MODEL RESEARCH OF INTERACTION PROCESSES OF TEXT SPACES

Konstantin I. Belousov, Tatyana N. Galinskaya

Abstract: The article discusses the problem of interaction of text spaces. When discussing the interaction of text spaces we assume that there exists a certain text model. The technique of semantic charting and the method of positional analysis allowed us to represent the successive-simultaneous semantic space of a text as its "semantic outline". Owing to the method of the prosodic analysis of a text, aimed at modeling its prosodic outline, there appears the possibility to analyze the cooperative interactions of these relatively independent text spaces. The system-approached research program presented in the work is aimed at the study of the text as a polyontological, self-organizing spatiotemporal linguistic object. The multiaspect text analysis is grounded on a) the positional analysis method, b) quantitative methods which in there turn comprise such methods as c) correlation methods, which determine the text aspects' level. By comparing and contrasting synchronically semantic connection intensity and mean sound intensity of the obtained data we received the results that allow us to be more specific in the discussion of the text structure as an evolving process. The search for explanatory tools of convergence, divergence, intersection, overlapping of various text structures is the key to understanding the complex material, ideal and social nature of text, its presentation as wholeness.

Keywords: *system activity approach, modeling, positional analysis, semantic charts, semantic graph of a text.*

Introduction

The concept "spaces" of the text is common to many works studying the text. However, it is used rather figuratively as an opposition of a certain (closed) reality to another reality (cf, e.g. the semantic space of a text, the emotive space of a text, etc). Due to the metaphorical character of the term, it remains unclear how "spaces of the text" are related to the text and to any other of its spaces. Thus the basic notion – the text – is ambiguously defined. The multiaspect text analysis is grounded on a) the positional analysis method, b) quantitative methods which in there turn comprise such methods as c) correlation methods, which determine the text aspects' level.

Studying the spaces of a text we usually start from the idea of a text as a functional system, hence, a text is usually regarded as being focused on "achieving certain goals, accomplishing a certain extralinguistic task" [Leont´ev, 2001]. Being a linguistic product by nature a text is created for pragmatic purposes and realized in the

field of pragmatics. Such understanding of a text is quite pertinent, but much narrower than the notion of the functional system as seen by P. K. Anochin [Anochin, 1999]. According to the theory of functional systems, developed by Anochin, the useful result (the pragmatic purpose in terms of linguistics) gains the status of a factor which causes the individual components to interact "in accordance with the type of their cooperation." [Anochin, 1999]. However, the study of the mechanisms controlling the interaction of the components (or subsystems) by which the favorable result for the whole system is gained, is skipped in the research dealing with the functionality of a text. Thus functionality becomes another linguistic metaphor. Hence it seems reasonable to introduce the prime statements regarding the subject of inquiry. We understand the text as an integral polyontological linguistic object which exists in forms of space-time, thus the main attributes of the text are:

1) Existence in space-time reality. The text entirely displays its existence abiding by the basic patterns of matter motion, and its content is also revealed according to these patterns. Unlike the meaning of a sign the text content extends in space-time since it is developed by the purposefully organized chronotheme sequence of techniques, actions and operations, united by various methods and approaches to the text.

2) We consider text as a phenomenon having various levels of existence (various ontologies) that lie between the two poles: material (to which physical attributes of a text, disclosing its acoustic-wave ontology) and conceptual.

3) The text organization on different ontological levels has common features as far as the text "accomplishes" extralinguistic tasks. This fact makes it necessary to introduce the concept of cooperation of text spaces. The analysis and the subsequent synthesis of text spaces appear feasible owing to the system activity approach, the method of positional text analysis as well as theoretical constructs (form, structure, etc.).

In compliance with goals and problems pursued, every studied *object* occupies a certain *object domain* (aspect) which can be realized independently. There can be several domains of this kind. By means of classification, abstraction, analysis techniques and synthetic procedures every object domain appears as a system of hierarchically arranged elements connected with each other by relations of various nature. Therefore, referring to the system/structure, we mean the system/structure of a certain aspect, not the whole object. First it is necessary to single out an aspect in the given object before carrying out its system analysis. And the focusing of the aspect is in the basis of the activity approach. Thus combining the system and activity approaches is inevitable when studying the objects of the objective reality. That is why we should not use the term "system object", but the term "system aspect", which implies the integrity of both the system and activity approaches.

Positional Analysis Method of the Text

The essence of the positional analysis is in the marking of the language units in a linear sequence. As long as a text is limited extensively (has the beginning and the end) the beginning is taken as "0", the end as "1", irrespective of the size of a text. Owing to this convention we can compare texts of diverse sizes as well as results of multiaspect descriptions of the same text. All language elements of a given text are positioned linearly in a row. The word is considered a counting unit. In order to locate an element, one should define the position of a word containing a required element (in case segments are smaller than a word) or coinciding with it (in case the word "level" is studied) [0; 1]. Then a simple arithmetic operation is performed, that is dividing the ordinal number of the pertinent word, by the total number of words in the text. The research based on the principles of the system and activity approach comprises the following steps:

1) singling out certain aspects within an object and presenting them as abstract systems;

2) the quantitative description of them including the spatiotemporal order, indicating the appearance and functioning of a system aspect element (description by means of the method of positional analysis). The description itself is the study of an attribute´s appearance intensity (or probability) in every point of the temporal development of the text;

3) converting *absolute* number values of the process rate in a certain spatiotemporal position into *relative* number values, i.e. the values within the [0; 1] interval (in order to compare description results for different aspects). This operation is performed for all text aspects. If we present the description results graphically, we obtain the "outline" of the text aspects (semantic, prosodic, etc); [we understand the term 'outline' as graphic representation of text aspects (dynamic intensity in text unfolding - see Figure 3)].

4) By means of correlation methods of statistics (the Pearson correlation method in particular) and methods of graphical representation the obtained constructs from different text aspects are systematized due to their general ontological basis. Graphical representation allows to compare similarly modeled text outlines.

Thus the research strategy comprises a number of operations such as idealization, modeling, object abstraction. Besides, one of the major operations is the statistic analysis. The correlation methods of this analysis is a tool to integrate multifarious research. In this respect the status of correlation as a research method becomes philosophically important, since correlation acts as integration means for various aspects of one object.

Form of a Text and Text Structures

For the above described procedure of text aspect integration, the concept "form of a text" is suggested. Every aspect is manifested as a system by means of abstracting, analysis and modeling and is "endowed" with its own structure. Undoubtedly, structures of individual aspects may be different. However, these structures should have some common features as these are aspect structures of one and the same object. *This commonalty is a form which is inherent in the object.* While structure is a derivative from the research activity – that is – its construct, the form is the phenomenon. When we study the form of a text in its certain aspects it transforms into its structure. Thus we deal with *form projection on the aspect domain*. That's why form can be reconstructed by comparing its constituent structures. We point out general and occasional in its structures, which is not possible in case of monoaspect object description (See Figure 1).

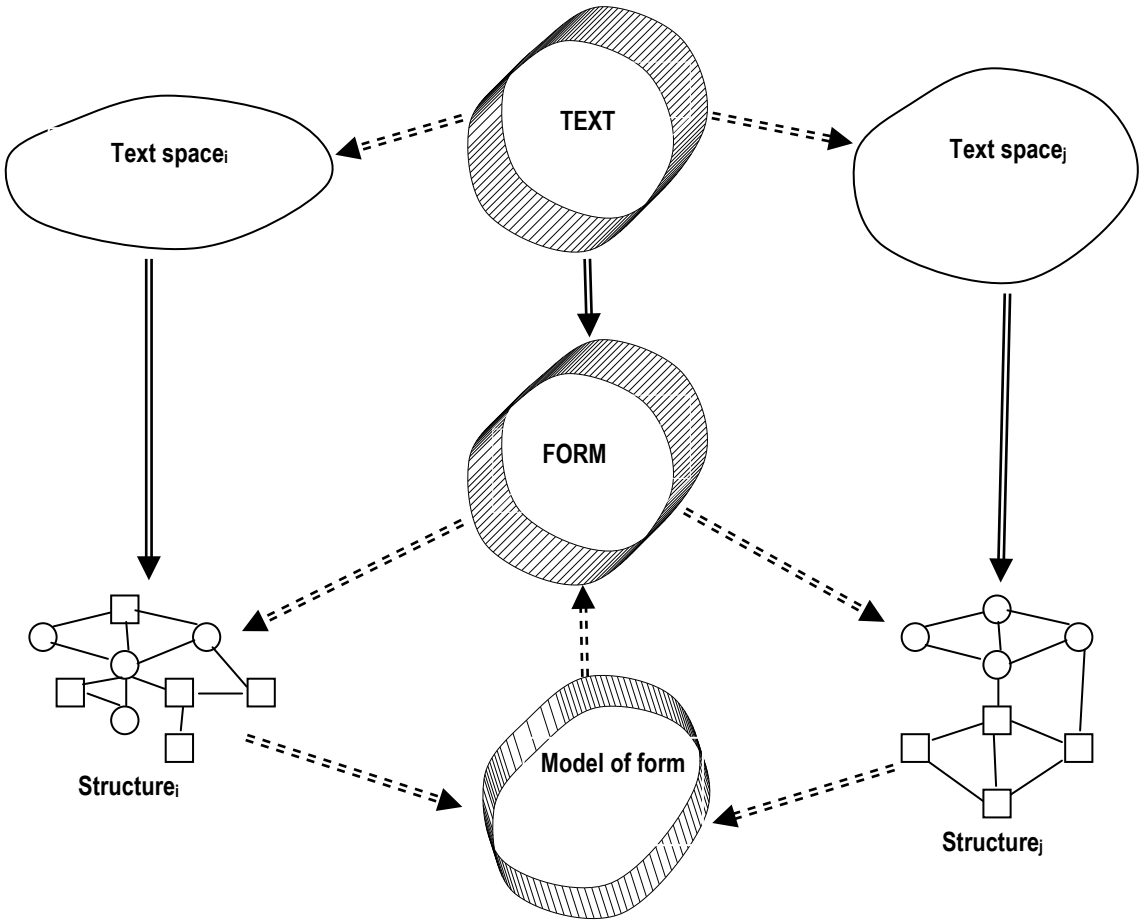

Figure 1. Form of the text and its structures

Semantic Charting of a Text

The study of the semantic text organization was carried out with the help of semantic charting of text – the method which has been elaborated by the author of the article. The semantic text charting is carried out according to the results of text assignments, performed by informants. The informants (21 philological faculty students) are given the text and assigned 1) to read the text and to define its theme; 2) to identify microthemes of the text; 3) to distribute the words of the text into semantic groups according to the identified microthemes, 4) each group should be entitled (the informants are expected to explain the grouping principle; the number of groups and words referred to them being unlimited and randomized within a certain text).

We want to specify again that the informants can include any word into any microtheme, which means that not only words but microthemes can be located in different parts of any text. Hence one can observe microtheme overlapping in some parts of text.

The obtained interpretations consist of semantic groups based on microthemes. The included words within one microtheme are linked with each other. Any word may be included in different microthemes. Word ability to be included into several microthemes allows us to speak about semantic connection intensity. The more frequently two words are included into one microtheme, the more semantic connection intensity they manifest. Semantic charting of the text shows semantic connection intensity of each word with other words of the text (based on all submitted interpretations) (see Table 1). In the given semantic chart the words from vertical and horizontal lines are taken from the text below. The figures at the crossing manifest the semantic connection intensity.

Лето умирает. Осень умирает. Зима – сама смерть. А весна постоянна. Она живет бесконечно в недрах вечно изменяющейся материи, только меняет свои формы. (В. Катаев)

Summer dies. Autumn dies. Winter is the death itself. And spring is constant. It lives infinitely in the womb of perpetually changing matter, it only changes its form. (V. Katayev)

It is evident that words can be connected either regularly or accidentally. The level of significance can be defined statistically. It should surpass the sum of mean value and mean square deviation of the semantic connection intensity for each word. For example, the word *лето* has the following semantic connection intensity with the other words: *лето* (0), *умирать* (5), *осень* (10), зима (8), *сама* (0), *смерть* (3), *а* (0), *весна* (5), *постоянна* (0), *она* (0), *жить* (2), *бесконечно* (2), *недра* (2), *вечно* (0), *изменяющийся* (0), *материя* (0), *только* (1), *менять* (1), *свои* (1), *форма* (0).

For instance, the cell of the column «жить» and the line «весна» gives the value «7». It means that these words were placed by the informants into one semantic group (microtheme) seven times (of possible 21 times according to the number of the informants).

The mean value of the semantic connection intensity of the word *лето* with other words equals to 2. The mean square deviation, equals to 2,88 (1-sigma). Hence semantic connection intensity can be regarded nonrandom as it exceeds 4,88. All the rest figures are accidental, that's why they are not considered.

Table 1. Fragment of the semantic charting Лето…

Word/word	лето	умирать	осень	зима	сама	смерть	а (but)	весна	постоянна	она (it)	жить	бесконечно	недра
лето (summer)	0	5	10	8	0	3	0	5	0	0	2	2	2
умирать (die)	5	1	7	4	0	7	0	0	0	1	2	1	1
осень (autumn)	10	7	1	9	1	5	1	3	0	1	0	1	0
зима (winter)	8	4	9	0	2	8	1	4	0	1	0	2	0
сама (itself)	0	0	0	2	0	3	6	3	2	8	0	2	4
смерть (the death)	3	7	4	8	2	0	1	0	0	1	2	3	1
а (but)	0	0	1	1	6	2	1	3	2	6	1	0	1
весна (spring)	5	0	4	4	3	0	2	0	8	3	7	4	2
постоянна (constant)	0	0	0	0	2	0	2	8	0	2	6	5	2
она (it)	0	1	1	2	8	1	6	4	2	0	2	1	3
жить (live)	2	1	0	0	0	2	2	7	6	2	0	6	3
бесконечно (infinitely)	2	2	1	2	2	3	0	5	5	1	6	1	1
недра (womb)	2	1	0	0	3	1	1	2	1	4	3	2	0

Semantic Graph of a Text

Due to the described technique the quantity of semantic connections of the words decreases substantially. It makes possible to use graphical means for the representation of semantic connections. To achieve it we place the words on the plane and show regular connections between them using connective lines (see Figure 2).

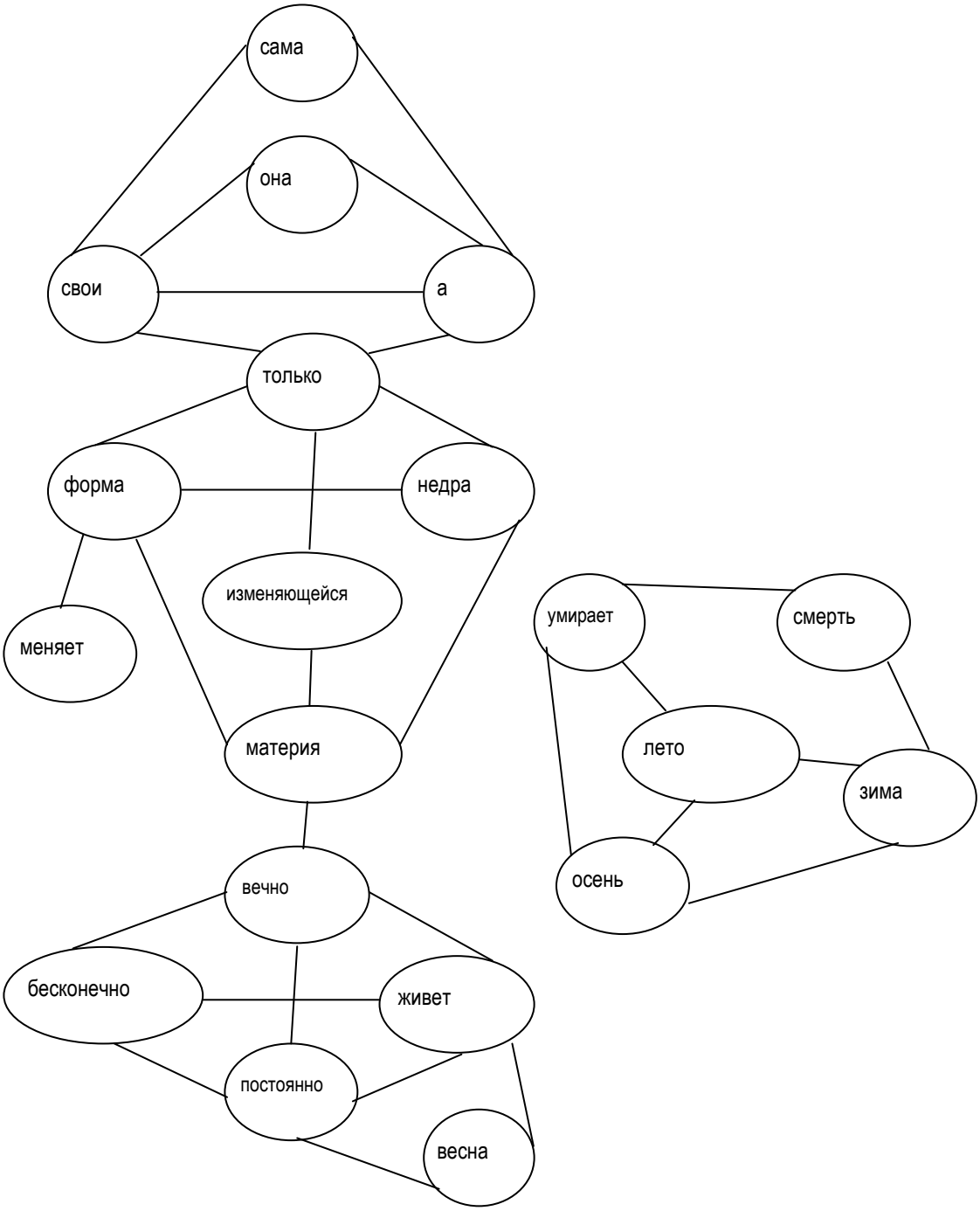

Figure 2. The semantic connections between the words of the text Лето…

In the Figure 2 we see that the semantic space structure consists of two substructures. The smaller one is the substructure, composed of the words *лето*, *осень*, *зима*, *смерть*, *умирать*, which can be conventionally denoted as the semantic field "death". This semantic field is formed only on the ground of contextual connections. The structure is homogeneous, no kernel element (the element having the greatest number of connections with other components of the same substructure) can be singled out in it. It is interesting to notice that this substructure is not connected with the dominant structure at all.

Evidently the "disconnected" semantic space provides the explanation of the fact, that there is considerable diversity in defining the theme of the text in the informants' interpretations (cf. some of the interpretations: *The text is about eternity of the spring. The text is about the spring being not only a season but something more. The text is about forms, which we, silly people, regard as different seasons. The text is about death of the seasons and eternal life of the spring. The text is about the infinity of the changeable creating spring .etc*).

As to the dominant substructure, it is discrete in the long run. So, we can single out the substructure formed by the words *вечно, бесконечно, постоянно, жить, весна*, which can be denoted as the semantic field "life". The substructure has the kernel element shown graphically. The kernel of the substructure, the word *вечно*, is the mediator between the given substructure and the main part of the dominant substructure. The dominant substructure also includes two smaller ones formed by the words 1) *материя, изменяющийся, недра, форма, только, менять* and 2) *только, свои, она, а, сама*. The difference between them is based on the functions performed by their components: those of the first mentioned substructure have a nominative function and components of the second substructure realize deictic and relative functions.

The tendency to be discrete is also seen in its syntactical organization. One can observe syntactical parallelism in the initial sentences which allows them to be more free compositionally than in case of chain connection. Having the adversative meaning the conjunction *а* also contributes to its discrete character as it divides the lingual matter into two parts. The conjunction *только* has the similar effect, although it has no adversative meaning, but it determines more precisely the notion *бесконечность* (*жизнь весны*) as infinity of its changing forms. And continuity is most clearly realized as the phenomenon, surmounting discontinuity which is observed most vividly in the syntagm *Она живет бесконечно в недрах вечно изменяющейся материи*. Here, on the sentence level, the text becomes syntactically deep, as opposed to the plane relations with coordinative connection only, realized in the first sentences (the tendency to discontinuity prevented them from expanding).

The kernel component in the substructure *материя, изменяющийся, недра, форма, только, менять* (field «материя») is the word *форма*, since it has the largest number of connections (6). However some of this word's connections are of little importance. These are the connections with the irredundant element *менять*, with the words *только*, *свои*. The latter components are the mediators of the substructure *только, свои, она, а, сама*, which is apparently important grammatically, rather than semantically. The most relevant element of the field «материя» ("matter") is the word *материя*, because it has the same function as the word *вечно*.

All this allows us to assume that the semantic space of the text is highly discrete. And surmounting this discontinuity allows the text to exist as a whole, being realized with the help of the mediating elements of the substructures "жизнь" ("life") and "материя" ("matter"). Thus the category of "wholeness" is realized by means of connections of kernel components *материя*, *вечно* and a number of adjacent components (*бесконечно, жить, постоянно, изменяющийся, форма, недра*).

Semantic Outline of a Text

With the semantic charting of the text and structural analysis of its semantic organization we can reconstruct the semantic outline of the text, which models the semantic connection intensity between words in the *progressive linear text development.* The following procedures have been implemented.

1. All the words (including recurrent ones) in their linear sequence in the text are placed in individual cells.

2. Only regular connections between words should be shown in the semantic charting of the text.

3. The regular semantic connections between text elements are put down in all cells situated between these words. If the words *лето "summer"* and *весна "spring"* have the intensity of connections which is equal to "5" (see Table 1) this meaning should be situated in the whole cell between these words. Thus we assume that *the semantic connection between two words in a text as a linear object inevitably covers all the space between the lexemes*. In case of recurrence of the words the procedure is to be done again.

4. Having completed this procedure the number of semantic connections are summed up. Summing is carried out within every individual column, having a word of the text in its top cell. Thus we can observe semantic connection density at any moment of the linear text development in the process of its perception (see the result of constructing semantic outline in Table 2).

Physical Outline of a Text

To scrutinize the *physical organization of a text* we used audio software which allowed us to analyse the prosodic aspects of the text. In the course of the research the informants were asked to recite the given text, as the recital assumes its interpretation. Thus the semantic interpretation is realized by sound intensity, pitch and duration of a sound. Only the first property was taken into account due to the systemic character of the speech intonation because of interdependency of all three above mentioned properties. Audio recording equipment allowed us to keep distance between mouth and microphone to obtain valuable material for the analysis. After the recording, the speech properties were analyzed by means of the specialized programs *Cool Edit Pro* and *Excel* according to the following algorithm: 1) the acoustic wave visually represented by Cool Edit Pro was segmented into parts, each of them being equal to the recorded words; 2) the points of maximum sound intensity (maximum amplitude) of the stressed syllables were determined within every segment, and the maximum values of the acoustic wave

intensity were put down in the Excel table. Thus every recital was processed and presented in the summary table, where every recital was represented as a dynamic outline of the text. Absolute values were converted into relative ones in order to compare different recitals with each other. Then the mean values were calculated from all relative values of sound intensity which we call the prosodic outline of the text.

Synchronization of Semantic and Physical Outlines of a Text

By comparing and contrasting synchronically semantic connection intensity and mean sound intensity of the obtained data we received the results given in Table 2.

Table 2. Semantic connection intensity and sound intensity values for the text *Лето…*

	Semantic connections	Sound intensity (mean)
Лето (summer)	0,33	0,96
умирает (dies)	0,60	0,75
Осень (autumn)	0,85	0,96
умирает. (dies)	0,82	0,73
Зима (winter)	0,77	0,76
сама (itself)	0,73	0,76
смерть.(death)	0,73	0,69
А (and)	0,60	0,71
Весна (spring)	0,70	0,84
постоянна.(is constant)	0,86	0,74
Она (it)	0,83	1,00
Живет (lives)	0,80	0,90
бесконечно (infinitely)	0,73	0,85
в недрах (in the womb of)	0,85	0,93
Вечно (perpetually)	0,90	0,72
Изменяющейся (changing)	0,93	0,84
материи,(matter)	1,00	0,67
Только (only)	0,93	0,93
Меняет (changing)	0,89	0,87
Свои (its)	0,79	0,72
Формы (forms)	0,42	0,63

For easier perception all the data from Table 2 are presented graphically (see Figure 3) The X-axis indicates the words in their linear sequence and their location in the text. The Y-axis indicates the intensity of semantic and prosodic processes (1 being the maximum value). Figure 3 shows the earlier stated differences and similarities in the above-mentioned aspects of the text.

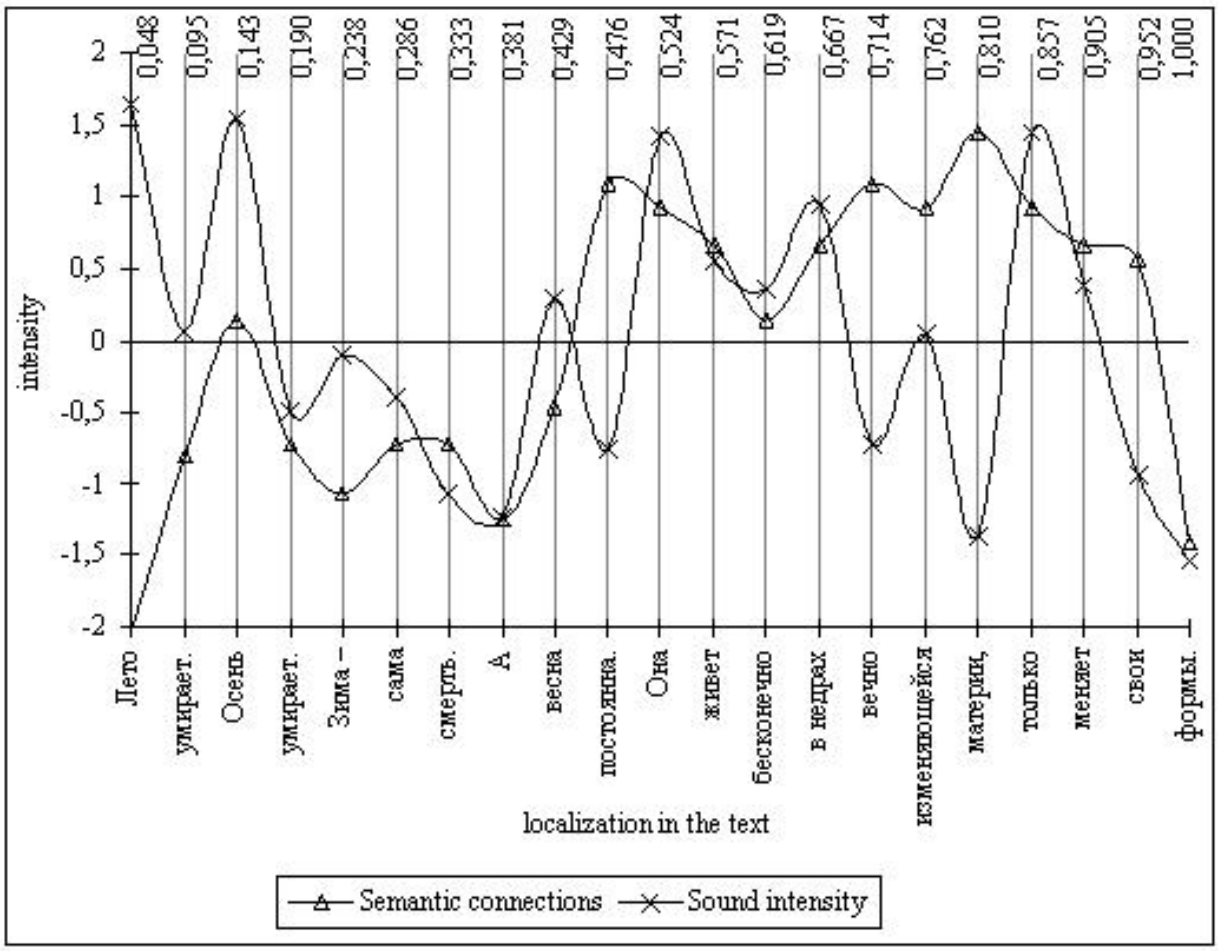

Figure 3. Intensity of semantic and prosodic processes in the linear development of the text.

We can see the three-part composition of the text in Figure 3. The three semantic maxima are situated between the semantic minima: *осень* – between *лето* and *а*, *постоянна* – between *а* and *бесконечно*, *материя* – between *бесконечно* and *формы*. According to the positional analysis of the text every text has the following position structure: the absolute beginning (AB) and the absolute end (AE), between which there is the harmonic centre of the text (HC, at the distance 0,618 from the beginning of the text), the harmonic centre of the initial zone (HCiz, at the distance 0,236 from the beginning of the text), the Setting (at the distance 0,146 from the beginning of the text), the absolutely weak positions (AWP_1, AWP_2, at the distance 0,236 to the right and to the left of the

HC) (see Figure 4). The mentioned above meanings are proportions of "golden section" which we use to define position structure of the text [Belousov, 2009].

We believe that a text is constructed according to the laws of harmony (proportions of "golden section") what affects text structure as a process and result of speech activity. This adjustment is known as text self-organization. One of self-organization markers is the creative attractor (the location in a text structure which has the greatest intensity of semantic process) of the text, understood as such extension in linear text space, in which self-organization processes are the most evident (explicit maximally). However the problem of the correspondence of the creative attractor to the units of language and semantic levels remains open. In Figure 3 we can see, that the HC of the text is the word *бесконечно*, HCiz – *зима*; AWP_1 is between *а* and *весна*, AWP_2 – between *только* и *меняет*. The position of the setting is between *осень* and *умирает*. The creative attractor occupies the interval between the HC and AWP_1 (*бесконечно в недрах вечно изменяющейся материи, только*). The above mentioned three-part composition of the text partially corresponds to its location structure: the initial zone of the text (from AB to AWP_1) is marked with the minimal number of semantic connections both to the left (the initial word) and to the right (the position between *а* and *весна*), i.e. it stands out of the whole structure; while the final zone of the text (from AWP_2 to AE) is not marked so clearly. There should be a considerable decrease of semantic connections intensity in absolutely weak positions, as their basic function is segmentation of the text structure. Something of this kind can be observed in the position AWP_1, while there are no distinctly seen mechanisms of discontinuity in AWP_2. Probably, the borders of the positions determined in the invariant must vary due to language substratum inertness: in our case the considerable drop in semantic connections intensity falls on the last word, and in this case AWP_2 should be between the words *свои* and *формы*. Regarding the harmonic center of the text (in Figure 3 HC is the word *бесконечно* "*infinitely*"), we see that semantic connection intensity in them appears to be of mean value. Harmonic centers had similar organization while we had analyzed other aspects of the text, its prosodic organization in particular.

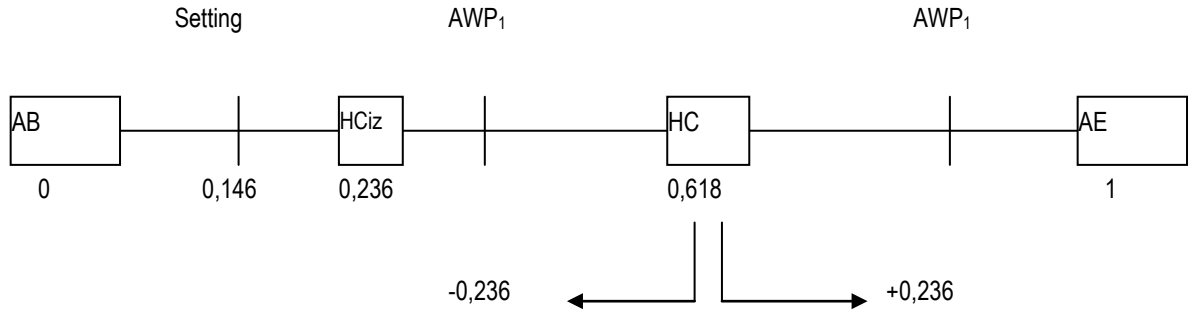

Figure 4. The position structure of the text

The creative attractor is situated in the interval with the maximum value of semantic connections intensity. In this case the attractor is in the interval where the intersection of all the semantic connections is the most evident, reflecting wholeness of this text in the utmost degree, what has already been discussed. However such processes are not characteristic for all texts.

As it is seen in Figure 3, the independent outlines begin to change interactively in the setting (0,146) of the text, and it can be interpreted as the attempt to coordinate their temporhythmical evolution parameters. We have already observed this function of the setting when we compared prosodic and emotional outlines of the text. It is interesting that the prosodic organization intensity in this interval is higher than the semantic one. In the HCiz the intensities of forming processes "equalize" for the first time and progress synchronically up to AWP_1 [cf. Belousov, 2008]. AWP_1, being the minimum point of the semantic and prosodic outlines, is the turning point of the development from which we can observe the synchronic growth of semantic and prosodic intensity. The prosodic processes dominate (their relative meanings exceed relative meanings of the semantic process (see Figure 3)) i.e. the semantic connection intensity tends to adjust to patterns of the physical aspect of the text development. Apparently such interaction can be called "suggestion". The coordinated intensity alteration of the two processes, registered in the interval around HC, indicates the harmonization of the processes, the stability point (the minimum) falling on HC. The area of the creative attractor (the maximum of semantic connections intensity) is the place of the most pronounced discrepancy of the semantic and prosodic outlines. The dominating role (their relative meanings the semantic process exceed those of the prosodic process) of the semantic component against the background of the prosodic phenomena shows the significance of cognitive mechanisms as opposed to emotive ones in this text self-organization. Strange as it may seem the area of the creative attractor comprises the words with abstract meaning: *бесконечно, недра, вечно, изменяющийся, материя*, which bring to the forefront rational-logical elements in the structuring of this text's wholeness. Besides, this interval has much more complicated syntactical structure than other intervals of the test. And finally in AWP_2 the intensity parameters of the both processes intersect and overlap each other to gradually dwindle synchronically up to the end of the text.

Conclusion

Thus the technique of semantic charting and the method of positional analysis allowed us to represent the successive-simultaneous semantic space of a text as its "semantic outline". Owing to the method of the prosodic analysis of a text, aimed at modeling its prosodic outline, there appears the possibility to analyze the cooperative interactions of these relatively independent text spaces.

By comparing and contrasting synchronically semantic connection intensity and mean sound intensity of the obtained data we received the results that allow us to be more specific in the discussion of the text structure as an evolving process. The search for explanatory tools of convergence, divergence, intersection, overlapping of various text structures is the key to understanding the complex material, ideal and social nature of text, its presentation as wholeness.

Bibliography

[Anochin, 1999] P.K. Anochin. Poleznyj rezul´tat kak organizujuščij faktor sistemy. In: Sinergetika i psichologija 2. Social´nye processy: 34-37. Moscow: JANUS-K2. 1999.

[Belousov, 2008] K.I. Belousov. Sinergetika teksta: ot strustury k forme. Moscow : Editorial URSS. 2008.

[Belousov, 2009] K.I. Belousov. Teorija i metodologija polistrukturnogo sinteza teksta. Moscow : Flinta: Nauka. 2009.

[Leont´ev, 2001] A.A. Leont´ev. Jazyk i rečevaja dejatel´nost´ v obščej i pedagogičeskoj psichologii. Moscow: Moskovskij psichologo-social´nyj institut. 2001.

Authors' Information

Konstantin I. Belousov - Orenburg State University, Department of Philology, Professor, Chief of Philological Modeling and Design Laboratory; P.O. Box 13, Prospekt Pobedy, Orenburg, 460018; e-mail: konstan-bel1515@yandex.ru

Major Fields of Scientific Research: Psycholinguistics and Psychosemantics, Semiotics, Semantics and Pragmatics, Cognitive Mechanisms for Text Processing

Tatyana N. Galinskaya - Orenburg State University, Department of Philology, Associate Professor; P.O. Box 15, Pereulok Yasnyi, Orenburg, 460036; e-mail: galinskaya@rambler.ru

Major Fields of Scientific Research: Psycholinguistics and Psychosemantics, Semiotics, Semantics and Pragmatics, Cognitive Mechanisms for Text Processing