



DOI: 10.22363/2687-0088-2021-25-1-147-164

Research article

## Linguistic creativity and discourse profiles of English language children's novels

Maria I. KIOSE

Moscow State Linguistic University  
*Moscow, Russia*

### Abstract

Contemporary discourse studies face the necessity to develop the methods of contrastive sub-discourse analysis which apply numeric and comparable data to diversify and describe sub-discourse types. The aim of the research is to propose a method of discourse profiling serving the purpose, and to further test the method in the contrastive study of linguistic creativity in different types of English language children's novels. The category of linguistic creativity being the leading form of language poesis receives parametric description on all language representation levels (in written form) and provides the discourse values for contrastive analysis. These values are explored in the fragments of non-autobiographic and autobiographic adventure novel sub-discourse types authored by J.C. George, F. Gibson, J.D. Ullman, and G. Durrell (a total amount of 120.000 signs) annotated manually for 52 linguistic creativity parameters on phonological, morphological, word-formation, lexical, syntactic, and graphic levels. The working hypothesis is that the linguistic creativity parametric activity distributions represent the sub-discourse profiles and may serve to contrast sub-discourse types by means of their vectors' contingency values. The analysis in individual parameter activity and in parameter groups activity demonstrated significant variance in sub-discourse construal, with autobiographic sub-discourse of G. Durrell manifesting several higher activity values in word-formation (occasional compounding), lexical use (the use of professional language, lexical tropes, allusive names, higher register style) and syntactic use (the use of parallel structures and syntactic intensifiers). In terms of morphological activity, the parameter values tend to be lower (morphological category shifts), the same stands true of some syntactic (the use of elliptical structures) and lexical parameters (the use of lower register types and proper names). The sub-discourse profiles demonstrate several common features, evidently typical of the discourse type itself, and the features differentiating non-autobiographic and autobiographic discourse subtypes. Vector correlation analysis revealed lower correlation values for autobiographic sub-discourse, which proves its specificity and testifies to the discourse profiling method applicability.

**Keywords:** *linguistic creativity, discourse profiles, children's novels, Gerald Durrell, vector analysis*

### For citation:

Kiose, Maria I. 2021. Linguistic creativity and discourse profiles of English language children's novels. *Russian Journal of Linguistics* 25 (1). 147–164. DOI: 10.22363/2687-0088-2021-25-1-147-164

## Лингвокреативность и дискурсивные профили художественной повести в англоязычной детской литературе

М.И. КИОСЕ

Московский государственный лингвистический университет  
*Москва, Россия*

### Аннотация

Для современной лингвистики дискурса актуальной является проблема отсутствия методов контрастивного анализа субдискурсов, применение которых позволило бы оперировать числовыми и сопоставимыми данными при их разграничении и описании. Цель исследования состоит в разработке метода дискурсивного профилирования, который тестируется при проведении контрастивного анализа лингвистической креативности в разных типах дискурса детской англоязычной повести; при этом лингвистическая креативность как ведущая форма языкового поэзиса изучается через параметрический анализ ее языковых маркеров на всех уровнях языковой репрезентации письменной формы дискурса. Материалом исследования выступают образцы неавтобиографического и автобиографического субдискурсов детской англоязычной приключенческой повести Дж. Джорджа, Ф. Гибсона, Дж. Ульмана и Дж. Даррелла общим объемом 120 тыс. знаков, которые подверглись аннотированию по 52 параметрам лингвистической креативности (с использованием процедуры ручного аннотирования) на фонологическом, морфологическом, словообразовательном, лексическом, синтаксическом и графическом уровнях. Гипотеза исследования заключается в том, что распределение частоты, или активности параметров дискурс-структурирующей категории лингвокреативности может быть использовано для проведения контрастивного анализа субдискурсов при установлении степени сопряженности их многомерных векторов лингвокреативности. Результаты анализа активности индивидуальных параметров и их групп позволили определить ряд значимых отличий в субдискурсах: автобиографический дискурс Д. Даррелла демонстрирует повышенную активность параметров на словообразовательном уровне (использование словосложений), лексическом уровне (использование терминологии, стилистических тропов, аллюзивных имен, переключений регистров), синтаксическом уровне (использование параллельных структур и интенсификаторов). Сниженная активность обнаружена у некоторых морфологических параметров (категориальные трансформации), синтаксических параметров (использование эллиптических конструкций) и лексических параметров (переключение на разговорный регистр, использование имен собственных). Наглядно представленные профили субдискурсов демонстрируют как отличительные, так и схожие проявления лингвокреативности в рассматриваемых дискурсах. Проведенный анализ корреляций векторов лингвокреативности субдискурсов показал сниженный коэффициент сопряженности в отношении автобиографического фрагмента, что свидетельствует в пользу результативности метода дискурсивного профилирования как способного разграничивать субдискурсы.

**Ключевые слова:** лингвистическая креативность, дискурсивные профили, детская повесть, Джеральд Даррелл, векторный анализ

### Для цитирования:

Kiose M.I. Linguistic creativity and discourse profiles of English language children's novels. *Russian Journal of Linguistics*. 2021. Vol. 25. № 1. P. 147–164. DOI: 10.22363/2687-0088-2021-25-1-147-164

## 1. Introduction

Contrastive discourse analysis (Neff et al. 2004, Tannen et al. 2015) has become a powerful method to explore the discourse production specificity with the studies centered around the discourse construal schemes (Gernbacher 1996, Van Dijk 1997), pragmatic categories (Aijmer & Lewis 2017), rhetorical functions (Mann & Thompson 1988, Gylling-Jørgensen 2013), and discourse markers (Taboada et al. 2012). Its results can serve educational purposes (Péry-Woodley 1990, Maxwell-Reid 2013), intercultural communication skills formation (Gordon & Chen 1991, Neff et al. 2004), social and cultural adaptation of discourse types (Silva & Grillo 2019). However, while the analysis of construal schemes, pragmatic and cultural functions leads to differentiating the discourse types, these methods fail to differentiate the discourse subtypes, as the features under observation may not follow any regular patterns or may occur occasionally. To develop the procedure of discourse subtypes differentiation, we need to explore the discourse markers of the category, which is common for the discourse in general, but which is subject to the slightest variations of the discourse construal. In this case the discourse markers distribution will reflect the sub-discourse construal differences.

One of the discourse categories that might suffice to explore the sub-discourse types is the category of linguistic creativity, which is expressed in the forms of language poesis in fiction and non-fiction texts. Among the many theories of linguistic creativity in discourse production, cf. among others (Sligh et al. 2005, Zawada 2006, Demjankov 2008, Iriskhanova 2009, Agres et al. 2015, Steels 2016, Swann & Deumert 2018, Simpson 2019, Sokolova 2019, Feschenko 2020), we addressed the ones which specify the linguistic means of creativity (Swann et al. 2011, Jones 2012, Carter 2016, Gridina 2018). In general, it is possible to trace two basic approaches to linguistic creativity means analysis, that is eliciting elementary particles (discourse markers, mostly stylistic means), and eliciting various shifts and modifications (code shifts and transfer). The first approach can be illustrated by the work of R. Carter; for example in (Carter 2016) the researcher compiles a list of discourse markers of linguistic creativity for written texts mostly expressed at the lexical level. The second approach is developed by T. Gridina in (Gridina 2018), who explores different modification schemes which can be traced in a definite discourse type. To detect the minor differences in differentiating sub-discourse types, we consider both approaches to compile a list of potentially significant parameters expressed in discourse markers and shifts which might display variance in the sub-discourses contrasted.

To proceed with the contrastive study, we develop the method of vector analysis in application to discourse profiles construal. The notion of discourse profiles suggested in construction grammar and structure building frameworks (Ariel 2004), developmental studies (Sherratt 2007, Singer 2013) and since recently in multimodal studies (Iriskhanova & Cienki 2018) is developed here to assess the relative activity of discourse parameters in the process of discourse production. Discourse and sub-discourse profiles are treated as the vector models of discourse

displaying the parametric activity distribution. Vector analysis which is an instrument in vector semantics (Mikolov et al. 2013, Paradis et al. 2015, Perek 2016), can be applied to discourse profile construal as it contributes to deducing the dependences of parametric activity in their correlation schemes and the degree of prototypicality of the discourses contrasted. The cosine of the angles between the two or several vectors representing the parametric activity values will indicate the similarity degree of the discourse profiles under observation. Higher correlation values will point out the discourse profiles with similar parametric activity distribution. Thus, detecting the general or common schemes of parametric activity distribution helps to list the more prototypical texts, i.e. the texts most typical of the discourse observed. The texts which display higher variance of parameter activity and lower correlation values with common parameter distribution schemes will be more marginal or less prototypic. Finally, we can range the discourses on the prototypicality scale in accordance with the correlation values their discourse profiles demonstrate. It means that vector analysis of discourse profiles is also an instrument to differentiate various discourses and sub-discourses and to give statistical evidence of their proximity.

The necessary requirements for such analysis are that the discourse parameters must be common for the discourses or sub-discourses contrasted and that to construe the model, the parameter activity ratio must be considered. In this study, the linguistic creativity parameters will serve to construe the discourse profiles through vector analysis. Discourse linguistic creativity vector is the linguistic parameters activity distribution within the discourse which construes the discourse profile. To test the method, we select the discourse subtypes demonstrating more and less prototypical features. The research data will be the samples of children's classical adventure novel of the 1950s in the English language literature included in the shortlists of best children's literature winning the Newbery awards. The choice of the book titles was controlled by the research task of exploring their discourse prototypicality, that is why their general discourse features had to display homogeneous nature. As an example of a potentially less prototypical book title we selected the world-famous Gerald Durrell's novel "My family and other animals" written at the same period which is known to demonstrate specificity in terms of its autobiographic character. We will find out whether this specificity makes it less prototypical than the more typical adventure stories of this period also written by men-writers, also featuring nature descriptions, also presenting the main character of a boy confronting some life problems in the natural surroundings. We will address these types of discourse as adventure novel sub-discourses and define the degree of their prototypicality in the contrastive analysis. To proceed, we will apply the method of linguistic creativity parameters elicitation and annotation in discourse developed in (Zykova & Kiose 2020). Then, sub-discourse parametric activity distribution will be assessed, and sub-discourse profiles construed. Their correlation patterns will serve to deduce the prototypicality of sub-discourse types.

## 2. Research methods and procedure

To conduct the discourse contrastive analysis with the aim of disclosing more and less prototypical sub-discourse profiles, we had to use the data which possesses the uniformity in genre, style, authorship, and issue period contrasting it with the data demonstrating some distinct variance, so that the two data samples will nevertheless present two variants of one discourse type, or its two sub-discourse types. The discourse type addressed was the English language adventure novels with its two sub-discourses, fiction narrative type which is non-autobiographic, and autobiographical type. To assess the prototypical discourse profile, we had to address several representatives of the sub-discourse type of fiction narrative and further on contrast the prototypicality results with the hypothetically less prototypical autobiographical narrative. The research data is the children's adventure novels selected from the list of Newbery award book titles for young adults.

To assess the prototypical discourse profiles, we had to select similar narration formats to avoid diversity. Thus, we addressed the novels issued at the same period (the second half of the 1950s) to avoid the diversity in discourse creation time. Next, we selected only natural surroundings as the basic scene and event frame to avoid the diversity in description formats. We also ensured that it was similarly the first-person narration of the main character (a boy of 12–14) and that the book was written by a male writer native speaker of English, and that the contrasting fragments might be of the same length. With these requirements, we selected three adventure novels “My Side of the Mountain” by Jean Craighead George (published in 1959), “Old Yeller” by Fred Gibson (published in 1956), and “Banner in the Sky” by James Ramsey Ullman (published in 1955). Out of these three novels we selected three fragments (chapters) of approximately 20,000 signs each (18,389, 20,062, 16,733, correspondingly). Parametric manual coding (annotation) procedure was then carried out and the prototypical discourse profile was construed.

To assess the prototypicality variation, we selected the autobiographical adventure novel “My Family and Other Animals” written by Gerald Durrell in 1956. It is also a boy's first-person narration about his adventures mostly in the natural surroundings, but the narrator here is the author himself. The fragment under consideration is 21,677 signs long (one chapter), equal in size with each of the three previously given fragments.

The study follows the research scheme comprising three steps: the linguistic creativity parameter system adaptation, manual coding, and statistical analysis procedure. We finally deduce the linguistic creativity vector values, which help construe the discourse profiles and contrast them. At the first step we compile the list of linguistic creativity parameters with the linguistic markers outlined, among others, in (Jones 2012, Carter 2016, Bergs 2019, Musolff 2019), which might serve to differentiate the sub-discourse types. In (Zykova & Kiose 2020) the list of 52 linguistic creativity parameters was developed to differentiate the Russian

language cinematic and fiction literature discourse. The list of parameters involved the ones which demonstrate stylistic specificity and code-switching of various types with these features being significant for discourse differentiation in terms of linguistic creativity potential. Due to multimodal character of discourse representation (Ponton 2016), the parametric activity was assessed at all the levels representing linguistic creativity in the discourse written form including phonological, morphological, word-formation, lexical, syntactic, and graphic parameters (Carter 2016). The results have sufficed to detect the differences in level activity and the activity of single parameters and their groups. At the same time, several similar features were also disclosed. In (Kiose 2020) this list of parameters was tested on Russian language fiction literature sub-discourse types, and vector instrument for discourse profiles assessment was introduced. Here we will specify the parameters for the English language discourse analysis as well as test the method of discourse profiling and vector analysis to deduce more and less prototypical discourse profiles. To do this, we apply the parameter adaptation procedure for the sub-discourse contrastive analysis in the English language.

The original parameter list (developed for the Russian language linguistic creativity assessment) involved the parameters of phonological, morphological, word-formation, lexical, syntactic, and graphic levels of linguistic creativity representation in application to the written discourse form under observation. The list of phonological parameters was adopted for the English language data with no changes made. In terms of phonological parameters, we assessed phonetic ellipsis, onomatopoeia (sound imitation), phonetic accent, sound repetition, phonetic occasionalism, metric, rhythm, rhyme, and phonetic error. The samples (1-3) show the examples of phonological parameters of linguistic creativity:

(1) *He would raise one hand in lazy salute, and across the still, blue water you could hear the plaintive **squeak** of the oar as it twisted, and the soft **clop** as it dug into the sea* (G. Durrell).

(2) *When his enemy was backed into the corner, George would dodge and twist round him with the agility of a wasp, **stabbing, thrusting, guarding*** (G. Durrell).

(3) *He **snorted, started** to run, then just **stood** there and **stared** at me* (F. Gibson).

In sample (1) *squeak* and *clop* employ the sound imitation techniques. Sample (2) is the illustration of a rhythm pattern in *stabbing, thrusting, guarding, it demonstrates* three subsequently used two-syllable words with the first-syllable stress pattern. In sample (3) we see the sound repetition in *snorted, started, stood, stared*. The list of graphic parameters has not undergone any changes, either. The list involved modifications of letters and with punctuation marks, modification with word spelling, graphic accent, and graphic errors. In (4–5) we give examples of the graphic parameters use:

(4) “He should be getting some sort of grounding in things like mathematics and **French... and his spelling’s appalling**” (G. Durrell).

(5) *The next day I told Bill **good-by**, and as I strode, warm and fed, onto the road, he called to me, “I’ll see you tonight”* (J. George).

In sample (4) elision marks stand for pauses in speech. Sample (5) contains an example of a spelling error in *good-by* instead of *good-bye* used to mark colloquial speech.

The lists of parameters for other language levels had to be slightly adapted to the English language typological structure and the discourse specificity. The list of morphological parameters which comprised the parameters of morphological categorial shifts, such as different part of speech shifts, was left the same although we had to consider the role of conversion in English and the necessity to preliminarily detect the direction of conversion shift in each example. We also had to consider grammaticalization and lexicalization when we assessed the morphological categorial shifts which made the number of categorial shifts in the English language examples much higher than in the Russian language data. In the list of word-formation parameters we had to neglect the opposition of novel and entrenched word building patterns, since compounding (being the main word building pattern) is much more typical in English than in Russian and its novelty degree is hard to be detected. The list of morphological parameters involved affixation demonstrating evaluation semantics, compounding, conversion, contraction and shortenings, word building error, and paradigmatic attraction. Samples (6–9) demonstrate some of the morphological and word-formation parameters use.

(6) *Rudi’s eyes moved over the **ice-walls**: almost vertical, smooth as glass. “Have you an ax?” he asked* (J.R. Ullman).

(7) “*Well, if you insist on stuffing him full of **useless** information, I suppose George would have a shot at teaching him,*” said Larry (G. Durrell).

(8) *The water keeps the leaf wet, and although the top dries up and **burns** down to the water level, that’s as far as the **burning** goes* (J. George)

(9) “**Lucky** I fell on the other side,” he grinned (J.R. Ullman).

In sample (6) there is a compound word *ice-walls*, in sample (7) – the negative affix-less in *useless*; in sample (8) we notice the use of paradigmatic attraction in *burns* and *burning*; in sample (9) there is a word building mistake in *lucky* where the adjective is used instead of an adverb in oral speech. In the lists of lexical and syntactic parameters we had to introduce several slight changes. In terms of lexical parameters, we annotated the register shifts, dialectal shifts, professional language shifts (use of terms), language shifts, semiotic system shifts (here – the use of numeric values and animal language), emotives, neologisms, lexical paradigmatic means (synonyms, antonyms, homonyms, etc.), polysemy, tropes, proper names, lexical errors, phraseology, phraseological errors, and allusion. We also followed the need to specify the register shifts to higher and lower registers which, as we previously indicated, might serve to demarcate the sub-discourse types in the

discourse under observation. Samples (10–11) show the use of higher and lower register shifts in the novels:

(10) *Then came the final moment, the upward and outward flick that would catch his opponent's weapon and twist it harmlessly to one side, the swift withdrawal, followed by the long, straight lunge that drove the point of his foil right through the **adversary's** heart* (G. Durrell).

(11) *"Well, Larry's so **bloody** unreasonable."* (G. Durrell).

In sample (10) the use of a formal word *adversary* instead of a more neutral *enemy* or *opponent* is marked as a higher register shift, whereas in (11) the use of *bloody* in *bloody unreasonable* is marked as a lower register shift. In the list of syntactic parameters, we had to specify the inversion parameter and marked as inversion any noun phrase or verb phrase group patterns with reverse word order. In samples (12–15) we give some examples of syntactic parameters and their discourse markers.

(12) ***Too many leaves had fallen and turned to loam, too many plants had grown up and died down** over the old home site* (J. George).

(13) *"Oh, no, Miss Turner, I want **to go live there**"* (J. George).

(14) *Scarcely **had we settled** into the strawberry-pink villa before Mother decided that I was running wild, and that it was necessary for me to have some sort of education* (G. Durrell).

(15) *"Nothing broken-no," said the voice. "Just shaken up some. And cold."* (J.R. Ullman)

In (12) we notice the use of the anaphor in *Too many leaves had fallen and turned to loam* and *too many plants had grown up and died down*. In (13) there is a syntactic error in the use of *I want to go live there* instead of *I want to go and live there*. In sample (14) there is the subject and auxiliary verb inversion in the emphatic construction *Scarcely had we settled*. Sample (15) provides several examples of elliptical sentences.

At the second step we use parameter codes in order to manually annotate the fragments selected for their further statistical study. To annotate the samples, we developed the following scheme which later underwent automatic statistical analysis in the program HETEROSTAT (Kiose & Efremov 2020) which assessed the parameters activity and calculated the correlation indices for the parameters. To demonstrate the annotation procedure, we will give an example of the annotation on a small fragment of F. Gibson's novel:

#269 I remember like yesterday how he strayed in <413> out of nowhere <302, 413, 410> to our log cabin on Birdsong Creek <302, 411, 601>. #270 He made me <413> so mad <406, 510> at first <413> that I wanted to kill him <406>.

Here the #numbers correspond to the sentence number in the collected textual data. The in-sentence numbers stand for the parameter codes, where the parameters of the same level start with the same number, where 300-type parameters stand for word building parameters, 400-type parameters are lexical parameters, 500-type are



syntactic parameters, and 600-type are the graphic ones. Here the parameter codes are the following: 302 is compounding, 406 is emotive, 410 is stylistic trope, 411 is proper name use, 413 is phraseology, 510 is syntactic intensifier.

At the third step, the coded texts were statistically processed for single and group parameter activity and for parameters correlation. The parameter activity ratio was calculated which allowed us to regard parameter values as the vector coordinates and conduct the correlation analysis between the parameter values of the four texts. We expected to find lower correlation indices of G. Durrell sub-discourse parameter values with the hypothetically more prototypical three text fragments parameter values, which in their turn were supposed to demonstrate higher correlation indices between the three sub-discourses. At the same time, to visualize the parametric activity results, we construed the sub-discourse profiles on the basis of the vector models of the sub-discourses. The models show the differences between the sub-discourse profiles as well as their similarities. In section 3 we will proceed to the results of the statistical parametrization analysis, prototypicality analysis through multiple correlation, and variance analysis.

### 3. Research Results

First, we will address the general statistical data on the four text fragments of the hypothetically more prototypical adventure novels by Jean Craighead George, Fred Gibson, and James Ramsey Ullman, contrasting them with the data of the hypothetically less prototypical sub-discourse of Gerald Durrell’s autobiographical novel. The general statistical data involves the number of words, sentences and general parameter activity results which are given in Table 1.

Table 1

General statistics on the four discourse fragments

Discourse fragments	Number of signs	Number of words	Number of sentences	General parameter activity
Jean Craighead George’s “My side of the mountain”	18,389	3,571	262	1389
Fred Gibson’s “Old Yeller”	20,062	4,014	282	1896
James Ramsey Ullman’s “Banner in the Sky”	16,733	3,191	390	1169
Gerald Durrell’s “My Family and other animals”	21,677	3,789	173	1665

In Table 2 we give the data on the four sub-discourse types and the parameter activity in absolute values and the ratio values (in relation to the general parameter activity).

The relative values of parameters demonstrating the linguistic creativity vectors underwent the correlation analysis to test the degree of contingency between the four samples parameter activity. The T-test analysis was carried out in JAMOVI program. We preliminary assessed the data degree of normality (Shapiro-Wilk test

of normality), which showed there was no violation of the assumption of normality (with p-values ranging from .02 to .035). The one-sample T-test demonstrated the following results for four samples: for Jean Craighead George’s “My side of the mountain” Student’s t was 3.01 (p = .03), for Fred Gibson’s “Old Yeller” Student’s t was 2.5 (p=.05), for James Ramsey Ullman’s “Banner in the Sky” Student’s t was 2.83 (p = .03), for Gerald Durrell’s “My Family and other animals” Student’s t was 2.27 (p = .07). Thus, the autobiographic sample of G. Durrell also demonstrated higher variance in parameter activity distribution as opposed to the four non-autobiographic samples.

Table 2

Linguistic creativity parameter values

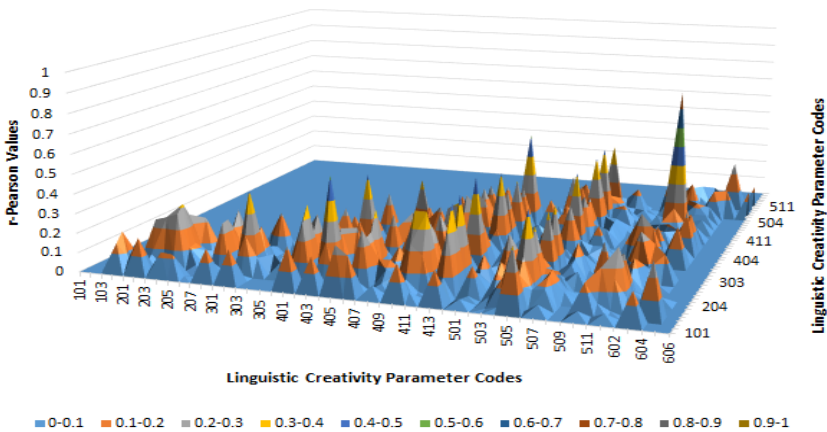
Parameter language levels	Phonological		Morphological		Word-formation		Lexical		Syntactic		Graphic	
	Absolute	Ratio	Absolute	Ratio	Absolute	Ratio	Absolute	Ratio	Absolute	Ratio	Absolute	Ratio
Jean Craighead George’s “My side of the mountain”	25	.018	381	.274	94	.068	516	.371	346	.249	27	.069
Fred Gibson’s “Old Yeller”	32	.017	622	.328	107	.056	778	.41	257	.136	100	.053
James Ramsey Ullman’s “Banner in the Sky”	23	.02	343	.293	66	.056	361	.309	342	.293	34	.03
Gerald Durrell’s “My Family and other animals”	51	.031	301	.181	129	.077	828	.497	338	.203	18	.011

Then we proceeded to testing the parameter values at different language levels to detect at which language levels the linguistic creativity parameters demonstrated higher variance. The one-sample T-test disclosed that the highest variance was observed at graphic (Student’s t is 3.2 at p = .04), syntactic (Student’s t is 6.56 at p = .007), phonological (Student’s t is 6.66 at p = .007), and morphological (Student’s t is 8.57 at p = .003) levels; whereas for word-formation and lexical levels the variance was lower (Student’s t is 12.59 at p = .001 and Student’s t is 10.08 at p = .002, correspondingly). We can therefore conclude that linguistic creativity parameter values distribution is contingent on the sub-discourse type, which in our case was clearly demonstrated by means of graphic, syntactic, phonological, and morphological parameter variance. However, we need to point out that with the similar parameter activity at word-formation and lexical levels the number of sentences in autobiographic discourse is twice smaller.

Now we will turn to parameter activity correlation values in the four samples. In HETEROSTAT program we calculated the binary correlation for all the parameters in the samples. In Fig. 1 we show the diagrams of parameter correlation values in non-autobiographic sub-discourse and autobiographic sub-discourse.

The results obtained prove that group parameter activity in autobiographic discourse type is much higher, which results from higher density of linguistic creativity markers used in twice smaller sentence number. We see higher correlation values in terms of all language levels, but the most significant (with the significant r-Pearson values exceeding .25 with the  $k = 102$  and  $p < 0.01$ ) were the following: parallel structures and professional language (terms) (.67), parallel structures and stylistic tropes (.63), as well as professional language and stylistic tropes on their own (.61), embedded structures and punctuation novelty (.71), compounds and categorial modifications of nouns (.58), compounds and professional language (.64), idioms and grammaticalization of verbs (.6), evaluative affixation and emotives (.52). Among the significant correlation pairs for non-autobiographic discourse type, we can name capitalization and personal names (.78), professional language and categorial modifications of nouns (.52), idioms and stylistic tropes (0.52). All in all, the results suffice to conclude that G. Durrell’s sample is far more sophisticated in terms of linguistic creativity use with the density of linguistic creativity markers and their specific choice.

**Group Linguistic Creativity Parameter Activity in Non-Autobiographic Discourse**



**Group Linguistic Creativity Parameter Activity in Autobiographic Discourse**

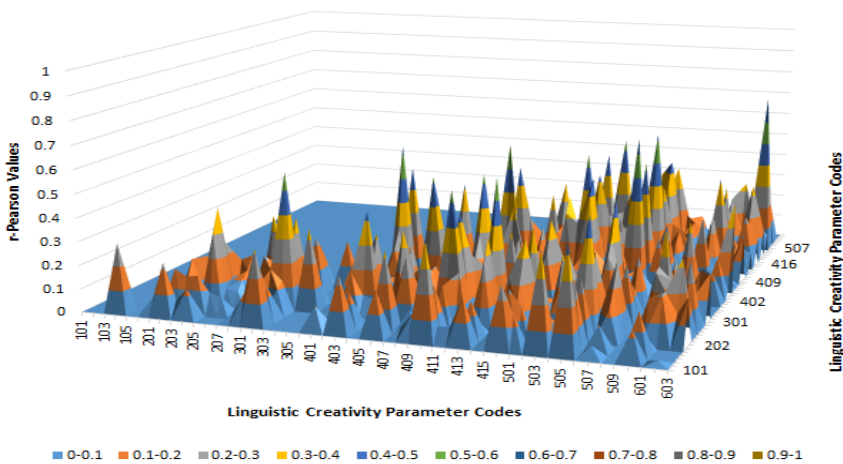
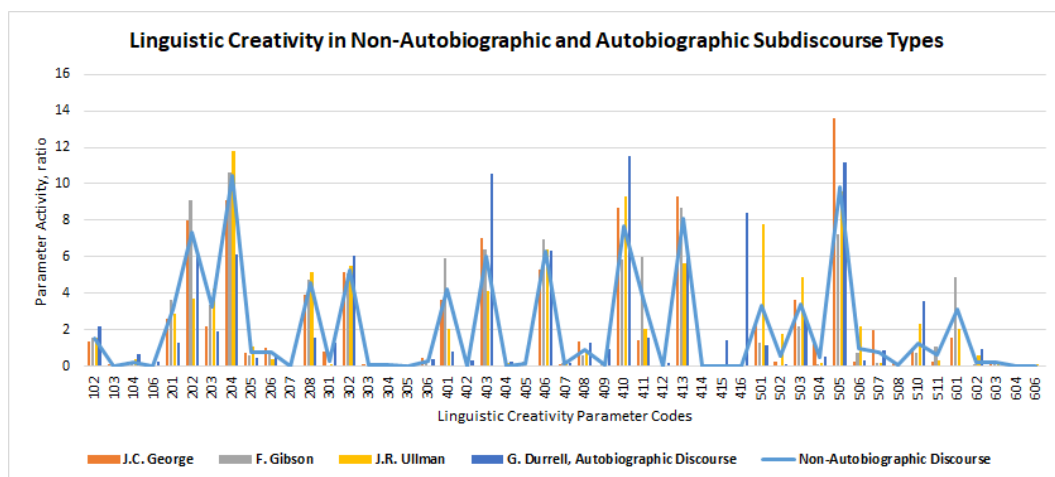


Figure 1. Group (Paired) linguistic creativity parameter activity in contrasting sub-discourse types

We now proceed to eliciting the parameters of the two discourse types which demonstrate significant variance in their activity. To minimize the effect of the total parameter activity within the samples, we will consider the relative parameter values in each sample. We will use the first three samples representing the same sub-discourse type (non-autobiographic) to deduce the mean values of linguistic creativity parameters for the non-autobiographic discourse type, whereas G. Durrell’s sample (autobiographic) parameter activity values will serve as a contrasting type of values. In Fig. 2 we present the diagram of linguistic creativity parameters relative values in 4 samples, three non-autobiographic samples and their mean values in non-autobiographic discourse in total, and G. Durrell autobiographic sample.



**Figure 2. Relative values of linguistic creativity parameters in the contrasting sub-discourse types**

The diagram shows that some parameters in autobiographic sub-discourse can be singled out, as their relative values in some cases exceed and in some are significantly lower than the mean values in non-autobiographic sub-discourse. The parameters with the values exceeding the mean values are professional language (code 403 – 10.57 vs 5.99), stylistic tropes (code 410 – 11.53 vs 7.66), formal register shifts (code 416 – 8.4 vs 0), parallel structures (code 505 – 11.7 vs 9.8) and interjections (code 510 – 3.54 vs 1.24). Among the parameters demonstrating lower values are adjective categorial shifts (code 201 – 1.26 vs 3.12), pronoun categorial shifts (code 203 – 1.92 vs 3.23), grammaticalization of the verb (code 204 – 6.13 vs 10.44), categorial shifts of auxiliaries (code 208 – 1.56 vs 4.58), colloquial register shifts (code 401 – 0.78 vs 4.2), use of personal names (code 411 – 1.56 vs 3.53), ellipsis (code 501 – 1.14 vs 3.3), appeal (code 506 – 0.3 vs 0.99), capitalization (code 601 – 0.06 vs 3.12).

#### 4. Discussion

The general statistics on the four discourse fragments given in Table 1 show that with the same number of signs the four fragments vary in the number of words used, which means that this indicator does not suffice to differentiate the

sub-discourse types. At the same time, the number of sentences has a more demonstrative character, with the lowest number of sentences in the novel by G. Durrell, which means that the sentence length in the novel is 1.8 times higher than in the adventure novels of a non-autobiographic type in general. General parameter activity does not demonstrate any specificity, but if we consider the fact that the number of sentences is almost two times lower, it means that the linguistic creativity potential of each sentence is almost two times higher, which cannot be neglected. The sentence parametric activity indices are 6.72, 5.3, 3.0, and 9.62, correspondingly. Here, the average sentence parametric activity for non-autobiographic narration is 5.0, whereas for autobiographic narration it is almost two times higher. It may comply with average parameter activity being the same within the same number of signs in the samples, but if we consider that, for instance, one sentence will contain not one but two metaphors, not one but two allusions, code-shifts, or sound imitations, it becomes obvious that it makes a sentence more “condensed” in terms of the information conveyed, thus requiring more cognitive effort on the part of the reader.

In terms of linguistic creativity parameter values in Table 2, in the phonological, morphological, word-formation, lexical, and graphic levels, the parameter activity of the autobiographical novel demonstrates specificity in some cases being significantly higher and in some lower than the parameter values within the range of non-autobiographical novel samples. If we contrast the average parameter ratio of the three non-autobiographical novels and compare it with the parameter ratio of the autobiographical one, we will come to the following results: in the autobiographical novel the activity of phonological parameters is 1.69 times higher, the activity of word-formation parameters is 1.28 times higher, the activity of lexical parameters is 1.37 times higher; whereas for morphological parameters it is 1.65 times lower, and for the graphic parameters it is 4.6 times lower. Thus, the parameter activity distribution demonstrates specificity being sub-discourse dependent. It also suffices to remind that higher phonological, word-formation and lexical parameter activity is observed in the two-time smaller sentences number, which makes it a highly significant result.

To illustrate the distribution of single parameters activity, we address the diagrams of the discourse profiles of four samples (see Fig. 3).

The discourse profiles demonstrate the redistribution of parameter activity in the sub-discourse samples contrasted. While the three samples of non-autobiographic discourse have higher values of morphological linguistic creativity parameters (200-codes), the autobiographic discourse does not demonstrate it. The latter also shows prevalence of professional language, ambiguity, and trope language, which may serve to single out this discourse type (several 400-codes).

The single parameter activity values obtained may be viewed as the vector coordinates of linguistic creativity vectors in the contrasting discourse types. Each vector is then represented by the relative values of the linguistic creativity parameter frequency within the samples under observation. Unfortunately, as the multiple vector coordinates are also subject to linguistic typological nature (and not

only to the discourse nature), the sub-discourse vectors demonstrate high correlation patterns. For this reason, correlation analysis applying all parameters' values does not suffice to distinguish between the discourse subtypes, although as we have shown, it is the linguistic creativity level activity and not the single parameters which may serve to detect significant variance in their distribution. The results of one-sample T-test have sufficed to conclude that the autobiographic sample by G. Durrell demonstrated higher variance in parameter level activity distribution, as opposed to the four non-autobiographic samples. At the same time, the results of r-Pearson parameter values correlation show that this sample is also far more sophisticated in terms of linguistic creativity use with the density of linguistic creativity markers and their specific choice.

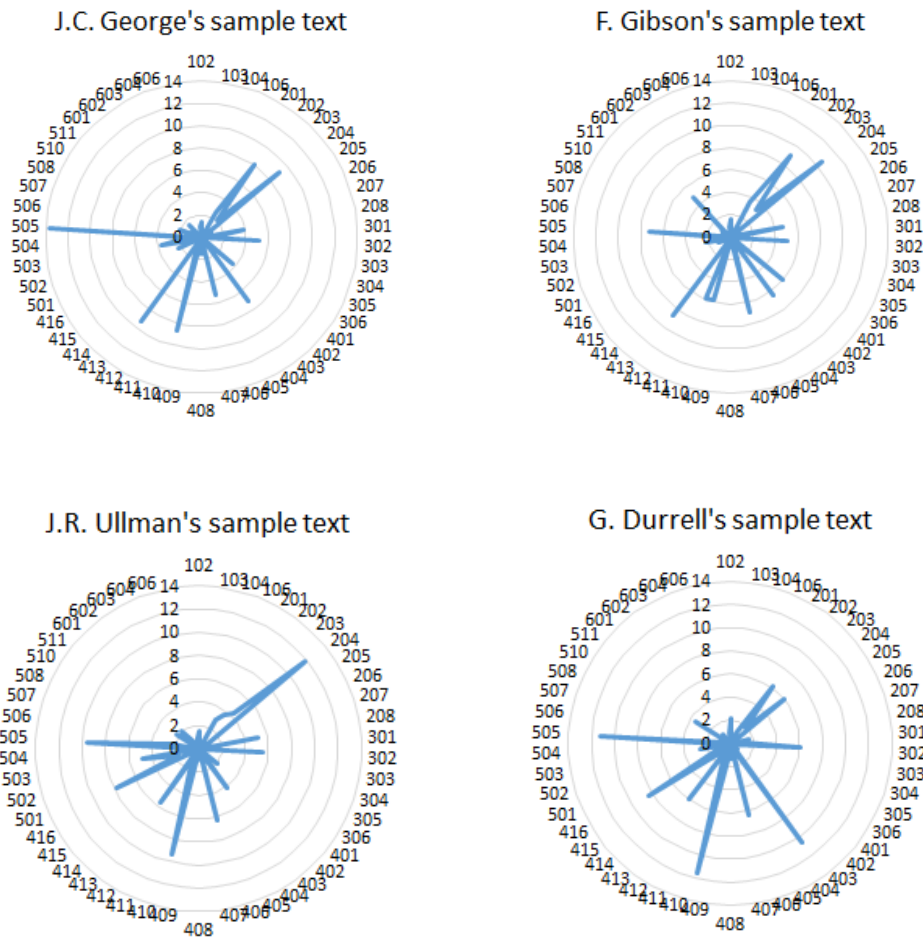


Figure 3. Discourse profiles of the four sample texts (J.C. George's, F. Gibson's, J.R. Ullman's non-autobiographic type, and G. Durrell's autobiographic type)

### 5. Conclusion

The study has proved the efficiency of vector methodology (Mikolov et al. 2013, Paradis et al. 2015, Perek 2016) in application to contrastive sub-discourse

analysis and the study of linguistic creativity as a discourse (Carter 2016). To proceed with the study, we used parametrization methods to identify the frequency and correlation patterns of the linguistic creativity parameters in contrasting discourse types, that is the non-autobiographic and autobiographic discourse subtypes of adventure novels. General and analytical statistical methods served to detect the parameter activity on all the linguistic creativity representation levels (in the discourse written form) and their correlation patterns. With the data obtained, we have arrived at the relevant conclusions about the prevalence of several single parameter activity as well as paired parameter activity. The study has shown that G. Durrell's autobiographic discourse exhibits higher values of professional language, stylistic tropes, formal register shifts, parallel structures and interjections; in terms of correlation patterns these are parallel structures and professional language, parallel structures and stylistic tropes, professional language and stylistic tropes, embedded structures and punctuation novelty, compounds and categorial modifications of nouns, compounds and professional language, idioms and grammaticalization of verbs, evaluative affixation and emotives.

Applying the relative parameter values of the discourse samples, we have developed the notion of discourse profiles suggested earlier in construction grammar (Ariel 2004), developmental and multimodal studies (Singer 2013, Iriskhanova & Cienki 2018). In our study, discourse profiles are viewed as multi-dimensional constructs visualizing the relative discourse activity of parameter values representing the category of linguistic creativity. The discourse profiles illustrating the results of vector analysis help analyze the variability of linguistic creativity parameter values distribution across the discourse types. The contrasted profiles of non-autobiographic and autobiographic discourse types help infer the relative activity of linguistic creativity representation levels and single parameter activity in relation to the discourse types. These results have sufficed to determine the variance in the distribution patterns of parameter activity in the contrasting discourse types, with autobiographic discourse showing weaker correlation with the distribution values of non-autobiographic discourse. Thus, vector analysis can serve to distinguish between discourse subtypes, at least in terms of the category of linguistic creativity.

To sum up, the method of discourse profiling has proved its efficiency in differentiating autobiographic and non-autobiographic discourse types, which means that it can be applied for contrastive sub-discourse analysis, probably to explore other discourse categories (rather than the category of linguistic creativity) with the aim of revealing the discourse specificity.

© Maria I. Kiose, 2021

 creative commons



This work is licensed under a Creative Commons Attribution 4.0 International License  
<https://creativecommons.org/licenses/by/4.0/>

## Acknowledgements

The research is sponsored by the Russian Science Foundation, grant № 19-18-00040 and is carried out in the Institute of Linguistics of the Russian Academy of Sciences.

## REFERENCES

- Agres, Kathleen, Stephen McGregor, Matthew Purver & Geraint Wiggins. 2015. Conceptualizing creativity: from distributional semantics to conceptual spaces. *Proceedings of the 6th International Conference on Computational Creativity*, 118–125. Aijmer, Karin & Diana Lewis (eds.). 2017. *Contrastive Analysis of Discourse-pragmatic Aspects of Linguistic Genres*. Berlin: Springer.
- Ariel, Mira. 2004. Accessibility marking: discourse functions, discourse profiles, and processing cues. *Accessibility in Text and Discourse Processing: A Special Issue of Discourse Processing* 37 (2). 91–116. DOI: 10.1207/s15326950dp3702\_2
- Bergs, Alexander. 2019. What, if anything, is linguistic creativity? *Gestalt Theory* 41 (2). 173–183. DOI: 10.2478/gth-2019-0017
- Carter, Ronald. 2016. *Language and Creativity. The Art of Common Talk*. 2nd edn. New York: Routledge.
- Demjankov, Valeriy Z. 2008. Yazykovoe tvorchestvo i rechevaya kreativnost' [Language art and speech creativity]. *Yazyk kak mediator mezhdz znanie i iskusstvom [Language as Mediator Between Knowledge and Art]*. 11–19. Moscow: Azbukovnik.
- Feschenco, Vladimir V. 2020. Ot lingvoestetiki k lingvoevristike: slovtvorchestvo v hudozhestvennom i nauchnom diskursah [From linguistic aesthetics to linguistic heuristics: verbal creativity in literary and academic discourses]. *Critique and Semiotics* 1. 92–113.
- Gernsbacher, Morton A. 1996. The structure-building framework: what it is, what it might also be, and why. In Bybruce K. Britton & Arthur C. Graesser (eds.), *Models of Understanding Text*, 289–311. New Jersey: Lawrence Erlbaum Associates.
- Gordon, Taylor & Tingguang Chen. 1991. Linguistic, cultural, and subcultural issues in contrastive discourse analysis: Anglo-American and Chinese scientific texts. *Applied Linguistics* 12 (3). 319–336.
- Gridina, Tatiana A. 2018. Smyslovaya perspektiva slova v igrovom hudozhestvennom tekste [Sense perspective of a word in fiction texts]. In Tatiana A. Gridina (eds.), *Lingvistika Kreativa [Linguistics of Creative-4]*. 270–281. Ekaterinburg: Ural State Pedagogical University.
- Gylling-Jørgensen, Morten. 2013. *The Structure of Discourse: A Corpus-Based Cross-Linguistic Study*, PhD Series. Frederiksberg: Copenhagen Business School (CBS).
- Iriskhanova, Olga K. 2009. O ponyatii kreativnosti i ego roli v metazykyke lingvisticheskikh opisanij [On the notion of creativity and its role in the metalanguage of linguistic description]. *Kognitivnye issledovaniya yazyka [Cognitive Studies of Language]* 5. 157–171.
- Iriskhanova, Olga K. & Alan Cienki. 2018. The semiotics of gestures in cognitive linguistics: contribution and challenges. *Issues of Cognitive Linguistics* 4. 25–36. DOI: 10.20916/1812-3228-2018-4-25-36
- Jones, Rodney. 2012. *Discourse and Creativity*. London: Routledge.
- Kiose, Maria I. 2020. Vektornye modeli lingvokreativnosti kak instrument ocenki variativnosti diskursa [Vector models of linguistic creativity and their applicability in discourse diversity assessment]. In Irina V. Zykova & Victoria V. Krasnykh (eds.), *Yazyk, Kultura,*



- Tvorchestvo: Mirovye Praktiki Izucheniya* [Language, Culture and Creativity: World Practices in Studies], 113–129. Moscow: Gnozis.
- Kiose, Maria I. & Andrej A. Efremov. HETEROSTAT. Program for discourse parameters variance calculation. Registered 21.09.2020. Registered Number 2020661240.
- Mann, William C. & Sandra A. Thompson. 1988. Rhetorical structure theory: toward a functional theory of text organization. *Text* 8 (3). 243–281.
- Maxwell-Reid, Corinne. 2013. The challenges of contrastive discourse analysis: reflecting on a study into the influence of English on students' written Spanish on a bilingual education program in Spain. *Written Communication* 28 (4). 417–435. DOI: 10.1177/0741088311421890
- Mikolov, Tomas, Kai Chen, Greg Corrado & Jeffrey Dean. 2013. Efficient estimation of word representations in vector space. *Proceedings of ICLR Workshop*, 430–443.
- Musolff, Andreas. 2019. Creativity in metaphor interpretation. *Russian Journal of Linguistics* 23 (1). 23–39. DOI: 10.22363/2312-9182-2019-23-1-23-39
- Neff, Joanne, Emma Dafouz, Mercedes Díez Prados, Rosa Prieto, Craig Chaudron. 2004. Contrastive discourse analysis: argumentative text in English and Spanish. In Carol L. Moder & Aida Martinovic-Zic (eds.), *Discourse Across Languages and Cultures*, 267–283. Amsterdam: John Benjamins.
- Paradis, Carita, Simone Löhndorf, Joost van de Weijer & Caroline Willners. 2015. Semantic profiles of antonymic adjectives in discourse. *Linguistics* 53 (1). 153–191. DOI: 10.1515/ling-2014-0035
- Perek, Florent. 2016. Using distributional semantics to study syntactic productivity in diachrony. A case study. *Linguistics* 54 (1). 149–188. DOI: 10.1515/ling-2015-0043
- Péry-Woodley, Marie-Paule. 1990. Contrasting discourses: contrastive analysis and a discourse approach to writing. *Language Teaching* 23 (3). 143–151.
- Ponton, Douglas M. 2016. Movements and meanings: towards an integrated approach to political discourse analysis. *Russian Journal of Linguistics* 20 (4), 122–139.
- Sherratt, Sue. 2007. Multi-level discourse analysis: a feasible approach. *Aphasiology* 21 (3–4). 375–393. DOI: 10.1080/02687030600911435
- Silva, Beatriz A. A. & Sheila V. C. Grillo. 2019. New paths for science: a contrastive discourse analysis of modifications in popularizing science through digital media Bakhtiniana. *Revista de Estudos do Discurso* 14 (1). 51–73.
- Simpson, Paul (eds.). 2019. *Style, Rhetoric and Creativity in Language*. Amsterdam: John Benjamins.
- Singer, Murray. 2013. Profiles of discourse recognition. *Discourse Processes* 50 (6). 407–429. DOI: 10.1080/0163853X.2013.822297
- Sligh, Allison C., Frances A. Connors & Beverly Roskos-Ewoldsen. 2005. Relation of creativity to fluid and crystallized intelligence. *Journal of Creative Behavior* 39. 123–136.
- Sokolova, Olga. 2019. *Ot Avangarda k Neoavangardu: Yazyk, Sub"ektivnost', Kul'turnye Perenosy* [From Avant-garde to Neo Avant-garde. Language, subjectivity, and Cultural Transfer]. Moscow: Kulturnaya Revolutsia.
- Steels, Luc. 2016. Meaning and Creativity in Language. In Mirko D. Esposti, Eduardo G. Altmann, François-David Pachet (eds.), *Creativity and Universality in Language*, 197–208. Berlin: Springer.
- Swann, Joan, Rob Pope & Ronald Carter (eds.). 2011. *Creativity in Language and Literature. The State of the Art*. Melbourne: Red Globe Press.
- Swann, Joan & Ana Deumert. 2018. Sociolinguistics and language creativity. *Language Sciences* 65. 1–8. DOI: 10.1016/j.langsci.2017.06.002
- Taboada, Maite, Susana D. Suárez & Elsa G. Alvarez. 2012. *Contrastive Discourse Analysis: Functional and Corpus Perspectives*. Sheffield: Equinox Publishing Limited.

- Tannen, Deborah, Heidi E. Hamilton & Deborah Schiffrin (eds.). 2015. *The Handbook of Discourse Analysis* 1. Chichester, UK: John Wiley & Sons.
- Van Dijk, Teun A. 1997. *Discourse as Structure and Process. Discourse Studies: A Multi-disciplinary Introduction*. London: SAGE Publications.
- Zawada, Britta. 2006. Linguistic creativity from a cognitive perspective. *Southern African Linguistics and Applied Language Studies* 24 (2). 235–254. DOI: 10.2989/16073610609486419
- Zykova, Irina. 2020. Verbal sources of cinematic metaphors: From cinematic performativity to linguistic creativity. *Lege artis. Language yesterday, today, tomorrow* 5 (1). 499–532.
- Zykova, Irina V. & Maria I. Kiose. 2020. Parametrizaciya lingvisticheskoj kreativnosti v mezhdiskursivnom aspekte: kinodiskurs vs. diskurs detskoj literatury [Linguistic creativity parametrization in contrasting discourse types: cinematic discourse vs. discourse of children's literature]. *Voprosy Kognitivnoj Lingvistiki [Issues of Cognitive Linguistics]* 2. 26–40. DOI: 10.20916/1812-3228-2020-2-26-40

### **Article history:**

Received: 31 July 2020

Accepted: 22 January 2021

### **История статьи:**

Дата поступления в редакцию: 31 июля 2020

Дата принятия к печати: 22 января 2021

### **Bionote:**

**Maria I. KIOSE** is Doctor habil. of Philology, Associate Professor, Leading Researcher of the Centre for Socio-Cognitive Studies of Moscow State Linguistic University. Her research interests embrace cognitive semantics, indirect naming in text, referential semantics, oculo-graphic and corpus-based research of text inference and generation process.

### **Contact information:**

Moscow State Linguistic University

38 Ostozhenka, Moscow, 119034, Russia

*e-mail*: maria\_kiose@mail.ru

ORCID: 0000-0001-7215-0604

### **Сведения об авторе:**

**Мария Ивановна КИОСЕ** – доктор филологических наук, доцент, ведущий научный сотрудник Центра социокогнитивных исследований дискурса Московского государственного лингвистического университета. В сферу ее научных интересов входят: когнитивная семантика, не прямое наименование в тексте, референциальная семантика, окулографические и корпусные методы анализа восприятия и порождения текста.

### **Контактная информация:**

Московский государственный лингвистический университет

Россия, 119034, Москва, ул. Остоженка, 38

*e-mail*: maria\_kiose@mail.ru

ORCID: 0000-0001-7215-0604