

THE USE OF PRONOUNS IN RUSSIAN CLASSROOM TEXTBOOKS: A QUNTITATIVE STUDY

Marina Solnyshkina¹, Maria Kazachkova², Elena Varlamova^{3*}, Liliya Ilyasova⁴

¹Doctor in Philology, Professor, Kazan Federal University, Russia, mesoln@yandex.ru,

²Ph. D. in Philology, Moscow State Institute of International Relations, Russia, mbkazachkova@yandex.ru,

³Ph. D. in Philology, Kazan Federal University, Russia, el-var@mail.ru,

⁴Ph. D. in Philology, Kazan Federal University, Russia, lily_ilyasova@mail.ru

*Corresponding Author

Abstract

Background: For decades scholars have been conducting research on English text complexity and influence of text metrics on its difficulty for different categories of readers. Recently, considerable steps have been taken towards better understanding of variations in metrics of Russian text of different genres. The current study is a pilot corpus-based analysis focused on differences in Russian texts on science and social studies as exemplified in two classroom textbooks for the 8th grade students of Russian secondary schools: (1) Social Studies by Nikitin A.F., Nikitina T.I. (hereinafter referred to as NIK); (2) Human anatomy, physiology, hygiene by A. M. Tsuzmer, O L. Petrishina, edited by V.V. Parin (hereinafter referred to as PAR). The present study extends the work in the area by analyzing differences in the use of pronouns in academic texts on Social studies and Science and their correlation with text complexity. This is the first study based on the sub-corpora of Russian Academic discourse Corpus (RAC) compiled by researchers of Kazan Federal University (see Solovyev et al 2018). Based on the intermediate results received in this study we plan to verify them on RAC the size of which has reached 1.75 mln. tokens.

Literature Review: Personal pronouns in writing have been viewed as related to the author's voice and position which many textbooks lack, thus demonstrating that textbook writers address readers 'impersonally' and express some general views on the existing paradigm in the area. Based on the analysis of research articles from a wide variety of disciplines, Hyland (2001) comes to the conclusion that modern writers add 'the dialogic nature of persuasion in research writing' while addressing readers with personal pronouns. The idea was also mentioned by Cherry (1998) in his analysis of academic writing where the author points to the importance of writer's voice. Kuo (1999) suggests that personal pronouns allow authors to share their contributions and expect 'solidarity with readers'. Reporting on a high proportion of personal pronouns in the social sciences discourse Hyland (2001) defines it 'a valuable rhetorical strategy' used by writers to establish academic credibility.

Regardless of the research on the number and range of pronouns used in the English academic discourse, no studies, to the best of our knowledge, have been conducted on the differences of the number and range of pronouns used in Russian classroom texts, which suggests that there is a need to have more studies to compare textbooks on Science and Social studies, thus contributing to solving the problem of defining text metrics correlating with text complexity.

Research Question: The research question that guides this study is as follows:

Are there significant differences in the number of personal, reflexive and possessive pronouns used in Russian classroom books on Science and Social studies?

Research Methodology: To address the research questions, we compiled RAC sub-corpus with the total number of tokens 100270: 41383 tokens in NIK and 58887 in PAR. Then the two above mentioned books by NIK and PAR were analyzed with regard to the number and range of pronouns used. Specifically, we use the following indices: (a) the range of personal, reflexive and possessive pronouns used in both books; (b) the number of each of the pronouns under study in both books (absolute frequency); (c) normalized to 1000 frequency of pronouns (relative frequency).

Results: The results received are as follows:

(a) the range of personal, reflexive and possessive pronouns used in both books is different: NIK (Social studies) uses 26 types of pronouns while PAR (Science) uses only 17 types;

(b) the number of each of the pronouns under study is higher in NIK with three exceptions: in PAR 'ono' (it) is used in 56 cases, 'oni' (they) is used in 147 cases, there are 47 occurrences of 'nash' (our) (sing, masc), while in NIK they are 36, 86 and 11 times correspondingly.

(c) Thus, RF of 'ono' (it) in PAR is 0,95 while in NIK it is only 0, 87, RF of 'oni' (they) is higher in PAR than in NIK: 7,047 vs 6,25 correspondingly, RF of 'nash' (our) (sing, masc) in PAR is 0,79 and in NIK it is 0, 26.

The average I RF in NIK is nearly 3 times higher than the corresponding metric in PAR: 9,2 vs 3,1.

All the above mentioned demonstrates a high tendency to impersonality in the Science textbooks (PAR) under study. This indicates that the authors of the textbook on Science (PAR) favour strategies of depersonalisation: they use agentless passive and impersonal constructions, which definitely decreases texts narrativity and hamper students' comprehension of the texts.

Discussion: The findings of the research provide strong support that the existing differences in distribution of personal, reflexive and possessive pronouns in science and social studies classroom books samples are related to text complexity. These findings have also important implications for understanding subject differences and corresponding metrics. First, they confirm that the language used by writers of different subjects is different in linguistic terms. Second, the differences in the number and relative frequencies of pronouns are large and consistent to suggest that each of the subjects may have a unique linguistic profile of features.

Conclusions: Further studies using Russian Academic Corpus are suggested to address limitations in the size and range of text used for the current research.

Acknowledgements: The research presented in Parts Related Work, Methods, Analysis was supported by the subsidy of the Russian Government to support the Program of Competitive Growth of Kazan Federal University.

The research presented in Parts Discussion and Conclusion of the article was financially supported by the Russian Science Foundation, grant № 18-18-00436.

Keywords: corpus, disciplinary differences, pronouns, Russian academic discourse, text complexity

1 INTRODUCTION

The so- called "classic" readability formulas rely on few metrics: number of words per sentence and the number of syllables per word (Chall et al 1995, Duran 2007, McNamara et al 2011, 2014, Solnyshkina et al 2016, Gafiyatova et al, 2017). Recently, researchers began applying Natural Language Processing techniques as a result of which text complexity research widened the range of the inter-correlated metrics. One of the most debated indices of text complexity are pronouns which are among the clusters of text metrics categorizing narrativity and cohesion. As for academic texts they are distinguished by many different linguistic features, and pronouns can be one of the elements for distinctions.

This paper aims at further investigating possible correlation of frequency of pronouns and text complexity/readability. For this purpose we chose two classroom textbooks for 8th grade students of Russian secondary schools: (1) Social Studies by Nikitin A.F., Nikitina T.I. (NIK); (2) Human anatomy, physiology, hygiene by A. M. Tsuzmer, O L. Petrishina, edited by V.V. Parin (PAR). Both textbooks are from the "Federal List of Textbooks Recommended by the Ministry of Education and Science of the Russian Federation to Use in Secondary and High Schools". The course on Social Studies as a compulsory subject in all high schools

of the Russian Federation is accomplished with a high stake matriculation exam after two school years: the 9th and the 11th grades. The choice of these textbooks was caused by a number of reasons: (a) the fact that the texts under study were relatively free of non alphabetical symbols, graphs, figures etc., (b) the availability on the textbooks on the Internet (School textbooks and manuals, 2017).

The study was conducted to answer the following research Question: Are there significant differences in the number of personal, reflexive and possessive pronouns used in Russian classroom books on Science and Social studies?

2 RELATED WORK

The modern view on classroom textbooks implies that they are not only supposed to present theories, contrast old views and modern research, they are expected to develop students' critical thinking. Thus we may predict that authors of classroom books do not only express their opinion on a number of issues, but demonstrate their personal 'ego' though the dialogue with the reader. There are also opponents of personal views in academic and especially scientific writing who support the opinion that personal views 'blur' objective writing thus contradicting the whole idea of scientific validity and preciseness.

Many researchers of language have pointed to the role of distributional frequencies in determining the relative accessibility or ease of processing associated with a particular lexical item or sentence (Ronald et al, 2007). These approaches are known by a number of names — constraint-based, competition, expectation-driven or probabilistic models — but all have in common the assumption that language processing is closely tied to a user's experience, and that distributional frequencies of words and structures play an important (though not exclusive) role in comprehension.

Pronouns in writing are typically considered as markers of the writer's point which is though quite frequent in fiction but rare in academic writing: writers of these genre prefer to present 'impersonally' choosing more general views on in the area. The review of publications in the field allows Hyland (2001) to conclude that contemporary writers are more known for 'the dialogic nature of persuasion in research writing' than writers of the previous centuries. Cherry (1998) also argued that in English academic writing the author's point of view is valued and very much encouraged by editors and reviewers.

As for pronouns they are proved to typically mark writer's 'solidarity with readers' (Kuo, 1999). Reporting on a high proportion of personal pronouns in social sciences discourse Hyland (2001) defines it 'a valuable rhetorical strategy' used by writers to establish academic credibility. Russian scholars Pyankova (1994), Vassileva (1998) and others argue that writers of Russian scientific texts are inclined to avoid personal pronouns and resort to passive and impersonal structures.

In the Russian discourse pronouns are the most commonly used words: "Frequency Dictionary of the Russian language" by L.N. Zazorina registers 29 pronouns among the first hundred of the most frequent words in Russian: ya (I 3839¹), on(he 13143), eto (this 6940), vy (you 6547), ty (you 6475), we (my 6220), etot (that 6061), ona (she 5836), oni (they 5712), to (that 4810), vse (all 1752), svoy (one's own 4204), kotoryi (which 3968), takoy (such 3230), tot (that 2881), nash (our 2506), sebya(ourselves 2205),ego (his 2136), sam (himself, 1861), kakoy (which 1823), moi (my 1668), kto (who 2568), tam (there 1105).

3 METHODS

As it was mentioned before, personal pronouns in writing have been viewed as related to the author's voice and position which many textbooks lack, thus exemplifying the thesis that textbook writers address readers 'impersonally' and express some general views on the existing paradigm in the area.

The research is performed on the sub-corpus of 100270 tokens which is a part of Russian Academic Corpus (RAC) comprising over 40 classroom textbooks used in secondary schools of the Russian Federation (see Gabitov et al 2017). Thus, the current study is a corpus-based analysis aimed at differences in science and social (non-science) writing.

The table below shows the size of the Corpus we compiled for the study: the total number of tokens is 100,270 with 41,383 tokens in NIK and 58,887 in PAR (Table 1).

¹ The figures in parentheses are raw frequencies in the general sample of over one million tokens (in Frequency Dictionary of Russian Ed. By L.N. Zazorina. - M.: Russian language, 1977.-C. 29-804).

| Books | Tokens |
|--|---------|
| Human anatomy, physiology, hygiene (PAR) | 58887 |
| Social Studies (NIK) | 41383 |
| Total | 100 270 |

Table 1. The Size of Corpus of textbooks on Human anatomy, physiology, hygiene by V. V. Parin and Social Studies by A.F. Nikitin for the 8th Grade of Russian secondary schools

A token is viewed in the work as an instance of a sequence of characters in some particular document that are grouped together as a useful semantic unit for processing. Usually it refers to the total number of words in a text, corpus etc, regardless of how often they are repeated. A type is the class of all tokens containing the same character sequence (Tokenization, 2008).

To analyze the Corpus and determine the relevant frequency of the pronouns in the texts, we employed AntConc (AntConc, 2014) which computes a list of words appearing with the keywords under study (personal, reflexive and possessive pronouns in the current study) thus providing not only a concordance but distributional patterns of the keywords.

We also computed Absolute Frequency, i.e. all occurrences of certain pronouns in the whole corpus, and Relative Frequency, i.e. the number of occurrences per thousand tokens. E.g. The absolute frequency of possessive pronouns in NIK is 381 (see Table 2 below). The size of NIK is 41383 tokens (see Table 1 above), thus, the relative frequency of possessive pronouns in NIK is $X = (381 \times 1000) : 41383 = 9,207$.

4. ANALYSIS

On the first stage of the analysis we classified all the registered pronouns into three categories: personal, reflexive and possessive. There are more pronouns in the textbook on Social Studies (1672, NIK) than in the textbook on Science (1492, PAR) The sizes of groups are presented in Table 2. As it can be observed that personal pronouns make the biggest part in both textbooks: of 1492 total pronouns in PAR 1230 are personal, 77 are reflexive and 185 are possessive. The sizes of the corresponding groups in NIK are similar: 1173 are personal pronouns, 118 are reflexive and 381 are possessive.

The second stage implied registering absolute and calculating the relative frequencies of three groups of the pronouns under study (See Table 2).

| Pronouns | PAR AF | PAR RF | NIK AF | NIK RF |
|------------|--------|--------|--------|--------|
| Personal | 1230 | 20,887 | 1173 | 28,345 |
| Reflexive | 77 | 1,307 | 118 | 2,851 |
| Possessive | 185 | 3,142 | 381 | 9,207 |
| TOTAL | 1492 | - | 1672 | - |

Table 2. Absolute and normalized relative frequency of pronouns in the textbooks.

As it might be expected, the range of personal, reflexive and possessive pronouns used in both textbooks is different, i.e. the relative frequency of each of the pronouns is higher in NIK with rates averaging RF 28.345; 2.851; 9.207 (see Table 2 above). The latter confirms the main hypothesis that social (non-science) texts are richer in pronouns. The highest difference is observed in the group of possessive pronouns.

Personal pronouns are used in both textbooks quite frequently: in every 1000 tokens there are either 20 -21 (PAR) or 28-29 (NIK) personal pronouns. The authors of both books address readers with '**vy**' (**you**):

E.g.

*I etu problemu **vam** udalos' reshit': vy kupili poderzhannyi mototsikl. Neozhidanno vvyasnyayetsya, chto v vashem vozraste po zakonu nel'zya sovershat' stol' znachitel'nyye sdelki kupli-prodazhi i poluchat' voditel'skiye prava (NIK) – And **you** managed to solve the problem, you bought a second-hand bike. It turns out that people of your age are prohibited by law to make such deals and gain driving licenses.*

***Vy** pomnite, chto u vzroslogo cheloveka kolichestvo krovi ravno primerno 5 l. – **You** remember that an adult has about 5 liters of blood in his blood vessels.*

Another pronoun which frequency is high in both books (see Table 3 below) is **my** (**we**). It is typically an inclusive "we" used to "unite" the writer and readers into one group. Inclusive implies that the pronoun refers

both to the writer and reader (Harwood 2005).

E.g.

My smozhem ne tol'ko vyzhit', no i dostoyno zhit' v nashem neprostom mire, tol'ko nauchivshis' derzhat'sya vmeste, podderzivat', tsenit' i uvazhat' drug druga (NIK). – **We** will be able not only to survive, but live happily in our difficult world, which is possible due mutual support.

Vo vremya yedy **my** otkusyvyayem pishchu pri pomoshchi reztsov i klykov (PAR). – While eating **we** bite off pieces of food with incisor teeth and cuspids.

But there are ample examples where the writer uses 'my' (we) to refer to himself only thus realizing the so-called exclusive 'we' functioning instead of ya ('I').

E.g.

My uzhe govorili o proiskhozhdenii cheloveka (sm. § 1 dannogo uchebnika) (NIK). - **We** have already talked about the origin of man (see § 1 of this textbook),

Interestingly in PAR we came across only one example of 'exclusive' we:

E.g.

Tak kak khristalik blagodarya sposobnosti izmenyat' svoyu kriviznu mozhet uvelichivat' ili umen'shat' prelomleniye luchey v glazu, **my** ostanovimsya tol'ko na yego roli. - Since the lens, due to the ability to change its curvature, can increase or decrease the refraction of the rays in the eye, **we** will focus only on its role.

The most prominent difference is in the use of possessive pronouns, which are on average three times more frequent in NIK than in PAR (see Table 2).

E.g.

Chelovek stol' sil'no otlichayetsya ot brat'yev **nashikh** men'shikh, chto nayti vse neobkhodimyye zven'ya v dlinnoy tsepoche, u nachala kotoroy stoit chelovekoobraznaya obez'yana, ne udayotsya (NIK). - A person is so different from **our** smaller brothers that it is impossible to find all the necessary links in a long chain, at the beginning of which there is an anthropoid ape.

Ogromnykh uspekhev dostigla v **nashi** dni khirurgiya (PAR). - The tremendous success was achieved in **our** days by surgery.

However, there are some clear differences and exceptions (see Tables 3 and 5). As we can see, AF of **it** (ono; oni; ikh; im; imi; o nikh) in PAR is 471 in comparison to 295 in NIK with; RF is 7,998 in PAR and 7,128 in NIK respectively (Table 3). Secondly, normalized RF of personal pronouns 'vy; vas; vam; vami; o vas' (you) is 2,751 and it is 1,619 in NIK.

Interestingly that pronoun we (my) is used in NIK two times more often in NIK than in PAR: 3.89 vs 1.935 (see Table 2 below).

Another metric to focus is the frequency of pronoun ya (I) which in the Science textbook (PAR) is used only once, in a quotation: "Uzhe v glubokoy starosti on pisal: «Vsyu moyu zhizn' **ya** lyubil i lyublyu umstvennyy trud i fizicheskiy i, pozhaluy, dazhe bol'she vtoroy. He was advanced in years when he wrote: "All my life I have loved and still love mental and physical labor and, perhaps, I love the second even". In NIK on the contrary ya (I) may occur both in direct speech and addresses of the author to readers: "Pochemu **ya** zanimayus' imenno etim? V chom smysl **moyego** truda? Kak nayti svoyo mesto v trudovom protsesse? Zachem **ya** zhivu? Nuzhen li **ya** lyudyam, obshchestvu, strane? (Why am I doing just this? What is the meaning of **my** work? How to find your place in the labor process? Why do I live? Do I need people, society, country?)". The writer of PAR is not using "I" at all, which produces "the impression that the writer is withdrawing from all responsibility for the academic essay" (Bakhtin 1986: 293).

| Personal pronouns | PAR AF | PAR Normalized RF | NIK AF | NIK Normalized RF |
|--|--------|-------------------|--------|-------------------|
| I (ya; menya; mne; mnoy; obo mne) | 1 | 0,034 | 63 | 1,522 |
| You (ty; tebe; tebya; toboy; o tebe) | 0 | 0 | 26 | 0,628 |

| | | | | |
|--|------------|--------------|------------|--------------|
| He (on; yego; yemu; im; o nem) | 369 | 6,266 | 406 | 9,810 |
| She (ona; yeye; yey; o ney) | 112 | 1,902 | 155 | 3,745 |
| It (ono; oni; ikh; im; imi; o nikh) | 471 | 7,998 | 295 | 7,128 |
| We (my; nas; nam; nami; o nas) | 114 | 1,935 | 161 | 3,890 |
| You (vy; vas; vam; vami; o vas) | 162 | 2,751 | 67 | 1,619 |
| Total | 1229 | 20,887 | 1173 | 28,345 |

Table 3. The range of personal pronouns used in both textbooks for the 8th Grade of Russian secondary schools.

Table 4 shows that absolute and normalized relative frequency of reflexive pronouns in NIK is substantially higher, namely 118 and 2,851, as compared with 77 and 1,307 in PAR.

| Reflexive pronoun | PAR AF | PAR Normalized RF | NIK AF | NIK Normalized RF |
|--|-----------|-------------------|------------|-------------------|
| oneself (sebya; sebe; soboy; o sebe) | 77 | 1,307 | 118 | 2,851 |

Table 4. Absolute and normalized frequency of reflexive pronouns used in the textbooks for the 8th Grade of Russian secondary schools.

Figures in Table 5 demonstrate differences in the range of 'nash (our, sing, masc, nominative case); nashego (our, sing, masc, genitive case); nashemu (our, sing, masc, dative case); nashim (our, pl, dative case); o nashem (our, sing, masc, accusative case); nash'e' (our, sing, neutr, nominative case) in PAR (AF are 47 and 14; RF are 0,798 and 0,238 correspondingly) and in NIK (AF are 11 and 9; RF are 0,265 and 0,217) are significant.

The texts and the data allow us to conclude that both Russian science and non-science writers tend to use the inclusive pronouns **my** (we), **nam** (us) and **nash** (our). As the examples below show (which is also true about the majority of all cases) the authors are using back-reference thus suggesting that the writer expects readers share his experience/opinion/assumptions.

E.g.

*No mnogoye meshayet **nam** byt' distsiplinirovannymi!* (NIK) - *But much prevents **us** from being disciplined!*

***My** poluchayem neobkhodimyye nam organicheskiye soyedineniya i mineral'nyye soli iz pishchi rastitel'nogo i zhivotnogo proiskhozhdeniya* (PAR). - ***We** receive the necessary organic compounds and mineral salts from vegetable and animal food.*

We may also agree with Harwood (2005) that inclusive pronouns are markers of "low-risk, discrete instances of textual authorial intervention" (Harwood 2005: 344). Hence, the data, which are quite visible in both textbooks, confirm the assumptions of Pyankova (1994) and Vassileva (1998) that the general tendency of the writers of these particular classroom textbooks is to disguise their personal opinion. However conclusions of the kind are to be verified on a larger corpus of a wider range of authors. Thus, it is obvious that one of the limitations of the study is a small corpus we conduct our research on: to further clarify or contradict the hypothesis what is needed is a similar research on a larger corpus.

| Possessive pronouns | PAR AF | PAR Normalized RF | NIK AF | NIK Normalized RF |
|--|--------|-------------------|--------|-------------------|
| My (Moy; moyego; moyemu; moim; o moyem) | 0 | 0 | 9 | 0,217 |
| My (moya; moyey; moyu; o | 1 | 0,017 | 9 | 0,217 |

| | | | | |
|--|-----------------------|--------------|-----------|--------------|
| moyey) | | | | |
| My (moyo) | 0 | 0 | 5 | 0,121 |
| My (moi; moih; moim; moimi; o moikh) | 0 | 0 | 2 | 0,048 |
| Your (tvoy; tvoyego; tvoyemu; tvoim; o tvoyem) | 0 | 0 | 4 | 0,096 |
| Your (tvoja; tvoyey; tvoyu; o tvoyey) | 0 | 0 | 4 | 0,096 |
| Your (tvoyo) | 0 | 0 | 1 | 0,024 |
| Your (tvoi; tvoikh; tvoim; tvoimi; o tvoikh) | 0 | 0 | 1 | 0,024 |
| My (svoy; svoyego; svoyemu; svoimi o svojom) | 23 | 0,391 | 69 | 1,667 |
| My/her own (svoya; svoyey; svoyu; o svoyey) | 26 | 0,442 | 64 | 1,547 |
| My/its own (svoyo) | 0 | 0 | 15 | 0,362 |
| My (svoi; svoikh; svoimi; o svoikh) | 17 | 0,289 | 104 | 1,767 |
| Our (nash; nashego; nashemu; nashim o nashem) | 47² | 0,798 | 11 | 0,265 |
| Our Nashe | 14 | 0,238 | 9 | 0,217 |
| Our (nasha; nashey;nash o nashey; nashi nashikh; nashim nash; o nashikh) | 51 | 0,866 | 56 | 1,353 |
| your (vash; vashego; Vashemu; vashim; o vashem) | 2 | 0,034 | 6 | 1,144 |
| Your (vasha; vashey vashu; o vashey) | 2 | 0,034 | 4 | 0,096 |
| your (vashe) | 0 | 0 | 1 | 0,024 |
| Your (vashi; vashikh vash; vash; o vashikh) | 2 | 0,034 | 9 | 0,217 |
| Total | 185 | - | 381 | - |

Table 5. The range of possessive pronouns used in the textbooks for the 8th Grade of Russian secondary schools.

An important finding of the research is the fact that plural pronouns are more frequent than singular (See table 5). Researchers provide a number of explanations of more frequent plural pronouns: (a) politeness, cooperation, academic courtesy [Gergokaeva 2008]; (b) "modesty" [Glushko 1979; Gnezdechko 2005]; (c) inclusive 'we' when one is presented as a member /part of a community [Boldyreva 2014]; (d) search for objectivity [Ivanov 1978]; (e) ideological dictate [Miroshnichenko 1995].

² In bold we mark the cases of higher metrics in PAR than in NIK.

5. DISCUSSION

The results of the study aimed at defining frequencies of pronouns in two Russian classroom books demonstrated that there is a strong tendency to use more pronouns in textbooks on Social science than in the Science textbook (Human anatomy, physiology, hygiene).

We view the distinction to be mostly caused by cross-cultural differences, not purely authorial. One of the most convincing explanations of using 'my' (we) instead of ya (I) may be collectivism which Russia has been practicing for many centuries: for many years of the Soviet power ya (I) was rooted out of the Russian discourse (cf. the Russian saying 'Ya is the last letter in the alphabet' which vividly demonstrating negative attitude to people emphasizing their personal opinion has always been popular among Russians). We argue that the choice of a pronoun is determined by the conventions of the community to which the writer belongs. The communities of writers working in Science and Humanities have numerous similarities but their differences appear to be the factors determining the choice of pronouns.

Although an idea of this kind is to be verified on a bigger corpus of Russian academic discourse. Thus, the perspective of the study lies in (1) testing the hypothesis that science texts contain fewer pronouns than social (non-science) texts; (2) determining relative frequencies of different types of pronouns in Russian science and social (non-science) texts.

We also suggest that relative frequencies of pronouns may correlate with text complexity. We believe that language distributional data such as that play an important role in understanding the nature of text complexity.

6. CONCLUSION

The article presents the results of the analysis aimed at cross-discipline quantitative (absolute and relative frequencies) and distributional (use in various contexts) characteristics of Russian pronouns in secondary school classroom texts on Social Studies and Science. The question that we planned to investigate was whether the behavior of pronouns differs in science and non-science texts, in relation to text complexity.

Summarizing the uses of first-person pronouns in Russian textbooks, we conclude that they demonstrate very rare instances of 'ya'(I), while the occurrences of 'my' (we) are three times more numerous. Russian authors of both textbooks on Science and Humanities prefer the pronoun of the first person plural (my (we), the third person both singular and plural (ono (it), on (he), ona (she), oni (they)).

The research may be useful in a number of areas including teaching Academic Russian and Pragmatics on tertiary levels. Differences in the use of pronouns across Science and Humanities may also be helpful to writers and researchers in the two above mentioned areas.

However, there are a number of limitations to our study and further investigation would be necessary before announcing correlation of pronouns frequency and text complexity.

ACKNOWLEDGEMENTS

The research presented in Parts Related Work, Methods, and Analysis was supported by the subsidy of the Russian Government to support the Program of Competitive Growth of Kazan Federal University.

The research presented in Parts Discussion and Conclusion of the article was financially supported by the Russian Science Foundation, grant № 18-18-00436.

REFERENCE LIST

- AntConc. (2018). Softpedia. <http://www.softpedia.com/get/Science-CAD/AntConc.shtml>
- Bakhtin, M. (1986) The problem of speech genres and the problem of the text in linguistics, philology and the human sciences: an experiment in philosophical analysis. In C. Emerson & M. Holquist (eds.) Bakhtin: speech genres and other late essays. [Trans. V. McGee.] Austin: University of Texas Press. 250-317.
- Boldyreva A.A. Peculiarities of Expressing Author's I in Academic Discourse [Osobennosti vyrazheniya avtorskogo «ya» v nauchnom diskurse], Available at <http://tp11999.narod.ru/> (accessed 13.04.2014).
- Frequency Dictionary of Russian Ed. By L.N. Zazorina. (1977). M.: Russian language.
- J. Chall and E. Dale. (1995) Readability revisited: The new Dale-Chall readability formula. Brookline Books.

- Cherry, R.D. (1998). "Ethos versus Persona: Self-representation in written discourse". *Written Communication* 15, pp. 384-410.
- N. Duran, C. Bellissens, R. Taylor, and D. McNamara. (2007). Quantifying text difficulty with automated indices of cohesion and semantics. In *Proceedings of the Cognitive Science Society*, volume 29.
- Gabitov, Azat I.; Solnyshkina, Marina I.; Shayakhmetova, Liliya Kh (2017) Text Complexity In Russian Textbooks On Social Studies in *REVISTA PUBLICANDO* V: 4 Issue: 13. pp.: 597-606.
- Gafiyatova, E.V., Korovina, I.V., Solnyshkina, M.I., Yarmakeev, I.E. (2017). Deictic elements as means of text cohesion and coherence in academic discourse. *Journal of Social Studies Education Research*.
- Gergokaeva, D.D. (2008). *Egocentrism of Linguistic Discourse [Egocentrism lingvisticheskogo diskursa]*. Moscow.
- Glushko, M.M. (1979). Structure of the academic text in terms of systemized nature of its pragmatic, semantic and syntactic characteristics. *Linguistic researches of Academic Discourse [Organizatsiya nauchnogo teksta v svete sistemnosti ego pragmaticheskikh, semanticheskikh i sintaksicheskikh kharakteristik. Lingvisticheskie issledovaniya nauchnoi rechi]*. Moscow.
- Gnezdechko, O.N. *Authorization of the Academic Discourse [Avtorizatsiya nauchnogo diskursa: kommunikativno-pragmaticheskii aspekt]*. Kiev, 2005.
- Harwood, N. (2005). "We do not seem to have a theory... The theory I present here attempts to fill this gap': Inclusive and exclusive pronouns in academic writing", *Applied Linguistics*, 26(3), pp. 343-375.
- Hyland, K. (2001). Bringing in the Reader: Addressee Features in Academic Writing. *Written Communication*, 18, pp. 549-574.
- Ivanov, V.V. (1978). *Odd or Even. Asymmetry of the Brain and Sign Systems [Chet i nechet. Assimetriya mozga i znakovykh system]*. Moscow.
- Kuo, C. (1999). "The use of personal pronouns: Role relationships in scientific journal articles". *English for Specific Purposes* 18, pp. 121-138.
- McNamara, D.S., Graesser, A., Cai, Zh., and Kulikowich, J. (2011). Coh-matrix easability components: Aligning text difficulty with theories of text comprehension. In annual meeting of the American Educational Research Association, New Orleans, LA.
- Miroshnichenko, A. (1995). *Interpretation of Discourse: Basics of Linguistic and Ideological Analysis [Tolkovanie Diskursa: Osnovy Lingvisticheskogo i Ideologicheskogo Analiza]*. Rostov-na-Donu.
- Pyankova, T. M. (1994). *ABC perevodchika nauchno-tekhnicheskoi literaturi [ABC of scientific and technical literature translation]*. Moscow: Letopis.
- Roland, D., Dick, F., & Elman, J. L. (2007). Frequency of Basic English Grammatical Structures: A Corpus Analysis. *Journal of Memory and Language*, 57(3), 348–379. <http://doi.org/10.1016/j.jml.2007.03.002>
- School textbooks and manuals. (2017). *Vse uchebniki*. <http://vseuchebniki.net/>
- Solovyev, V. D., Solnyshkina, M.I., Ivanov, V.V., Timoshenko, S. V. (2018). Complexity of Russian Academic Texts as the Function of Syntactic Parameters, 19th International Conference on Computational Linguistics and Intelligent Text Processing, CICLing 2018, March 18 to 24, 2018, Hanoi, Vietnam
- Tokenization. (2008). Cambridge University Press. <https://nlp.stanford.edu/IR-book/html/htmledition/tokenization-1.html>
- Vassileva, I. (1998). "Who am I/who are we in academic writing? A contrastive analysis of authorial presence in English, German, French, Russian and Bulgarian", *International Journal of Applied Linguistics*, 8(2), pp. 163-185.