

WUT 2020

10th International Conference “Word, Utterance, Text: Cognitive, Pragmatic and Cultural Aspects”

SEMIOTIC AND SYNERGETIC METHODS OF TEXT ANALYSIS

Natalia Olizko (a), Natalya Mamonova (b), Maria Samkova (c)*

*Corresponding author

(a) Chelyabinsk State University, 129 Bratiev Kashirinykh st., Chelyabinsk, Russia, olizko@yandex.ru

(b) Chelyabinsk State University, 129 Bratiev Kashirinykh st., Chelyabinsk, Russia, nat2.mv@gmail.com

(c) Chelyabinsk State University, 129 Bratiev Kashirinykh st., Chelyabinsk, Russia, maria.a.samkova@gmail.com

Abstract

The article attempts to systematize methods most applicable to text analysis. The purpose of the research is to present methods within the quantitative and positional approach and the semiotic-synergetic approach to text analysis. The method of positional analysis as a probabilistic-statistical method is developed by G. G. Moskal'chuk and her colleagues. This is one of the methods within the psycholinguistic approach. It is based on the properties of the golden ratio. The method of positional text analysis is the most credible with texts that contain more than five words and is most effective in finding a text invariant. The semiotic-synergetic approach is based on Yuri Lotman's doctrine of semiosphere and a concept of synergetics of art created by A. V. Voloshinov. Within the semiotic-synergetic approach, the fractal theory as a special methodological variation is invented. Fractal modeling enables us to detect and investigate order and dynamic processes in the chaos of the text systems. From this perspective, text can be described as a complex self-organizing system, the fundamental characteristics of which are openness, non-linearity, and dissipativity. The results of the research prove that the linguosynergetic paradigm and the methodology of text analysis have been developed in the last decade but still remain an area that requires further study and clarification.

2357-1330 © 2020 Published by European Publisher.

Keywords: Synergetic linguistics, linguosynergetics, psycholinguistic approach, semiotic-synergistic approach, method of positional analysis, fractal modelling.



1. Introduction

In the second half of the 20th century, the fundamental changes in natural science give rise to the theory of self-organization or synergistics. The development of the science is often related to the investigations of statical and idealized abstractions or phenomena which evolve accurately but slowly or even remaining invariant for an observer. However, it is later noticed that if an equilibrium is upset in any system, this destabilizes system or an open system exchanging energy, matter, and information with the environment. In this case, a dynamic system can be classified as chaotic if a system is open, unstable, and sensitive to the slightest fluctuations, which are able to form qualitatively new structures that may cause restructuring of the whole system and the way it performs. In other words, the scenarios of the system's evolution become ambiguous.

Synergetics is a new interdisciplinary direction in modern science. The word 'synergetics' is derived from a Greek word 'synergètikós' meaning joint action, interaction. Synergetics combines "the theory of chaos, the theory of self-organization, the theory of complex systems, the theory of disasters, the theory of nonlinear oscillations and waves, and others" (Satiya & Hofstadter, 2016, p. 2).

In recent years, there has been a tendency to use mathematical and the methods of the natural science in humanities. Synergistics serves this aim because it is a kind of metascience, combining common laws of different fields of science: mathematics, physics, chemistry, philosophy, etc. Such synergetic ideas as nonlinearity, openness, and dynamic structurality of a language system are developed by linguists.

Synergetic linguistics or linguosynergetics (linguistic synergetics) is an actively developing paradigm of language cognition as a human, social, biopsychic, cognitive, and social phenomenon. Synergetic linguistics is the approach to study language, discourse, and text as a self-organizing system, which are closely interrelated with the external environment (the system of language, communicants' consciousness, communicative situation). In modern science, there is no strict distinction of terms 'synergetic linguistics', 'linguosynergetics', and 'linguistic synergetics.' They are used interchangeably.

The goal of this article is to study the results of the paradigm development in the last decade and to present the methods most applicable to text analysis.

2. Problem Statement

The most noticeable trends in synergetic linguistics are the psycholinguistic, contradictory-synergetic, and semiotic-synergistic approach. The psycholinguistic synergetic approach (Pishhal'nikova, 2018) is a pioneering approach that integrates synergetics, speech theory, and text analysis (Korbut, 2019; Moskal'chuk, 2016). The approach is based on the analogy of physiological, mental processes of meaning generation, text production and perception.

The contradictory-synergetic approach introduces the energy dimension. Its goal is to "study synergy (energy merging) of language units' self-movement taking into account their contradictions as a mechanism for the development of the system" (Myshkina, 2010, p. 87). There has been proposed the dynamic-system modelling of self-movement of language units' meaning and their energy properties. The

findings and results within the contradictory-synergetic approach differ significantly from synergetics. Therefore, we will not discuss this method.

The semiotic-synergetic approach (Olizko & Danilova, 2018) integrates synergistic and semiotic research principles. As a system component, discourse is included in the semiosphere (all sign sphere including text, language, and culture). Discourse self-organization implies moving from less complex forms to more complex ones. The semiotic approach, which considers stages and forms of system development, helps to outline the sequence of development stages, to identify relationships and interactions between different simple and complex forms, to explain the emergence of qualitatively new levels in discourse self-organization.

Based on the synthesis of synergetics and linguistics, new concepts and categories appear. However, the methodology of text analysis is not sufficiently developed.

3. Research Questions

Conducting the formal and content analysis of the current state, we will find out what achievements in this paradigm can be successfully applied for the text analysis. The data for the content analysis is collected via Google Scholar (2020), Elibrary (2020), which is the Russian scientific electronic library integrated with the Russian scientific citation index, and Mendeley (2020), which is a reference manager integrated with Scopus and ScienceDirect.

The keywords we searched for in titles, annotations, and keywords, were both in Russian and in English: semiotic-synergetic approach, linguosynergetics, linguistic synergetics, and synergetic linguistics. The time period is ten years.

Table 01. The number of publications in the thematic research field

Time period	Google Scholar	Elibrary	Scopus
2009-2010	1385	64	8
2011-2012	2106	60	11
2013-2014	2973	80	15
2015-2016	3096	91	17
2017-2018	2979	66	31
2019-2020	1041	28	1

Table 01 shows the keywords were most frequently used in publications in Russia in 2015-2016. The term that most scientists in Russia use is linguosynergetics. Most publications within this thematic research field were registered in Scopus in 2017-2018. Among the non-Russian scientists, the term synergetic linguistics is used more frequently than others. A quick overview of the articles results in the conclusion that Russian scientists expand the methodology of synergetics applying the fundamentals of synergetics to language analysis. While non-Russian scientists go beyond the traditional notions of self-regulation and self-organization and apply synergetic principles within the quantitative linguistics and language teaching.

The present paper has two research questions: 1) how the paradigm has developed in the last decade and 2) what the most effective methods of text analysis are.

4. Purpose of the Study

Within the psycholinguistic approach, a text is the result of the speech-making activity and therefore it is considered to have space-time parameters that integrate such notions as probability, symmetry/asymmetry, phase space, density, the distribution and sequence of text elements. A word form and a position as elements of detecting the text procedural structure are the units that measure text space-time parameters. The effectiveness of communication, emotional and psychological impact through text is largely due to its positional structure. The task is to determine the qualitative and quantitative probabilistic parameters, understand the patterns of spatial and temporal distribution as well as the functioning of text strong and weak positions. Such a quantitative and positional approach to text analysis optimizes the procedure of textual big data analysis, creates and improves digital tools and computer programs.

The task within the semiotic-synergetic approach is to reveal the interaction of the author's intentions, the reader's possible reactions and a text structure. Text as a developing synergistic system is the result of the interaction of different semantic structures within the semiosphere. In the space of the semiosphere, each specific fragment of the author's text is not a final version, but a combination of an infinite number of meanings. As a result, the reader has some associations that can generate a different from the author's message.

The purpose of this paper is to study and prove the effectiveness of the methods within the quantitative and positional approach and the semiotic-synergetic approach to text analysis.

5. Research Methods

The method of positional analysis is one of the methods within the psycholinguistic approach. It is a probabilistic-statistical method developed by G. G. Moskal'chuk and her colleagues based on the properties of the golden ratio as the optimal ratio of "compositional division into internal semantic parts" (as cited in Korbut, 2005, p. 30). The method detects the periodicity of the text structural organization, the patterns of text formation as the distribution of structure-forming elements that are repeated word forms which mark the strong and weak positions of a text, and the invariant that combines "the proportionality of dividing the whole text into composite zones of the beginning, the harmonic center, and the end in accordance with the proportions of the golden ratio" (Moskal'chuk, 2018, p. 61).

Based on the golden ratio, the intervals and the main positions of a text are determined. A text as a unit is divided into unequal segments, the first and the main position is placed at a distance of 0.618. It divides the text into larger and smaller parts. Then the text is divided by the same proportion, and a series of intervals is built: $1 - 0.618 = 0.382$; $0.618 - 0.382 = 0.236$; $0.382 - 0.236 = 0.146$; $0.236 - 0.146 = 0.09$. The following proportional volume of intervals is obtained: 0.618; 0.382; 0.236; 0.146, 0.09.

If the topic of a conversation changes and the communication moves from one text functional and semantic unit to another, the harmony of the text structural organization is achieved due to the optimality of the ratio based on the abovementioned proportion. The golden ratio is identified in the text structure when the principles of symmetry and asymmetry are implemented. They are predetermined by the structure of text lexical units and grammatical structures. The description of the text structuring process

begins with the symmetrical sections with an arranged structure and asymmetric sections with an unarranged structure. Symmetrical and asymmetric sections are identified by determining the text intervals with frequent lexical and syntactic repetitions as “text-forming factors, physical parameters of a text that determine its structure and self-similarity” (Moskal'chuk, 2018, p. 74). Repetition reveals “the structural, semantic, rhythmic, and intonation organisation of a whole text” (Moskal'chuk, 2018, p. 24).

The methodological basis of the semiotic-synergetic approach is, first, Yuri Lotman's doctrine of semiosphere, according to which semiotic space is “a multi-layered intersection of various texts that together form a certain layer, with complex internal relations, with different degrees of translatability and spaces of non-translatability” (Kull & Velmezova, 2018; Lotman, 2000, p. 30). Secondly, there is a concept of synergetics of art by Voloshinov (2002) that concerns the main provisions of fractal aesthetics and reflects the experience of analyzing art works from the standpoint of nonlinear thinking. The choice of this methodological approach is determined by the specifics of the studied subject since a text is associated with the complex activity of non-linear, self-developing, and self-similar semantic structures that function within the semiosphere. A text can be described as a complex self-organizing system, the fundamental characteristics of which are openness, non-linearity and dissipativity.

6. Findings

The method of positional analysis is applied to the numerical data (coordinates from 0 to 1) of the positions in the text that appear to correlate with a series of golden ratio proportions within the statistically proved limits (0.618; 0.236; 0.146; 0.09). This allows us to explicate and identify the text structure and form, regardless of its size, which is unified due to proportion. Therefore, the “positional” text model includes the harmonic center (0.618 of the volume of the whole text estimated in word forms), the harmonic center of the beginning zone (0.236 of the volume of the whole text), the beginning (0.146 of the volume of the whole text), and absolutely weak positions: the proportions of 0.236 of the whole text, deferred from the harmonic center of the text to the right ($0.618 + 0.236$) and left ($0.618 - 0.236$). The method of positional analysis shows controversial results with short texts like tweets which are less than six words: “the word form can be marked with a strong text position (the harmonic center of the beginning zone) and the absolutely weak position at the same time” (Khitina, 2018, p. 147).

The results of the positional method are usually compared with the results of the psycholinguistic experiment. The psycholinguistic experiment taken by Sherbakova and Golovina (2018) identifies the features of the perception of the text structure. 40 participants got a questionnaire with three texts of different genres and styles. The task to underline the semantically important words in the texts. The experiment resulted in 120 reaction texts and 1013 semantically important words. In the following stage, the results of both analyses were compared to find out whether these words match the words that were indicated with the method of positional analysis. This experiment reveals that a person better perceives certain areas of a text in which the golden ratio is located regardless of external factors. The most sensitive to perception is the area of the pre- and post-harmonic center of a text (Sherbakova & Golovina, 2018). However, the data obtained from the psycholinguistic experiment on the reaction on tweets (Khitina, 2018) differ from the results of the method of positional analysis. Each person selected the most significant units but not all of them coincide with the positionally marked ones. Khitina (2018) concludes

that everything depends on the interpretation and interprets: the more complex the text, the more diverse the interpretation is.

The mixed results are shown in Dorofeeva's (2004) research of 906 Russian poetic texts. For each of the poetic texts, a positional model is determined, the distribution of the repetition in the intervals is revealed, and lexical (thematic) repetitions as elements of symmetry are calculated in each interval of the analyzed texts. As a result, the text model of the dynamics of symmetry elements does not always correlate with the thematic repetition in the texts.

Within the semiotic-synergetic approach, the fractal theory as a special methodological variation has a fairly simple model called a fractal. It programs a rather complex multilevel system as a text or a discourse, through the procedure of identifying recursive semantic components and the procedure of identifying subjects of measurement of the fractal concept and subjects of measurement of theoretical constructs of a linguistic research. In this sense, Mandelbrot (1983) introduces a single complex representation into disparate clumps of models and facts, creating a fractal order of the interpreted world, more precisely, launching a mechanism of self-development, self-organization of this order.

The term 'fractal' comes from the Latin word 'fractus' that means consisting of fragments, and the corresponding Latin verb 'frangere' that means to break, i.e. to create the wrong fragments (Mandelbrot, 1983). The famous Norwegian physicist Feder (2013) clarifies the concept by calling "fractal as a shape made of parts similar to the whole in some way" (p. 11). Fractality of the object is defined by self-similarity: a small part of the fractal contains information about the whole fractal object; dynamism: an ability to self-develop; irregularity: fractal structures with increasing scale do not become simpler, the object will have an equally complex structure at all levels; recursivity: the process of repeating something in a self-similar way; and divisibility.

Fractal modeling enables to detect and investigate order and the dynamic processes that produce elements of chaos in the text systems and different types of discourses. Ideal self-similarity is considered as a given algorithm reproduced on different scales, at different levels, conditionally marked system micro-level, macro-level, and mega-level under analysis. An ideal self-similarity is achieved only in mathematical fractals. In linguistics, the system we consider also consists of micro-level (intertext), macro-level (discourse), and mega-level (interdiscourse). The ability of each level to be in relative equilibrium is determined by order parameters, i.e. the factors that control the functioning of the text and discourse system. The order parameters of the system convey elliptically the general sense of behavior and targets-attractions of the text/discourse system (Mamonova & Olizko, 2017). Order parameters that specify the macro-level of a language system are associated with the author's idea as a creative attractor that provides a stable state and gradual development of discourse. Variables that specify the language of the underlying micro-level that are various kinds of intertextual inclusions serve at the macro-level discourse as a structural chaotic material. The mega-level of the interface above the macro-level is organized by system-forming concepts, which act with control parameters at the macro-level (Mamonova & Olizko, 2017).

Iterations on the system micro-level and macro-level in the area of the creative attraction direct action (mega-level) can be represented in the form of fractal models of self-organization: 'concentric circles,' spiral,' 'rhizome,' and 'tree.' The creative attractor is represented by the local concept of the

investigated text system or discourse system. The system micro-level is described by semantic components directly present, appearing and disappearing as the system of discourse (macro-level) develops in the area of the creative attractor. In the border areas of different creative attractors (for example, the change of creative attractor) there are the phenomena of competition for the organization of semantic components, which in turn creates ambivalence of semantic development of text /discourse systems. At the same time, there is a transition from chaos to ordering and back, the same semantic components become a part of semantic structures controlled by one creative attractor and another.

Recursion cannot be dispensed with if fractality is discussed. Recursion provides an infinite repetition of anything in a self-similar way. The main feature of semantic structures of text/discourse systems is the combination of stability of simple speech patterns (in static), which provides stability of semantic components at the system micro-level, with instability in space-time coordinates. This phenomenon illustrates the predisposition of the text/discourse system to states of space-time dynamic chaos, which predetermines the ability of the semantic system to fractally self-organize. These processes are accompanied by the formation of dissipative structures, which are specified by the parameters of the studied system. In discourse analysis, such parameters should include the features of participants, goals, and spacetime as a collection of space and time coordinates.

7. Conclusion

We conclude that the method of positional text analysis is the most credible with texts that contain more than five words and is most effective in finding a text invariant. It definitely should be combined with the psycholinguistic experiment and the quantitative methods.

The fractal analysis allows the discourse to be explored as a dynamic semantic system of chaos by highlighting elements of order in the chaos itself. Determination of fractal structures within the semiotic-synergetic approach detects variability of language semantics with a sufficiently high proportion of accuracy at the most initial stages. The principles of synergetic linguistics explain why certain structures presented in the form of fractal models are repeated at particular stages of the semantic development of the language, discourse, and text, and reveal their importance in the processes of self-organization in nonlinear systems of different scales.

The research on the development of the synergetic linguistics paradigm in the last decade shows mixed results. Therefore, it is still at the developing stage and requires the instrumentalization of principles as well as further improvement of methods and techniques.

References

- Dorofeeva, V. A. (2004). *Strukturno-informatsionnaya ustojchivost' teksta k deformatsiyam ego ob"ema* [Structural and informational stability of the text to deformations of its volume]. (Doctoral dissertation). Kemerovo State University, Kemerovo, Russia.
- Elibrary (2020). Retrieved January 28, 2020, from <https://elibrary.ru/defaultx.asp>
- Feder, J. (2013). *Fractals*. Springer Science & Business Media.
- Google Scholar (2020). Retrieved January 28, 2020, from <https://scholar.google.com/>
- Khitina, M. V. (2018). Issledovanie vozmozhnosti primeneniya metoda pozitivnogo analiza teksta (Na materiale tvitov) [Study of the possibility of using the method of positive text analysis (based on the material of tweets)]. In V. A. Pishhal'nikova, L. R. Komalova, N. N. Germanova (Eds.),

- Yazykovoe bytie cheloveka i ehtnosa. Materialy XV Berezinskikh chtenij* [Linguistic existence of a person and an ethnic group. Proceedings of the XV Berezinsky readings] (pp. 144–147). Institut nauchnoj informatsii po obshhestvennym naukam RAN.
- Korbut, A. Yu. (2005). *Tekstosimmetrika kak razdel obshhej teorii teksta* [Textanimation as a section of the general theory of text]. [Doctoral dissertation]. Altai State University, Barnaul, Russia.
- Korbut, A. Yu. (2019). Psikholingvisticheskie pomekhi v protsesse vospriyatiya i ponimaniya teksta [Psycholinguistic hindrances in the process of perception and understanding of the text]. In L. N. Rozhentsova (Eds.), *Vektory razvitiya filologii v kontekste modernizatsii sovremennogo filologicheskogo obrazovaniya. Materialy vserossijskoj nauchno-prakticheskoy konferentsii* [Vectors of philology development in the context of modernizing modern philological education. Proceedings of the all-Russian scientific and practical conference] (pp. 33–39). Irkutsk: Asprint.
- Kull, K., & Velmezova, E. (2018). The paradox of the “semiotics of life”: Yuri Lotman’s later works. *Slovo.ru: Baltic Accent*, 9(4), 6–14. <https://doi.org/10.5922/2225-5346-2018-4-1>
- Lotman, Yu. M. (2000). *Semiosfera* [Semiosphere]. Iskusstvo-SPB.
- Mamonova, N. V., & Olizko, N. S. (2017). Fraktaly kak mehanizm vozniknoveniya porjadka iz haosa [Fractals as a mechanism for the occurrence of order from chaos]. *Filologicheskie nauki: Voprosy teorii i praktiki* [Philological Sciences: Issues of Theory and Practice], 5(71), 107–111.
- Mandelbrot, B. B. (1983). *The Fractal Geometry of Nature*. W. H. Freeman and Company.
- Mendeley (2020). Retrieved January 28, 2020, from <https://www.mendeley.com/>
- Moskal'chuk, G. G. (2018). *Struktura teksta kak sinergeticheskij protsess* [Text structure as a synergetic process]. 3ed. Editorial URSS.
- Moskal'chuk, G. G. (2016). Kak voznikayut sinergeticheskie ehffekty v tekste [How do synergistic effects occur in the text]. *PNRPU Linguistics and Pedagogy Bulletin*, 2, 8–16.
- Myshkina, N. L. (2010). Issledovanie funktsionirovaniya yazykovoj sistemy na osnove printsipov kontradiktno-sinergeticheskoy lingvistiki [Study of the functioning of the language system based on the principles of contradict-synergetic linguistics]. In N. L. Myshkina (Eds.), *Sinergeticheskaya lingvistika vs. lingvisticheskaya sinergetika. Materialy Mezhdunar. nauch.-prakt. konf* [Synergetic linguistics vs. linguosynergetics. Proceedings of International scientific and practical conference] (pp. 85–93). Izd-vo Perm. gos. tekhn. un-ta.
- Olizko, N. S., & Danilova, K. A. (2018). Semiotic and cognitive interpretation of the postmodern myth. In I. V. Denisova. (Eds.), *The European proceedings of social & behavioural sciences (EPSBS)* (pp. 537–542). Future Academy. <https://doi.org/10.15405/epsbs.2018.04.02.76>
- Pishhal'nikova, V. A. (2018). Nekotorye tendentsii razvitiya sovremennoj zarubezhnoj psikholingvistiki [Some trends in the development of modern foreign psycholinguistics]. *Vestnik of Moscow state linguistic university. Humanitarian sciences*, 14 (809), 93–104.
- Satija, I. I., & Hofstadter, D. (2016). The fractal family. In I. I. Satija, *Butterfly in the quantum world* (pp. 1–24). Morgan & Claypool Publishers.
- Sherbakova, M. V., & Golovina, E. V. (2018). Metod pozitsionnogo analiza kak sposob izmereniya vnutritekstovogo prostranstva-vremeni [Method of positional analysis as a way to measure intra-text space-time]. *Mir lingvistiki i kommunikatsii: ehlektronnyj nauchnyj zhurnal* [World of linguistics and communication: electronic scientific journal], 1, 156–170.
- Voloshinov, A. V. (2002). Ob ehstetike fraktalov i fraktal'nosti iskusstva [On the aesthetics of fractals and fractality of art]. In V. A. Koptsik (Eds.), *Nelinejnoe myshlenie v nauke i iskusstve* [Nonlinear thinking in science and art] (pp. 213–246). Progress-Traditsiya.