to the synergy laws. A synergetic process in linguistics is evolving and assists the language enrichment, as the introduction in the language of synergetic ideas replenishes the language vocabulary with new definitions, categories, terms, that imposes its internal transformation and by this forms linguistic-synergetic scientific trends.

Studying the text hierarchic structure and methods of its information analysis is one of those urgent problems that are dictated by the need of using an objective, quantitative assessment of the grammatical system and semantic-and-syntactic organization of the text, as well as in the comparative analysis of kin and non-kin languages. In this aspect there was carried out the information-entropy analysis of a large mass of present day texts belonging to different genres, subjects and styles of the Kazakh language.

The approach to the text as to a hierarchic system permits to consider the text both from the point of view of its components analysis and from the point of view of their synthesis at the highest language level. In the home science the idea of the text integrity as of a hierarchically organized structure was for the first time presented by N.I. Zhinkin. "Any speech, – he noted, – can be reduced to a system of predicates that, consistently complimenting each other, reveal the structure and the ratio of the signs of the unknown before object of reality". "The text, – he wrote, – is separated into a hierarchic network of subjects, subsubjects, sub-subsubjects and micro-subjects"[1].

Thus, a text is structured according to certain laws integrity, consisting of the language units, i.e. sentences combined by a single subject and forming larger units, superphrase unities, thematic pieces of the text, paragraphs, chapters, sections, etc. that serve for presenting a certain completed content and some information. When defining the information quantity

there is considered the text that consists of letters, words, word combinations, sentences, etc. Each letter occurrence is described as a consistent realization of a certain system. The information quantity expressed by the indicated letter in its absolute value is equal to the entropy that characterized the system of possible choices and that was taken off as a result of selecting a certain letter.

It is known that for entropy calculation it is necessary to have the complete distribution of possible combinations probabilities. Therefore for entropy calculation of this or that letter it is needed to know each possible letter occurrence probability.

The language entropy is an important for linguistics measure. Entropy is a common measure of probabilistic-linguistic ties in the texts of a certain language. In this connection there was carried out comparison of the data characterizing the numerical estimation of these measures in Kazakh.

The information-entropy analysis of the text structure was carried out based on Shannon's entropy using the formula of probability classic determining.

In the general characteristic of the entropy-information (entropy is a measure of disorder, information is a measure of disorder elimination) analysis of the texts there was used Shannon's statistical formula for determining the text perfectness, harmony:

$$H = - \sum_{i=1}^N p_i \log_2 p_i, \quad (1)$$

where p_i is probability of detecting any system unit in their totality N ; $\sum_{i=1}^N p_i = 1$, $p_i \geq 0$, $i = 1, 2, \dots, N$.

Thus, we carried out a linguistic analysis of the texts containing 500 characters of scientific, journalistic, official, informal and artistic styles of the Kazakh speech.

To calculate the texts information there were counted the probabilities of occurring one letter, two-letter, three-letter, four-letter, five-letter and six-letter combinations. In counting there were taken into account 43 letters (42 letters, 1 blank) of the Kazakh alphabet, all the rest characters (brackets, quotation marks, commas, etc.) were not considered. Numerical data contained in the text are written in words.

The calculation of probability (p) of different letters occurrence in the text is achieved by calculating a relative frequency of individual letters. To determine the probability of occurring one letter in the Kazakh text there was used the classic formula of determining probability:

$$P(\text{oneletter}) = \frac{m}{n},$$

P is a relative frequency;

M is the number of one letter occurrence in the text;
 n is the number of all letters occurrence in the text.

As a result for the Kazakh language there were obtained the following values (in bits). Then, according to Shannon's formula

$$H_0 = \log 43 = 5,4 \text{ bit},$$

where H_0 is the maximum value of the text entropy consisting in receiving one letter of the Kazakh text (information contained in one letter) under the condition that all letters are considered equally probable.

When summarizing, we'll note that on the basis of Kazakh texts there were obtained the information characteristics of the letters that are in different positions; there were obtained the letter distributions of the text entropy, and given the possibility to estimate quantitatively the information ratio in the text. This permits to come to the conclusion that information entropy can be used to any language for revealing the information distribution in the text.

Entropy distribution in the Kazakh text

Entropy (E)	Scientific style of speech (SS)	Journalistic style of speech (JS)	Official style of speech (OS)	Informal style of speech (IS)	Artistic style of speech (AS)
H_1	4,3598	4,4253	4,3443	4,3873	4,3438
H_2	2,3444	2,7267	2,6006	2,7843	2,7468
H_3	0,852	1,0687	1,0225	1,0557	1,2596
H_4	0,2813	0,3301	0,2665	0,3187	0,414
H_5	0,1882	0,1198	0,2012	0,1265	0,1091
H_6	0,1657	0,0657	0,095	0,056	0,0414

From this there can be concluded that the dynamics of the text information entropy reduces when transiting to the higher level of organization; at this the text information content increases that proves the language development according to the law of preservation the sum of information and entropy.

References

1. Zhinkin N.I. Language. Speech. Creativity. M.: Labirint, 1998. – 368 p.
2. Shannon C. Works on theory of informatics and cybernetics / Under ed. of R.L. Dobrushin and O.B. Lupanov. M.: Publ. of Foreign literature, 1963. – 827 p.

The work is submitted to the International Scientific Conference "Modern science education", France, Paris, October 14–21, 2014 came to the editorial office on 11.09.2014.

PROFESSIONAL LINGUISTIC TRAINING OF A NON-LINGUISTIC COLLEGE STUDENT WITHIN A SYSTEM OF MULTI-LEVEL EDUCATION

Pendyuhova G.K.

Branch of "Moscow Psychologic-Social University",
Sterlitamak, e-mail: sterlitamak@mpsu.ru

Official documents of UNESCO identify educational system of Russia as a unique one for its fundamentality, scientific potential of the country remains high, regardless of the "lack of brain".

Russian fundamental education was created at foundation of knowledge paradigm. Many years of

practice have revealed significant disadvantages of this approach.

Comparative researches of education quality among university graduates of post-Soviet countries (Russia, Belarus, Ukraine) and developed countries of the West (USA, France, Canada, Israel), carried out by the World bank, have established that our students show very high results (9–10 points) according to such criterions as "knowledge" and "comprehension", and extremely low scores according to criterions "practical application of knowledge", "analysis", "synthesis", "evaluation" (1–2 points). Students from developed Western countries demonstrated completely opposite results [1].

According to objectives of Bolognese process, a review of state of qualifications in Russian education took place at the foundation of training professionals for a specific labour market in continuously complicating post-industrial society.

An employer prefers graduates with a good linguistic training who have an expertise in their professional area, feel comfortable in a linguistic environment and possess skills of communicating with foreign colleagues in context of professional activity.

In terms of multi-level educational model a problem of increasing efficiency of foreign language training among future specialists can be solved not only through organizational methods, but also via selecting the correct methodical strategy [3]. This strategy can involve linguistic-didactic, psycho-linguistic, and other aspects of linguistic training at each stage of education: profile class of