

РАЗДЕЛ II. ПРИКЛАДНАЯ ЛИНГВИСТИКА  
SECTION II. APPLIED LINGUISTICS

УДК 004.8:811.1/.2

DOI: 10.18413/2313-8912-2024-10-3-0-4

Гатиятуллина Г.М.<sup>1</sup>   
Гафиятова Э.В.<sup>2</sup>   
Данилов А. В.<sup>3</sup> 

Динамика уровня сложности как функция  
морфологических параметров текста на русском языке

<sup>1</sup> НИЛ «Текстовая аналитика», Казанский (Приволжский) федеральный университет  
ул. Кремлевская, 18, Казань, 420008, Россия  
*E-mail:* [ggaliya-m@mail.ru](mailto:ggaliya-m@mail.ru)  
ORCID: 0000-0001-6752-2100

<sup>2</sup> Казанский (Приволжский) федеральный университет  
ул. Кремлевская, 18, Казань, 420008, Россия  
*E-mail:* [rg-777@yandex.ru](mailto:rg-777@yandex.ru)  
ORCID: 0000-0003-3190-4566

<sup>3</sup> НИЛ «Текстовая аналитика», Казанский (Приволжский) федеральный университет  
ул. Кремлевская, 18, Казань, 420008, Россия  
*E-mail:* [tukai@yandex.ru](mailto:tukai@yandex.ru)  
ORCID: 0000-0002-2358-1157

*Статья поступила 21 июля 2024 г.; принята 15 сентября 2024 г.;  
сентября 30 июня 2024 г.*

**Источник финансирования:** Исследование выполнено при поддержке гранта РНФ 24-28-01355 «Жанрово-дискурсивные характеристики текста как функция лексического диапазона».

**Аннотация.** Представленное исследование нацелено на (1) выявление предикторов морфологической сложности и (2) идентификацию маркеров, дифференцирующих предметные области учебного текста на русском языке. Общий объем корпуса, включающий двадцать учебников по биологии и обществознанию трех уровней сложности, соответствующих 6 – 7, 8 – 9, 10 – 11 классам общеобразовательной российской школы, составил 941963 словоформы. Лингвистическая сложность текстов рассчитывалась при помощи модифицированной для русского языка формулы читабельности Флеша–Кинкейда, а взаимозависимость параметров оценивалась на основе корреляционного анализа с использованием программы STATISTICA. Расчеты значений лингвистических параметров, включающих дистрибуцию имени существительного, прилагательного, глагола, а также индекс читабельности, производились при помощи текстового профайлера для русского языка RuLingva ([rulex.kpfu.ru/](http://rulex.kpfu.ru/)), а метрики частотности девербативов и деадъективов имен существительных были выделены авторами исследования вручную.

Дистрибутивный анализ частотности метрик субстантивации производился с нормализацией текста на 10000 словоупотреблений. Последовательное контрастирование индекса читабельности указанных текстов с метриками «дистрибуция имени существительных», «лексическая плотность», «девербация», «деадъективация» продемонстрировало их линейную взаимозависимость. Читабельность изучаемых текстов находится в обратно пропорциональной зависимости от дистрибуции глагола. Морфологический анализ подтвердил высокий уровень номинативности текстов и стабильный рост субстантивов, объективируемый в увеличении доли суффиксов девербации и деадъективации от 6 к 11 классу. Метрики лексической плотности, дистрибуции имени прилагательного и отдельных суффиксов субстантивов также выступают в качестве дискриминант, дифференцирующих учебные тексты по биологии и обществознанию указанных уровней сложности. Результаты исследования будут полезны специалистам в области текстовой аналитики, компьютерной лингвистики, жанроведения, а также разработчикам тестов и авторам учебно-методических материалов. Перспектива исследования видится в изучении долей сложных терминов-лексем, образованных от латинских и греческих корней, в учебных текстах. Предполагается, что данный параметр имеет потенциал выступать в качестве предиктора лингвистической сложности и дискриминанта предметной области.

**Ключевые слова:** Учебные тексты; Читабельность; Русский язык; Учебник; Морфологические параметры; Номинализация.

**Информация для цитирования:** Гатиятуллина Г. М., Гафиятова Э. В., Данилов А. В. Динамика уровня сложности как функция морфологических параметров текста на русском языке // Научный результат. Вопросы теоретической и прикладной лингвистики. 2024. Т. 10. № 3. С. 70-90. DOI: 10.18413/2313-8912-2024-10-3-0-4

UDC 004.8:811.1/2

DOI: 10.18413/2313-8912-2024-10-3-0-4

Galiya M. Gatiyatullina<sup>1</sup>   
Elzara V. Gafiyatova<sup>2</sup>   
Andrei V. Danilov<sup>3</sup> 

**Text complexity increase in Russian texts as a function of morphological changes**

<sup>1</sup> Text Analytics Laboratory, Kazan Federal University,  
18 Kremlevskaya St., Kazan, 420008, Russia  
*E-mail:* [galiya-m@mail.ru](mailto:galiya-m@mail.ru)  
ORCID: 0000-0001-6752-2100

Kazan Federal University,  
<sup>2</sup> 18 Kremlevskaya St., Kazan, 420008, Russia  
*E-mail:* [rg-777@yandex.ru](mailto:rg-777@yandex.ru)  
ORCID: 0000-0003-3190-4566

<sup>3</sup> Text Analytics Laboratory, Kazan Federal University,  
18 Kremlevskaya St., Kazan, 420008, Russia  
*E-mail:* [tukai@yandex.ru](mailto:tukai@yandex.ru)  
ORCID: 0000-0002-2358-1157

Received 21 July 2024; accepted 15 September 2024; published 30 September 2024

**Acknowledgments:** The research was supported by the RSF grant 24-28-01355 “Genre-discourse characteristics of the text as a function of lexical range”.

**Abstract:** The study aims to identify (1) morphological complexity predictors and (2) domain inherent markers able to differentiate subject areas of academic text in Russian. The total size of the corpus, including textbooks on biology and social studies of three levels of complexity, corresponding to 6-7, 8-9, and 10-11 grades of the Russian school, amounted to 941963 tokens. The linguistic complexity of the texts was assessed using the Flesch-Kincaid readability formula modified for the Russian language, and the interdependence of the parameters was measured based on the correlation analysis conducted with STATISTICA. Calculation of linguistic parameters values, including distribution of nouns, adjectives, verbs, and readability index, were performed using RuLingva (rutex.kpfu.ru/), a text profiler for the Russian language, while the frequency metrics of deverbatives and deadjectives were identified by the contributors manually. To ensure comparability of the metrics, the distributional analysis of deverbatives and deadjectives was performed in the corpus normalized to 10000 tokens. Metrics “noun distribution”, “lexical density”, “deverbation”, “deadjectivation” demonstrated linear interdependence with readability and as such can be viewed as complexity predictors. Inverse correlation was revealed between text readability and verb distribution. Morphological analysis confirmed a high level of texts nominativity and a stable growth of substantives frequency. The latter explicates in an increase in the frequency of deverbation and deadjectivation suffixes in texts from the 6<sup>th</sup> to the 11<sup>th</sup> grade. Metrics of lexical density, adjective distribution and substantive suffixes demonstrate ability to discriminate academic texts domains. The research findings are applicable in text analytics, computational linguistics, genre studies, and can be useful for test developers and textbook writers. The authors view the research prospect in the study of compounds of Latin and Greek origin in academic texts. The identified parameters may be used as linguistic complexity predictors and domain discriminants.

**Key words:** Academic texts; Readability; The Russian language; Textbooks; Morphological parameters; Nominalization.

**How to cite:** Gatiyatullina, G. M., Gafiyatova, E. V., Danilov, A. V. (2024). Text complexity increase in Russian texts as a function of morphological changes, *Research Result. Theoretical and Applied Linguistics*, 10 (3), 70-90. DOI: 10.18413/2313-8912-2024-10-3-0-4

## Введение

Одной из детерминант успеха чтения и понимания текста является его сложность или читабельность. Сложность текстов на английском языке стала объектом исследования почти 70 лет тому назад (Flesch 1949; McLaughlin 1969; Kincaid et al. 1975; Just and Carpenter 1987; Alderson 2000; Koda 2005; Crossley et al. 2008), а исследования сложности текстов на русском языке начали активно изучаться

с 1970-х гг. и в настоящее время продолжают находиться в фокусе современных исследователей (Микк 1970, Мацковский 1976, Оборнева 2006, Шпаковский 2007, Solnyshkina and Kiselnikov 2015).

Для оценки читабельности текста существуют более 20 формул (DuBay, 2007). Однако преимущественное большинство этих формул являются функциями двух метрик – «длина

предложения» и «длина слова». Наиболее распространенной является формула удобочитаемости по Р. Флешу, разработанная для английского языка: УФ (формула удобочитаемости) (англ) =  $206.835 - (1.015 \times \text{средняя длина предложения}) - (84.6 \times \text{среднее число слогов})$ . Для русских художественных текстов формула Р. Флеша была адаптирована И.В. Оборневой (2006): УФ (рус) =  $206,836 - (1,52 \times \text{средняя длина предложения}) - (65,14 \times \text{среднее число слогов})$ . Оценки сложности научных и учебно-научных текстов на русском языке производится в настоящее время по формуле  $ФК_{\text{Мод}}$  (модифицированная формула Флеша-Кинкейда) =  $0.36 \times \text{средняя длина предложения} + 5.76 \times \text{среднее число слогов в слове} - 11.97$  (Solnyshkina et al. 2018).

Известно, что формулы читабельности даже в случае их использования для текстов соответствующих функциональных стилей демонстрируют весьма ограниченную достоверность. Именно поэтому, по признанию отечественных и зарубежных ученых, работающих в рамках дискурсивной комплексологии, вопрос о спектре параметров, влияющих на сложность текста, остается исследовательской нишей (Vahrusheva et al., 2023).

Представленное исследование осуществлено с целью тестирования двух **гипотез**: (1) о взаимозависимости читабельности учебно-научного текста, с одной стороны, и следующих морфологических параметров, с другой стороны: дистрибуция существительных, прилагательных и глагола, лексическая плотность, доля существительных, образованных при помощи девербации и деадъективации; (2) о способности морфологических параметров дифференцировать предметные области.

### Обзор литературы

Письменные научные тексты, в том числе лингвистические характеристики

вузовских и школьных учебников не раз становились предметом исследований (Biber 1989; Biber et al. 1999; Biber and Gray 2013, 2016; Halliday and Martin 2003; Banks 2005, 2008; Fang, Schleppegrell and Cox 2006). Для учебных и научных текстов на английском языке выявлена высокая дистрибуция имен существительных и именных словосочетаний. Выраженная предметность (термин Б. Н. Головина), именной характер (Жеребило 2010) или номинативность (Валгина, 2003) научного стиля резко контрастирует с разговорной речью, в которой существенно преобладают глаголы и придаточные конструкции (Biber 1989; Biber and Gray 2013, 2016).

Выявленные зависимости подтверждаются и результатами исследований развития речи: именно номинативные группы представляют особую трудность для понимания учениками начальной школы, поскольку в привычной устной речи существительные используются преимущественно для обозначения предметов и людей, а глаголы – для обозначения действий. Способность заменять придаточные предложения номинативными группами появляется у носителей языка значительно позднее способности формировать простые предложения. Например, фраза *the rate of glass crack growth* (скорость роста трещин стекла) значительно труднее для понимания, чем *how quickly cracks in glass grow* (как быстро разрастаются трещины на стекле) (Halliday and Martin, 2003: 82). Учебный дискурс предполагает высокую степень субстантивации глаголов, поэтому, по мнению М. Халлидея, вопрос восприятия дискурса учебником является отчасти вопросом взросления: даже отучившись пару лет в средних классах ребенок может испытывать трудности с пониманием текста, особенно текстов по точным и техническим предметам, так как именно в них наиболее часто встречаются субстантивированные глаголы и отглагольные существительные (Halliday

and Martin, 2003: 82). Таким образом М. Халлидей фактически предлагает список предикторов сложности, которые в значительной степени обуславливают трудности восприятия учебного дискурса.

По мнению ряда ученых научный дискурс характеризуется двумя активными процессами: девербацией, т.е. образованием отглагольных существительных со значением отвлеченного действия (Biber 2013), и деадъективацией, образованием имен существительных со значением отвлеченного качества (Martin 1991). По мнению О. Д. Митрофановой, такие суффиксы как *-ость*, *-ние* (*-ение*, *-ание*) в общелитературном языке не имеют форм множественного числа и предполагают высокую абстрактность, однако под влиянием языка науки данные номинативы могут обнаруживать множественное число для повышения конкретизации и индивидуализации (Митрофанова 1973: 62-63).

Номинализация или субстантивация, т.е. процесс перехода лексем из класса глаголов, прилагательных, предлогов, союзов в класс существительных, позволяет вместить или «упаковать» больше информации на предложение или сделать предложения более содержательными (Eggins 2004: 96). Это связано с потенциалом «номинативной группы», т.е. части клаузы, которая состоит из существительных и слов, служащих модификатором к имени существительному. С. Эггинс (2004: 95) также пишет, что номинализация обладает двумя текстовыми преимуществами: риторической организацией текста и лексической плотностью. Высокая степень лексической плотности, с одной стороны, наполняет текст содержанием, но с другой стороны, приобретает особую значимость

при оценке трудности восприятия текста. Оценка лексической плотности текста рассчитывается как отношение суммы словоупотреблений знаменательных частей речи, т.е. имени существительного, прилагательного, глагола и наречия, ко всему объему текста в словоупотреблениях (Biber et al. 2021; To et al. 2013).

Модели субстантивации, выявленные для русского языка немногочисленны, но весьма продуктивны (см. Петрова 2009; Мамонтов, Мамонтова 2012). Девербация имеет место при образовании существительных, например, *чтение*, *чихание* (V+ а(е)ни-е); *занятие*, *принятие* (V+ яти-е); *бахвальство*, *баловство* (V+ ств-о); *акция*, *дивизия* (V+ц/зи-я). Существительные *героизм*, *эгоизм* (Adj+изм); *выносливость*, *заносчивость* (Adj+ость); *широта*, *доброта* (Adj+ота); *глубина*, *кривизна* (Adj+ и(з)на) образованы по словообразовательным моделям деадъективации.

#### **Материал и методы исследования**

##### *Материал исследования.*

Исследование осуществлено на материале российских учебников по биологии и обществознания. Общий объем подкорпуса текстов по биологии составил 399555 словоупотреблений (см. Таблицу 1 и Список иллюстративных источников).

Гуманитарный блок исследования представлен двумя линейками российских учебников по обществознанию для 6-11 классов: (а) под редакцией А. Ф. Никитина и (б) под редакцией Л. Н. Боголюбова. Объем подкорпуса российских учебников по обществознанию составил 542408 словоупотреблений (см. Таблицу 1 и Список иллюстративных источников).

Все учебники в составе корпуса были разделены на три уровня сложности: уровень I – 6-7 классы, уровень II – 8-9 классы, уровень III – 10-11 классы.

**Таблица 1.** Объем корпуса исследования  
**Table 1.** Size of the research corpus

Уровень	Учебник	Количество словоупотреблений	Учебник	Количество словоупотреблений	Итого
<b>Биология</b>					
I	БС 6 *	26870	БП 6**	37164	64034
	БС 7	27956	БП 7	47969	75925
II	БС 8	30136	БП 8	52173	82309
	БС 9	30010	БП 9	50927	80937
III	БС 10-11	27408	БП 10-11	68942	96350
Итого					<b>399555</b>
<b>Обществознание</b>					
I	ОН 6	16843	ОБ 6 ****	16444	33287
	ОН 7	23111	ОБ 7	23855	46966
II	ОН 8	50492	ОБ 8	40986	91478
	ОН 9	44508	ОБ 9	43231	87739
III	ОН 10	101716	ОБ10	102140	203856
	ОН 11	39289	ОБ11	39793	79082
Итого					<b>542408</b>

\*Буквами БС маркированы учебники по биологии Л. Н. Сухоруковой и В. С. Кучменко. Цифра указывает класс.

\*\* БП маркируют учебники по биологии В. В. Пасечника и А. А. Каменского.

\*\*\*ОН маркируют учебники по обществознанию под редакцией А. Ф. Никитина.

\*\*\*\* ОБ маркируют учебники по обществознанию под редакцией Л. Н. Боголюбова.

*Методы и алгоритм исследования.*

1. Расчеты метрик морфологических параметров (дистрибуция имени существительного, прилагательного, глагола) и индекса читабельности (ФКмод) осуществлялись при помощи текстового анализатора RuLingva [rulex.kpfu.ru/]. Индекс читабельность ранжирует тексты по годам обучения, т.е. текст, имеющий читабельность 6 (ФКмод=6), соответствует лингвистическим и когнитивным способностям школьников 6 класса и т.д. RuLingva производит морфологический анализ текста и позволяет выгружать списки знаменательных частей речи. Для последующего сравнения каждая метрика была нормализована на 1000 словоупотреблений (см. Таблицу 2).

2. В группе имен существительных каждого текста вручную были выделены девербативы и деадъективы.

3. Для расчета дистрибуции количество каждого суффикса было нормализовано на 10 тыс. словоупотреблений

4. Расчеты среднестатистических значений, корреляции читабельности с морфологическими параметрами (дистрибуция имени существительного, прилагательного, глагола, лексическая плотность, степень субстантивации), а также оценка статистически значимых различий в показателях текстов изучаемых дисциплин осуществлялись в программе Statistica [statsoft.com/]. Для визуализации динамики параметров, коррелирующих с читабельностью текста, были построены

гистограммы и графики (см. Рис. 1-5). Все графики и гистограммы строились на основе метрик каждого из сегментов (глав) в составе изучаемых текстов учебников. Количество изученных сегментов указано в Таблице 2. Например, в учебниках обществознания уровня I

проанализировано 203 отрывка, на уровне II – 143 и т.д.

### Результаты

Рассчитанные при помощи текстового анализатора Rulingva метрики изучаемых лингвистических параметров демонстрируют сильную корреляцию с индексом читабельности (см. Табл. 2).

**Таблица 2.** Лингвистические параметры текстов трех образовательных уровней (I – III)  
**Table 2.** Linguistic features of texts of three grade levels (I – III)

параметр	Уровень I (6-7 классы)			Уровень II (8-9 классы)			Уровень III (10-11 классы)		
	Ср. стат. био (N=203)	Ср. стат. общ (N=76)	р-значение, U-критерий Манна-Уитни	Mean био (N=143)	Mean общ (N=112)	р-значение, U-критерий Манна-Уитни	Mean био (N=145)	Mean общ (N=110)	р-значение, U-критерий Манна-Уитни
класс	6,458	6,526	0,4	8,6	8,473	0,1	10,1	10,5	0,8
Количество слов (Tokens)	702,6	1074,5	<.001*	934,3	1588,4	<.001*	670,3	2355,9	<.001*
Ср. количество слов в предложении	11,4	12,7	<.001*	13,8	14,5	.003*	14,6	16,6	<.001*
Ср. количество слогов в слове	2,6	2,4	<.001*	2,8	2,657	<.001*	2,8	2,8	0,2
Сущ\1000	397,8	341,4	<.001*	406,9	379,9	<.001*	403,1	391,2	<.001*
Прил\1000	151,2	118,0	<.001*	153,7	141,1	<.001*	145,9	163,6	<.001*
Гл.\1000	130	145,7	<.001*	124,8	127,7	0,4	123,7	115,9	<.001*
лекс.плотность\1000	718	657,7	<.001*	725,5	692,4	<.001*	715,2	713,3	0,6
Местоим\1000	54	103	<.001*	55,9	79,7	<.001*	59,5	68,7	<.001*
Индекс ФК (мод) SIS	7,2	6,8	<.001*	8,9	9,04	<.001*	9,3	8,8	<.001*
Наст.в.\1000	90,4	62,9	<.001*	83,0	63,5	<.001*	74,7	64,1	<.001*
Прош.в.\1000	24	41,9	<.001*	24,8	34,1	<.001*	33,0	29,2	0,7

\*  $p < .05$  – статистически значимые различия

Рассмотрим каждый из параметров.

#### Читабельность.

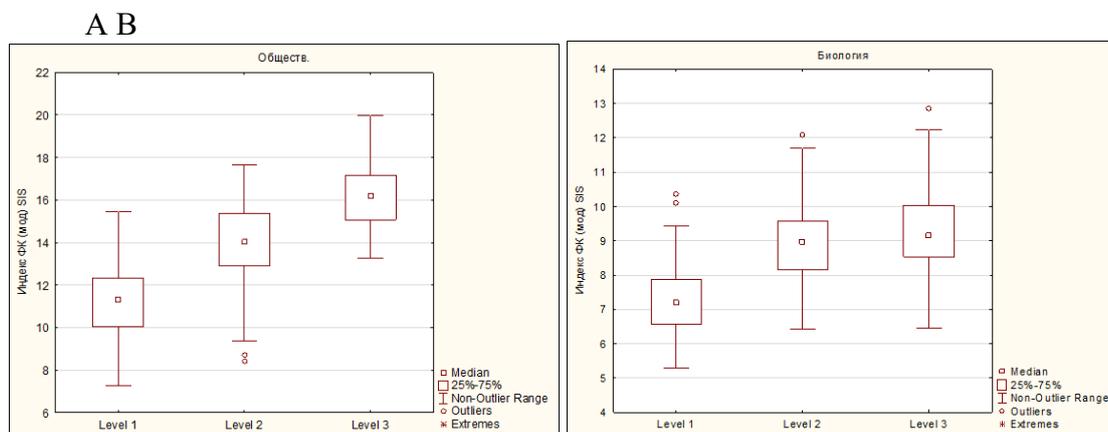
Анализ текстов показал, что для текстов обеих дисциплин характерна динамика роста от уровня I к уровню II, однако читабельность текстов по

обществознанию к уровню III снижается на 0,6, а тексты по биологии, наоборот, незначительно растут к уровню III. Читабельность текстов по

обществознанию для 6-7 классов (уровень I) составляет 6,8, для 8-9 классов (уровень II) читабельность вырастает до 9,04, а для

10-11 классов (уровень III) снижается до 8,8 (рис. 1 А).

**Рисунок 1.** Динамика индекса читабельности текста по обществознанию (А) и биологии (В)  
**Figure 1.** Readability rate in Russian texts of textbooks on Social Studies (A) and Biology (B)



Метрики, на основании которых осуществляется расчет читабельности, демонстрируют положительную динамику: длина слова на уровне I равна 2,4 слога, на уровне II – 2,7 слога, а на уровне III – 2,8 слога. Длина предложения в текстах по обществознанию составляет 12,7 слов на уровне I, в текстах уровня II среднестатистическая длина предложения составляет 14,5 слов, а на уровне III – 16,6 слов.

Читабельность текстов по биологии показывает положительную динамику: в текстах по биологии уровня I читабельность составляет 7,2, в текстах уровня II – 9, а на уровне III – 9,3 (рис. 1 В). Аналогичным образом линейно растут метрики параметров, формирующих читабельность. Длина слова на I и II уровнях выше, чем в текстах по обществознанию: на уровне I равна 2,6

слога, на уровне II – 2,8 слога, а на уровне III – 2,8 слога. Длина предложения в текстах по биологии в целом ниже, чем в текстах обществознанию, и составляет 11,4 слов на уровне I, на уровне II длина предложения составляет 13,8 слов, а на уровне III – 14,6 слов.

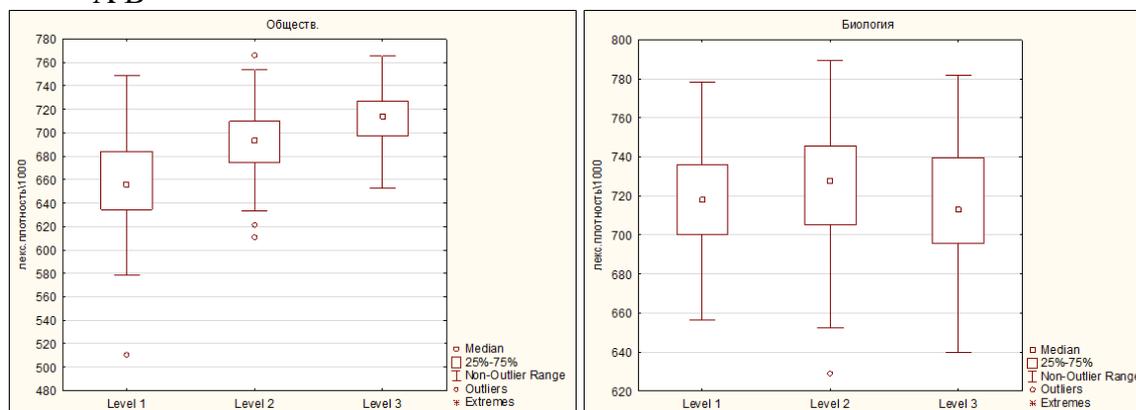
#### *Лексическая плотность.*

Динамика степеней лексической плотности в текстах обеих дисциплин отличается в значительной степени. В текстах по обществознанию лексическая плотность растет от уровня к уровню: на уровне I – 658 знаменательных частей речи на 1000 словоупотреблений, в текстах уровня II лексическая плотность выросла до 692 слов знаменательных частей речи на 1000 словоупотреблений, а на уровне III лексическая плотность составила 713,3 (рис. 2 А).

**Рисунок 2.** А) Лексическая плотность на 1000 слов (Обществознание); В) Лексическая плотность на 1000 слов (Биология)

**Figure 2.** А) Lexical density per 1000 words (Social Studies); В) Lexical density per 1000 words (Biology)

А В



В текстах по биологии метрики лексической плотности высоки в текстах всех уровней. Однако стабильный рост от уровня к уровню не выявлен. В текстах уровня I лексическая плотность составляет 718, далее растет в текстах уровня II, но на уровне III падает ниже уровня I и составляет 715,2 слов знаменательных частей речи на 1000 словоупотреблений (рис. 2 В). При этом необходимо отметить, что даже наиболее минимальное значение лексической плотности в текстах по биологии выше, чем максимальное значение лексической плотности в текстах по обществознанию.

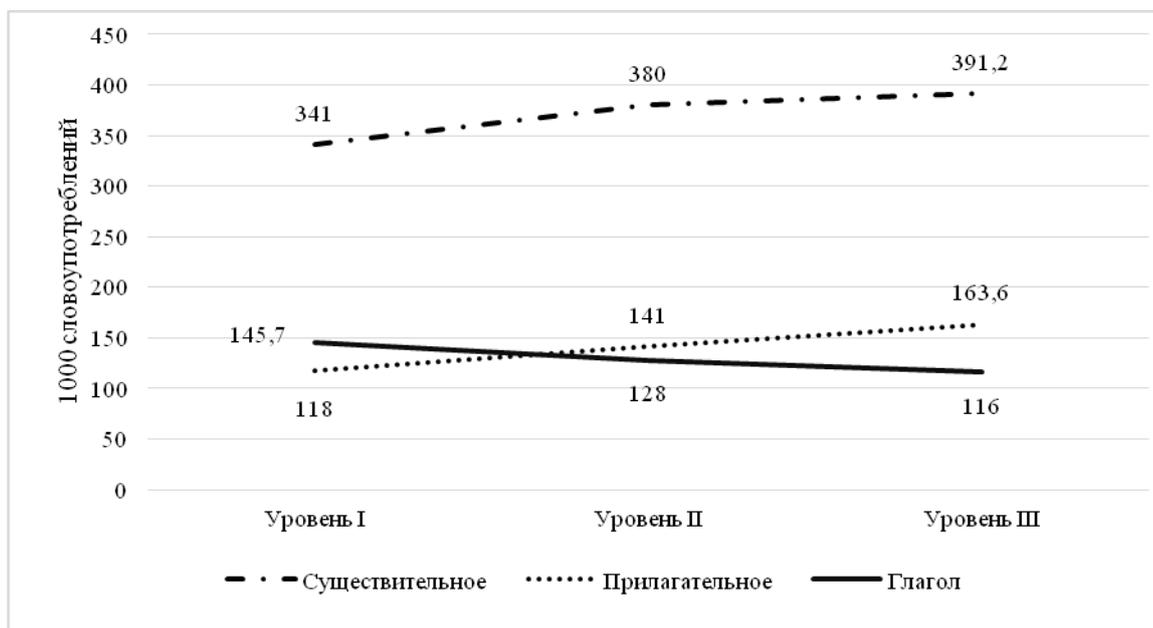
#### *Дистрибуция частей речи.*

Анализ дистрибуции частей речи показал существенное преобладание имен существительных в текстах обеих дисциплин. Тем не менее, в текстах по

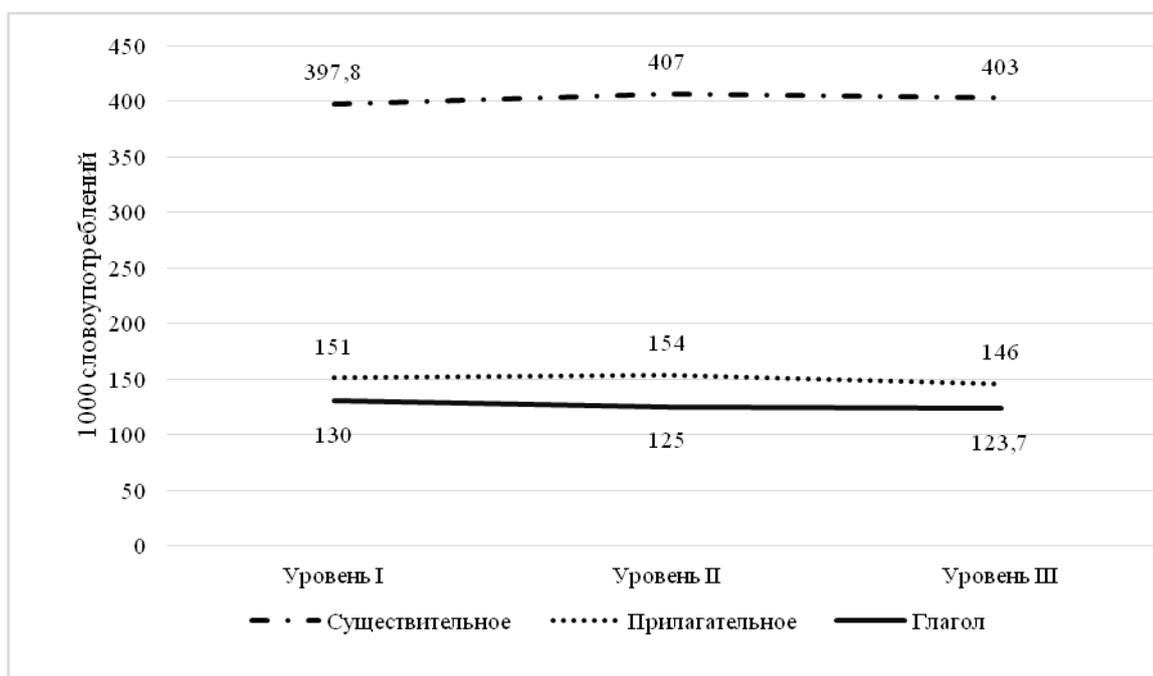
обществознанию динамика роста существительных наиболее выражена. В текстах уровня I дистрибуция имен существительных составляет 341 вхождение на 1000 словоупотреблений, в текстах уровня II этот параметр имеет значение 380 вхождений на 1000 словоупотреблений, а в текстах по обществознанию уровня III дистрибуция имен существительных составляет 391,2 вхождение на 1000 словоупотреблений (см. Рисунок 3).

В текстах по биологии дистрибуция существительных в целом выше, однако динамика роста неравномерна: на уровне I – 397,8 вхождений на 1000 словоупотребления, на уровне II – 407 вхождений, а в текстах уровня III – 403 вхождения (см. Рисунок 4).

**Рисунок 3.** Динамика дистрибуции частей речи в текстах по обществознанию  
**Figure 3.** Part of Speech distribution rate in Social Studies texts



**Рисунок 4.** Динамика дистрибуции частей речи в текстах по биологии  
**Figure 4.** Part of Speech distribution rate in Biology texts



Дистрибуция имен прилагательных в изучаемых текстах несколько ниже дистрибуции имен существительных: в текстах по обществознанию динамика роста дистрибуции прилагательных равномерна: 118 (I уровень) → 141 (II

уровень) → 163, 6 (III уровень) на 1000 словоупотреблений. Для сравнения в текстах биологии: 151 (I уровень) → 154 (II уровень) → 146 (III уровень) на 1000 словоупотреблений.

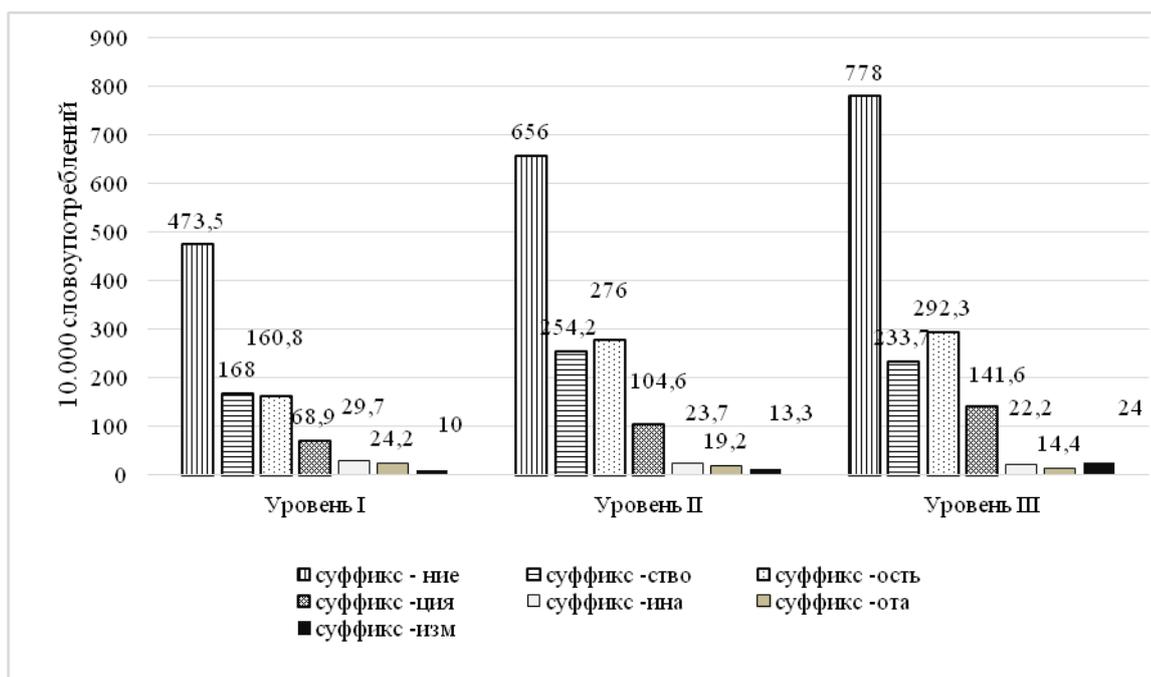
В текстах обеих дисциплин динамика роста дистрибуции глаголов стабильно отрицательная. Обществознание: 145,7 (I уровень) → 128 (II уровень) → 116 (III уровень) на 1000 словоупотреблений. В текстах по биологии: 130 (I уровень) → 125 (II уровень) → 123,7 (III уровень) на 1000 словоупотреблений.

*Номинативность текстов.*

В текстах обеих дисциплин выявлен активный процесс субстантивации. В текстах по обществознанию совокупность суффиксов субстантивации в текстах уровня I составляет 935,1 вхождений на 10 тыс. словоупотреблений, в текстах уровня II дистрибуция суффиксов субстантивации

растет до 1347 вхождений на 10 тыс. словоупотреблений, а в текстах по обществознанию уровня III суффиксы субстантивации вырастает до 1506,2 вхождений на 10 тыс. словоупотреблений. В текстах по биологии дистрибуция суффиксов субстантивации на всех уровнях в целом ниже, чем в текстах по обществознанию: на уровне I – 929 вхождений, в текстах уровня II наблюдается рост суффиксов субстантивации до 1285,4 на 10 тыс. словоупотреблений, а на уровне III – снижение до 1273,2 вхождений (см. Рисунок 5).

**Рисунок 5.** Динамика суффиксов субстантивации в текстах по обществознанию  
**Figure 5.** Distribution rate of nominalizations in Social Studies school textbooks



Динамика роста дистрибуции суффикса *-ние* является наиболее стабильной: в текстах по обществознанию уровня I – 473,5 вхождений на 10 тыс. словоупотреблений, уровня II – 656 вхождений, а в текстах уровня III его частотность возрастает до 778 вхождений (см. Рисунок 5).

Вторую позицию по частотности среди суффиксов субстантивации в

изучаемых текстах по обществознанию занимают суффиксы *-ство* и *-ость*. Суффикс *-ство* демонстрирует значительный рост от уровня I до уровня II: 168 → 254,2. При этом дистрибуция суффикса *-тво* незначительно падает в текстах уровня III и составляет 233,7 вхождений на 10 тыс. словоупотреблений. Динамика роста суффикса *-ость* стабильна на всех уровнях: 10, 13,3, 24 вхождений на 10 тыс. словоупотреблений.

уровне I, 276 вхождений – на уровне II и 292,3 – на уровне III.

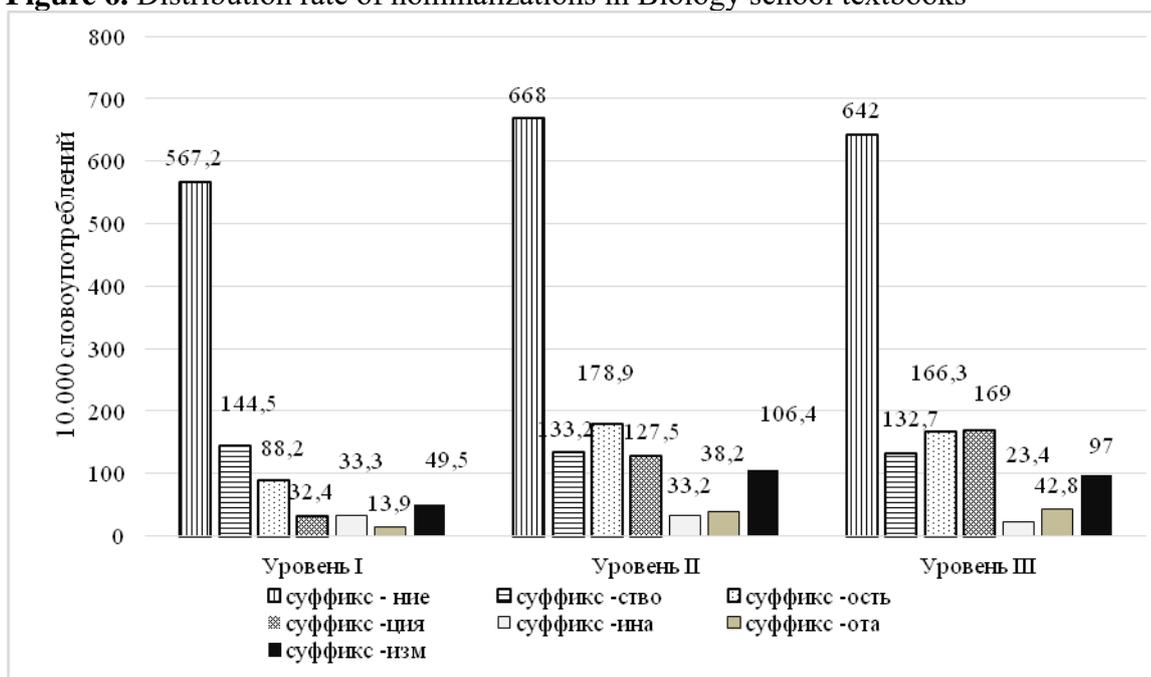
Из относительно менее частотных суффиксов (-ция, -ина, -ота, -изм) два суффикса -ция и -изм показывают стабильную тенденцию роста. Суффикс -ция: 68,9 вхождений на уровне I, 104,6 вхождений – на уровне II, 141,6 вхождений на уровне III в текстах по обществознанию. Суффикс -изм: 10 – на уровне I, 13,3 – на уровне II, 24 – на уровне III.

Суффиксы -ина и -ота, наоборот, демонстрирует стабильное снижение: -ина имеет 24,2 вхождения в текстах по обществознанию на уровне I, снижается до

23,7 в текстах уровня II, а на уровне III – до 22,2 вхождения. Дистрибуция суффикса -ота в текстах по обществознанию на уровне I составляет 24,2 вхождения, в текстах уровня II падает до 19,2, а на уровне III – снижается до 14,4.

Аналогично текстам по обществознанию в текстах по биологии наиболее частотным суффиксом субстантивации является суффикс -ние, однако динамика его роста не стабильна. Так, в текстах по обществознанию уровня I он имеет 567,2 вхождений, в текстах уровня II – 668 вхождений, а в текстах уровня III его дистрибуция незначительно снижается до 642 (см. Рис 6).

**Рисунок 6.** Динамика суффиксов субстантивации в текстах по биологии  
**Figure 6.** Distribution rate of nominalizations in Biology school textbooks



Второе место по частотности в текстах биологии занимают суффиксы -ство и -ость, однако динамика их дистрибуции является разнонаправленной. Суффикс -ство стабильно снижается в текстах по биологии от уровня I к уровню III: зафиксировано 168 вхождений суффикса -ство на уровне I, 133,2 вхождений – на уровне II, а на уровне III дистрибуция суффикса -ство также

незначительно падает и составляет 132,7 вхождения.

Динамика изменений частотности суффикса -ость не стабильна: в текстах по обществознанию на уровне суффикс -ость используется в среднем 88,2 раза, на уровне II его частотность поднимается до 178,9, а в текстах уровня III – падает до 166,3 вхождений.

Среди менее частотных суффиксов субстантивации (-ция, -ина, -ота, -изм) два

суффикса *-ция* и *-ота* показывают стабильную тенденцию роста, а суффикс *-ина* демонстрирует стабильное снижение. Что касается дистрибуции суффикса *-изм* в текстах по биологии, его метрики выстраиваются по параболе с ветвями вниз: от 49,5 к 106,4 и к 97.

### Обсуждение

Результаты представленного исследования в значительной степени способствуют расширению базы данных о референсных диапазонах метрик морфологических параметров изученных текстов. Их сопоставление с данными более ранних публикаций высвечивает меж- и внутриязыковые сходства и различия текстов разных жанров, с одной стороны, а также исследовательские ниши, с другой стороны.

Тексты современных российских учебников по обществознанию и биологии характеризуются приблизительно равной длиной слова: 2,7 в биологических текстах и 2,8 в текстах обществознания на уровне II и 2,8 для обеих дисциплин на уровне III. Длина предложения также приблизительно одинакова для обеих дисциплин: 14,5 слов в предложении по обществознанию и 13,8 по биологии на уровне II, на уровне III их длина предложения одинакова и составляет 16,6 слов. Незначительные различия в длине предложения и слова дают несколько более высокую читабельность текстов по биологии по сравнению с текстами по обществознанию (см. Рисунок 1 и 2). Однако в целом читабельность русскоязычных текстов ниже читабельности англоязычных текстов аналогичных дисциплин. Так, тексты уровня III по дисциплине *Social Studies* имеют индекс читабельности 13,3, а по дисциплине *Biology* – 13,04, что в целом значительно выше индекса читабельности текстов уровня III сравниваемых русскоязычных учебников (Gatiyatullina et al. 2020).

Лексическая плотность текстов биологии и обществознания на русском языке существенно выше по сравнению с

учебниками *Biology* и *Social Studies*: диапазон лексической плотности в англоязычных учебниках – 475 вхождений на 1000 словоупотреблений на уровне I в учебниках *Biology* и максимальное значение – 501 вхождение также в учебниках *Biology* (Gatiyatullina et al. 2020). В русскоязычных учебниках диапазон лексической плотности имеет диапазон от 658 до 718 вхождений. Очевидно, что данный факт можно объяснить принадлежностью языков сравниваемых текстов к разным типам: артиклевый английский аналитичен vs безартиклевый русский флективен.

Сравнение данных по дистрибуции существительных, прилагательных и глаголов с результатами (1) О. Д. Митрофановой (1973) в научных текстах и (2) О. Б. Сиротининой (2009) по предметным областям не представляется возможным, т.к. в известных нам публикациях они представлены только в относительных величинах (процентах). Однако в целом можно утверждать, что изученные нами тексты уровня III, в большей степени приближенные к текстам, изученным российскими учеными ранее, имеют значения, сопоставимые с опубликованными О. Д. Митрофановой (1973) и О. Б. Сиротининой (2009). В письменных общенаучных текстах дистрибуция существительных, по материалам О. В. Сиротининой (2009) составляет 374 вхождения на 1000 словоупотреблений, а имени прилагательного – 152. Выявленные нами метрики аналогичны: доля имен существительных колеблется в диапазоне от 341 до 403, имя прилагательное – от 118 до 163 (см. Рисунок 3, 4). В целом такие различия могут быть детерминированы тем, что изученные нами тексты имеют предметную специализацию, а изученные О. В. Сиротининой (2009) тексты являются общенаучными.

Что касается дистрибуции глагола, то здесь наши данные значительно выше полученных М. Н. Кожинной (1972). В

биологических текстах уровня III дистрибуция глагола демонстрирует минимальное значение и составляет 123,5 вхождений на 1000 словоупотреблений, а текстах по обществознанию – 116 вхождений. В научных текстах, изученных М. Н. Кожинной (1972), дистрибуция глагола составляет только 90 вхождений на 1000 словоупотреблений. Данные различия косвенно указывают на более высокий уровень номинативности и, как следствие, степени сложности научных текстов. Учебным текстам, напротив, за счет большего присутствия в их составе глаголов, свойственна нарративность, рассчитываемая как отношение глаголов к существительным (Vahrusheva et al. 2023).

Говоря о процессе, номинализации необходимо отметить положительную динамику роста именных суффиксов в текстах обеих дисциплин. Это в целом подтверждает выводы, сделанные А. С. Мамонтовым и Е. Ю. Мамонтовой (2012) относительно частотности суффикса *-ние* в публицистических текстах. Вторую по частотности позицию после суффикса *-ние* в публицистических текстах (газета Московский комсомолец), изученных А.С. Мамонтовым и Е.Ю. Мамонтовой (2012), занимает суффикс *-ция*. В нашем корпусе по обществознанию суффикс *-ция* показывает стабильный рост от уровня I к уровню III, однако его доля уступает частотности суффиксов *-ость* и *-ство*: он занимает только четвертую строчку среди суффиксов субстантивации. Что касается *-ция* в текстах по биологии, то там он занимает четвертое место на уровне I, третье – на уровне II, а в текстах уровня III – уже второе место. В целом эти данные согласуются с результатами исследования А. С. Мамонтова и Е. Ю. Мамонтовой (2012). Статистический анализ демонстрирует, что преимущественное большинство изученных нами параметров коррелируют с индексом читабельности, т.е. могут использоваться в качестве предикторов сложности текстов.

### Заключение и перспектива исследования

В данном исследовании были изучены школьные учебники с целью (1) выявления предикторов морфологической сложности и (2) идентификации маркеров, дифференцирующих предметные области учебного текста на русском языке. В корпус вошли двадцать учебников по биологии и обществознанию трех уровней сложности, соответствующих 6 – 7, 8 – 9, 10 – 11 классам общеобразовательной российской школы, общим объемом 941963 словоформы. Анализ текстов школьных учебников по обществознанию и биологии на русском языке показал стабильную динамику роста индекса читабельности текстов от уровня к уровню и положительную взаимозависимость индекса читабельности учебно-научного текста, с одной стороны, и морфологических параметров, включая дистрибуцию существительных, прилагательных и долей существительных, образованных девербацией и деадъективацией, с другой. Индекс читабельности текстов имеет негативную корреляцию с дистрибуцией глагола. Морфологический анализ продемонстрировал высокий уровень номинативности текстов и динамику стабильного роста субстантивации в текстах за счет суффиксов девербации и деадъективации. Диапазоны значений лексической плотности и их динамика в текстах различных уровней сложности, а также дистрибуция существительных, прилагательных и отдельных суффиксов субстантивации выявили различия метрик учебников двух дисциплин, которые могут быть использованы в качестве маркеров предметной дифференциации при анализе больших языковых данных. Перспектива исследования видится в изучении долей сложных терминов-лексем, образованных от латинских и греческих корней, поскольку для биологии такой тип терминообразования является одним из наиболее активных. Валидация данного

параметра в корпусах разных предметных областей способна продемонстрировать его корреляцию с лингвистической сложностью текстов и дифференцировать тексты разных предметных областей.

#### Список источников

Обществознание. 6 класс: учеб. для общеобразоват. учреждений / Н. Ф. Виноградова, Н. И. Городецкая, Л. Ф. Иванова и др.; под ред. Л. Н. Боголюбова, Л. Ф. Ивановой. М.: Просвещение, 2012. 111 с.

Обществознание. 7 класс: учеб. для общеобразоват. учреждений с приложением на электронном носителе / Л. Н. Боголюбов, Н. И. Городецкая, Л. Ф. Иванова и др.; под ред. Л. Н. Боголюбова, Л. Ф. Ивановой. М.: Просвещение, 2013.

Боголюбов Л. Н. Обществознание. 8 класс: учебник для общеобразоват. учреждений. Под ред. Боголюбова Л. Н., Городецкой Н. И. М.: Просвещение, 2010. 255 с.

Боголюбов Л. Н. Обществознание. 9 класс: учебник для общеобразоват. учреждений. Под ред. Боголюбова Л. Н. М.: Просвещение, 2014. 255 с.

Обществознание. 10 класс: учеб. для общеобразоват. организаций: базовый уровень / Л. Н. Боголюбов, Ю. А. Аверьянов, А. В. Белявский и др.; под ред. Л. Н. Боголюбова и др. М.: Просвещение, 2014. 350 с.

Обществознание. 10 класс: учебник для общеобразовательных организаций: профильный уровень / Л. Н. Боголюбов и др.; под ред. Л. Н. Боголюбова, А. Ю. Лазебниковой, Н. М. Смирновой. 8-е изд. М.: Просвещение, 2014. 415 с.

Обществознание. 11 класс: учебник для учащихся общеобразовательных учреждений: профильный уровень / Л. Н. Боголюбов и др.; под ред. Л. Н. Боголюбова, А. Ю. Лазебниковой, К. Г. Холодковского. 3-е изд. М.: Просвещение, 2010. 414 с.

Обществознание. 6 класс: учебник для общеобразовательных учреждений / А. Ф. Никитин. 4-е изд., стер. М.: Дрофа, 2011. 125 с.

Обществознание. 7 класс: учебник / А. Ф. Никитин. 6-е изд., стер. М.: Дрофа, 2014. 94 с.

Никитин А. Ф., Никитина Т. И. Обществознание. 8 класс. Учебник. М.: Дрофа, 2014. 256 с.

Никитин А. Ф., Никитина Т. И. Обществознание. 9 класс. Учебник. М.: Дрофа, 2014. 208 с.

Обществознание. 10 класс: базовый уровень: учебник / А. Ф. Никитин. 10-е изд., стер. М.: Дрофа, 2014. 238 с.

Обществознание. 11 класс: базовый уровень: учебник / А. Ф. Никитин. 6-е изд., стер. М.: Дрофа, 2013. 237 с.

Сухорукова Л. Н. Биология. Живой организм. 5-6 классы: учеб. для общеобразоват. организаций / Л. Н. Сухорукова, В. С. Кучменко, И. Я. Колесникова. 3-е изд. М.: Просвещение, 2014. 143 с.

Сухорукова Л. Н. Биология. Разнообразие живых организмов. 7 класс: учеб. для общеобразоват. Организаций / Л. Н. Сухорукова, В. С. Кучменко, И. Я. Колесникова. М.: Просвещение, 2014. 159 с.

Сухорукова Л. Н. Биология. Человек. Культура здоровья: учеб. для 8 кл. общеобразоват. учреждений / Л. Н. Сухорукова, В. С. Кучменко, Т. А. Цехмистренко. М.: Просвещение, 2009. 160 с.

Сухорукова Л. Н. Биология. Живые системы и экосистемы. 9 класс: учеб. для общеобразоват. учреждений / Л. Н. Сухорукова, В. С. Кучменко. М.: Просвещение, 2010. 143 с.

Сухорукова Л. Н. Биология. 10-11 классы: учеб. для общеобразоват. учреждений / Л. Н. Сухорукова, В. С. Кучменко, Т. В. Иванова. М.: Просвещение, 2011. 127 с.

Биология. Бактерии, грибы, растения. 6 кл.: учеб. для общеобразоват. учреждений / В. В. Пасечник. 14-е изд., стереотип. М.: Дрофа, 2011. 304 с.

Биология. 7 класс: учеб. для общеобразоват. организаций / В. В. Пасечник, С. В. Суматохин, С. Г. Калинова; под ред. В. В. Пасечника. 3-е изд. М.: Просвещение, 2014. 256 с.

Биология. 8 класс: учеб. для общеобразоват. учреждений / В. В. Пасечник, А. А. Каменский, Г. Г. Швецов ; под ред. В. В. Пасечника. М.: Просвещение, 2010. 255 с.

Биология. Введение в общую биологию и экологию: Учеб. для 9 кл. общеобразоват. учеб. заведений / А. А. Каменский, Е. А. Криксунов, В. В. Пасечник. 3-е изд., стереотип. М.: Дрофа, 2002. 304 с.

Общая биология. 10-11 класс: учеб. для общеобразоват. учреждений / А. А. Каменский, Е. А. Криксунов, В. В. Пасечник. М.: Дрофа, 2005. 367 с.

### Corpus Materials

Vinogradova, N. F., Gorodetskaya, N. I., Ivanova, L. F. et al. (2012) *Obshchestvoznaniye*. 6 klass: ucheb. dlya obshcheobrazovatelnykh uchrezhdeniy [Social science. 6th grade: textbook for general educational institutions], in Bogolyubov, L. N., Ivanova, L. F. (eds.), *Prosvescheniye*, Moscow, Russia. (*In Russian*)

Vinogradova, N. F., Gorodetskaya, N. I., Ivanova, L. F. et al. (2013) *Obshchestvoznaniye*. 7 klass: ucheb. dlya obshcheobrazovatelnykh uchrezhdeniy [Social science. 7th grade: textbook for general educational institutions], in Bogolyubov, L. N., Ivanova, L. F. (eds.), *Prosvescheniye*, Moscow, Russia. (*In Russian*)

Bogolyubov, L. N., Gorodetskaya, N. I. (2010) *Obshchestvoznaniye*. 8 klass: ucheb. dlya obshcheobrazovatelnykh uchrezhdeniy [Social science. 8th grade: textbook for general educational institutions], *Prosvescheniye*, Moscow, Russia. (*In Russian*)

Bogolyubov, L. N. (2014) *Obshchestvoznaniye*. 9 klass: ucheb. dlya obshcheobrazovatelnykh uchrezhdeniy [Social science. 9th grade: textbook for general educational Institutions], *Prosvescheniye*, Moscow, Russia. (*In Russian*)

Bogolyubov, L. N. (2014) *Obshchestvoznaniye*. 10 klass: ucheb. dlya obshcheobrazovatelnykh uchrezhdeniy [Social science. 10th grade: textbook for general educational institutions: general level], in L. N. Bogolyubov, Yu. A. Averyanov, A. V. Belyavsky (eds.), *Prosvescheniye*, Moscow, Russia. (*In Russian*)

Bogolyubov, L. N. (2014) *Obshchestvoznaniye*. 10 klass: ucheb. dlya obshcheobrazovatelnykh organizatsiy: profilnyy uroven [Social science. 10th grade: textbook for general education organizations: advanced level], in Bogolyubov, L. N., Lazebnikova, A. Yu., Smirnov, N. M. (eds.), 8th ed., *Prosvescheniye*, Moscow, Russia. (*In Russian*)

Bogolyubov, L. N. (2010) *Obshchestvoznaniye*. 11 klass: ucheb. dlya obshcheobrazovatelnykh organizatsiy: profilnyy uroven [Social science. 11th grade: textbook for students of general education institutions: profile level], in Bogolyubov, L. N., Lazebnikova, A. Yu., Kholodkovsky, K. G. (eds.), 3rd ed., *Prosvescheniye*, Moscow, Russia. (*In Russian*)

Nikitin, A. F. (2011) *Obshchestvoznaniye*. 6 klass: ucheb. dlya obshcheobrazovatelnykh uchrezhdeniy [Social science. 6th grade: textbook for educational institutions], 4th ed., Drofa, Moscow, Russia. (*In Russian*)

Nikitin, A. F. (2014) *Obshchestvoznaniye*. 7 klass: ucheb. dlya obshcheobrazovatelnykh uchrezhdeniy [Social science. 7th grade: textbook], 6th ed., Drofa, Moscow, Russia. (*In Russian*)

Nikitin, A. F. (2014) *Obshchestvoznaniye*. 8 klass: ucheb. dlya obshcheobrazovatelnykh uchrezhdeniy [Social science. 8th grade. Textbook], Drofa, Moscow, Russia. (*In Russian*)

Nikitin, A. F. (2014) *Obshchestvoznaniye*. 9 klass: ucheb. dlya obshcheobrazovatelnykh uchrezhdeniy [Social science. 9th grade. Textbook], Drofa, Moscow, Russia. (*In Russian*)

Nikitin, A. F. (2014) *Obshchestvoznaniye*. 10 klass: ucheb. dlya obshcheobrazovatelnykh uchrezhdeniy [Social science. 10th grade: textbook for general education organizations: general level], 10th ed., Drofa, Moscow, Russia. (*In Russian*)

Nikitin, A. F. (2013) *Obshchestvoznaniye*. 11 klass: ucheb. dlya obshcheobrazovatelnykh uchrezhdeniy [Social science. Grade 11: textbook for general education organizations: general level], 6th ed., Drofa, Moscow, Russia. (*In Russian*)

Sukhorukova, L. N., Kuchmenko, V. S., Kolesnikova, I. Ya. (2014) *Biologiya. Zhivoy organizm*. 5-6 klassy: ucheb. dlya obshcheobrazovatelnykh organizatsiy [Biology. Living organism. Grades 5-6: textbook for general educational institutions], 3rd ed., *Prosvescheniye*, Moscow, Russia. (*In Russian*)

Sukhorukova, L. N., Kuchmenko, V. S., Kolesnikova, I. Ya. (2014) *Biologiya. Raznoobraziye zhivyykh organizmov*. 7 klass: ucheb. dlya obshcheobrazovatelnykh organizatsiy [Biology. Diversity of living organisms. 7th grade: educational. for general education organizations], *Prosvescheniye*, Moscow, Russia. (*In Russian*)

Sukhorukova, L. N., Kuchmenko, V. S., Tsekhmistrenko, T. A. (2009) *Biologiya. Chelovek. Kultura zdorovya: ucheb. dlya 8 kl. obshcheobrazovat. uchrezhdeniy* [Biology. Human. Culture of health. 8th grade: textbook for general educational institutions], Prosvescheniye, Moscow, Russia. (*In Russian*)

Sukhorukova, L. N., Kuchmenko, V. S. (2010) *Biologiya. Zhivyye sistemy i ekosistemy. 9 klass: ucheb. dlya obshcheobrazovat. uchrezhdeniy* [Biology. Living systems and ecosystems. 9th grade: textbook for general educational institutions], Prosvescheniye, Moscow, Russia. (*In Russian*)

Sukhorukova, L. N., Kuchmenko, V. S., Ivanova, T. V. (2011) *Biologiya. 10-11 klassy: ucheb. dlya obshcheobrazovat. uchrezhdeniy* [Biology. Grades 10-11: textbook for general educational institutions], Prosvescheniye, Moscow, Russia. (*In Russian*)

Pasechnik, V. V. (ed.) (2011) *Biologiya. Bakterii, griby, rasteniya. 6 kl.: ucheb. dlya obshcheobrazovat. uchrezhdeniy* [Biology. Bacteria, fungi, plants. 6th grade: textbook for general educational institutions], 14th ed., Drofa, Moscow, Russia. (*In Russian*)

Pasechnik, V. V., Sumatokhin, S. V., Kalinova, S. G. (2014) *Biologiya. 7 klass: ucheb. dlya obshcheobrazovat. organizatsiy* [Biology. 7th grade: textbook for general educational institutions], 3rd ed., Prosvescheniye, Moscow, Russia. (*In Russian*)

Pasechnik, V. V., Kamensky, A. A., Shvetsov, G. G. (2010) *Biologiya. 8 klass: ucheb. dlya obshcheobrazovat. organizatsiy* [Biology. 8th grade: textbook for general educational institutions], Prosvescheniye, Moscow, Russia. (*In Russian*)

Kamensky, A. A., Kriskunov, E. A., Pasechnik, V. V. (2002) *Biologiya. Vvedeniye v obshchuyu biologiyu i ekologiyu: Ucheb. dlya 9 kl. obshcheobrazovat. ucheb. zavedeniy* [Biology. Introduction to general biology and ecology: 9th grade: textbook for general educational institutions], 3rd ed., Drofa, Moscow, Russia. (*In Russian*)

Kamensky, A. A., Kriskunov, E. A., Pasechnik, V. V. (2005) *Obshchaya biologiya. 10-11 klass: ucheb. dlya obshcheobrazovat. uchrezhdeniy* [General biology. 10-11 grade: textbook for general educational institutions], Drofa, Moscow, Russia. (*In Russian*)

## Список литературы

Валгина Н. С. Теория текста. М.: Логос, 2003. 280 с.

Жеребило Т. В. Словарь лингвистических терминов. Изд. 5-е, испр. и доп. Назрань: ООО «Пилигрим», 2010. 486 с.

Кожина М. Н. О речевой системности научного стиля сравнительно с некоторыми другими / М. Н. Кожина. Пермь: Изд-во Пермского гос. ун-та, 1972. 395 с.

Мамонтов А. С., Мамонтова Е. Ю. Лингвостатистический анализ использования девербативов в различных видах текстов русского и английского языков // Вестник Российского университета дружбы народов. Серия: Лингвистика. М.: РУДН. 2012. №2. С. 107–110.

Мацковский М. С. Проблемы читабельности печатного материала / М. С. Мацковский // Смысловое восприятие речевого сообщения в условиях массовой коммуникации; отв. ред. Т. М. Дридзе, А. А. Леонтьев. М.: Наука, 1976. С. 126–142.

Микк Я. А. О факторах понятности учебного текста: автореф. дисс. ... канд. пед. наук: 13.00.00. Тарту, 1970. 22 с.

Митрофанова О. Д. Язык научно-технической литературы. М.: Изд-во Моск. ун-та. 1973. 147 с.

Оборнева И. В. Автоматизированная оценка сложности учебных текстов на основе статистических параметров: дисс. ... канд. пед. наук.: 13.00.02. М. 2006. 65 с.

Петрова Н. Е. Субстантивация и деадвербиализация наречий в современном русском языке / Н. Е. Петрова // Известия Уральского государственного университета. Сер. 2, Гуманитарные науки. 2009. № 1/2 (63). С. 35–42.

Разговорная речь в системе функциональных стилей современного русского литературного языка. Грамматика: / О. Б. Сиротинина, Э. А. Столярова, Н. Г. Мартыненко и др.; под ред. О. Б. Сиротининой. 2. изд., стер. Москва, УРСС. 2009. 309 с.

Шпаковский Ю. Ф. Оценка трудности восприятия и оптимизация сложности учебного текста (на материале текстов по химии): автореф. дисс. ... канд. филол. наук: 10.02.19. Минск, 2007. 22 с.

Alderson J. Ch. Assessing reading. Cambridge University Press, 2000. 398 p.

- Banks D. On the historical origins of nominalized process in scientific text // *English for Specific Purposes*. 2005. Vol. 24. Iss. 3. Pp. 347–367. DOI: 10.1016/j.esp.2004.08.002
- Banks D. *The Development of Scientific Writing: Linguistic Features and Historical Context*. London: Equinox, 2008. 221 p.
- Biber D., Gray B. *Grammatical complexity in academic English: Linguistic change in writing*. Cambridge University Press, 2016. 276 p.
- Biber D., Gray B. Nominalising the verb phrase in academic scientific writing / D. Biber and B. Gray // *The Verb Phrase in English* / Bas Aaarts et al. (eds.). London. Cambridge University Press. 2013. Pp. 99–132.
- Biber D. *Grammar of spoken and written English* / D. Biber, S. Johansson, G. Leech, S. Conrad and E. Finegan. John Benjamins, Amsterdam. 2021. 1220 p.
- Biber D. *The Longman grammar of spoken and written English* / D. Biber, S. Johansson, G. Leech, S. Conrad and E. Finegan. Longman, London. 1999. 1204 p.
- Biber D. *A typology of English texts* // *Linguistics*. Vol. 27. Pp. 3–43.
- Crossley S. A. et al. *Toward a new readability: A mixed model approach* / S. A. Crossley et al. // *Proceedings of the 29th annual conference of the Cognitive Science Society* / D. S. McNamara and G. Trafton (eds.). Cognitive Science Society, 2007. Pp. 197–202.
- DuBay W. H. *The Classic Readability Studies*. Costa Mesa, 2007. 240 p. URL: <https://files.eric.ed.gov/fulltext/ED506404.pdf> (Accessed 21 July 2024)
- Egins S. *An introduction to systemic functional linguistics*. London, 2004. 384 p.
- Fang Z., Schleppegrell M. J., Cox B. E. *Understanding the Language Demands of Schooling: Nouns in Academic Registers* // *Journal of Literacy Research*. 2006. Vol. 38. Iss. 3. Pp. 247–273. [https://doi.org/10.1207/s15548430jlr3803\\_1](https://doi.org/10.1207/s15548430jlr3803_1)
- Flesch R. *A new readability yardstick* // *Journal of Applied Psychology*. Vol. 32. 1948. P. 221–233.
- Gatiyatullina G. *Computing Russian Morphological distribution patterns using RusAC Online Server* / G. Gatiyatullina, M. Solnyshkina, V. Solovyev, A. Danilov, E. Martynova, I. Yarmakeev // *13th International Conference on Developments in eSystems Engineering (DeSE)*. 2020. Pp. 393–398. DOI: 10.1109/DeSE51703.2020.9450753
- Halliday M. A. K., Martin J. R. (eds.) *Writing Science: Literacy and Discursive Power*. London: Falmer Press, 2003. 434 p.
- Just M. A., Carpenter P. A. *The psychology of reading and language comprehension*. Boston; London: Allyn and Bacon. 1987. 518 p.
- Kincaid J. P., Fishburne R. P., Rogers R. L., Chissom B. S. *Derivation of new readability formulas (Automated Readability Index, Fog Count, and Fesch Reading Ease Formula) for Navy enlisted personnel* // *Research Branch Report*. 1975. Pp. 8–75.
- Koda K. *Insights into second language reading: A cross-linguistic approach*. Cambridge: Cambridge University Press. 2005. 320 p.
- Martin J. R. *Nominalization in science and humanities: Distilling knowledge and scaffolding text* // *Functional and Systemic Linguistics: Approaches and Uses* / Eija Ventola ed. Berlin, New York: Mouton de Gruyter, 1991. Pp. 307–338. <https://doi.org/10.1515/9783110883527.307>
- McLaughlin G. H. *SMOG Grading – A new readability formula* // *Journal of Reading*, 1969. Vol. 12. Iss. 8. Pp. 639–646.
- Solnyshkina M., Ivanov V., Solovyev V. *Readability formula for Russian texts: A modified version* // *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2018. Pp. 132–145. DOI: 10.1007/978-3-030-04497-8\_11
- Solnyshkina M. I., Kiselnikov A. S. *Text complexity: Study phases in Russian linguistics* // *Vestnik Tomskogo Gosudarstvennogo Universiteta, Filologiya*. 2015. Vol. 38 (6). Pp. 86–99. DOI: 10.17223/19986645/38/7
- To V., Fan S., Thomas D. P. *Lexical density and readability: A case study of English textbooks* // *The International Journal of Language, Society and Culture*. 2013. Vol. 37. Iss. 7. Pp. 61–71.
- Vahrusheva A., Solovyev V., Solnyshkina M., Gafiaytova E., Akhtyamova S. *Revisiting Assessment of Text Complexity: Lexical and Syntactic Parameters Fluctuations* // *Speech and Computer. SPECOM 2023. Lecture Notes in Computer Science* / A. Karpov, K. Samudravijaya, K. T. Deepak, R. M. Hegde, S. S. Agrawal, S. R. M. Prasanna, (eds). Springer, 2023. Vol. 14338. [https://doi.org/10.1007/978-3-031-48309-7\\_35](https://doi.org/10.1007/978-3-031-48309-7_35)

## References

- Valgina, N. S. (2003). *Teoriya teksta [Theory of text]*, Logos, Moscow, Russia. (In Russian)
- Jerebilo, T. V. (2010). *Slovar lingvisticheskikh terminiv [A Dictionary of linguistic terms]*. Issue 5, revised and added, Piligrim, Nazran, Russia. (In Russian)
- Kozhina, M. N. (1972). *O rechevoy sistemnosti nauchnogo stilya sravnitel'no s nekotorymi drugimi [On the speech systematicity of scientific style in comparison with some others]* Perm State University Publishing House, Perm, Russia. (In Russian)
- Mamontov, A. S., Mamontova, Ye. Ye. (2012). *Lingvostatisticheskiy analiz ispol'zovaniya deverbativov v razlichnykh vidakh tekstov russkogo i angliyskogo yazykov [Linguistic and statistical analysis of the use of deverbatives in various types of texts in Russian and English]*, *Vestnik RUDN*, No. 2., 107–110. (In Russian)
- Matskovskiy, M. S. (1976). *Problemy chitabel'nosti pechatnogo materiala [Problems of readability of printed material]*, in Dridze, T. M. and Leontyev, A. A. (eds.) *Meaningful perception of a speech message in the conditions of mass communication*, Nauka, Moscow, 126–142. (In Russian)
- Mikk, Ya. A. (1970). *O faktorakh ponyatnosti uchebnogo teksta [On the factors of understandability of educational text]*, Abstract of Ph.D dissertation, University of Tartu, USSR. (In Russian)
- Mitrofanova, O. D. (1973). *Yazyk nauchno-tekhnicheskoy literatury [Language of scientific and technical literature]*, Moscow University Press, Moscow, Russia. (In Russian)
- Oborneva, I. V. (2006). *Avtomatizirovannaya otsenka slozhnosti uchebnykh tekstov na osnove statisticheskikh parametrov [Automated assessment of the complexity of educational texts based on statistical parameters]*, Abstract of PhD dissertation, Moscow, Russia. (In Russian)
- Petrova, N. Ye. (2009). *Substantivatsiya i deadverbializatsiya narechiy v sovremennom russkom yazyke [Substantivization and deadverbialization of adverbs in the modern Russian language]*, *Vestnik Uralskogo gosudarstvennogo universiteta: Seriya gumanitarnye nauki*, 1/2 (63), 35–42. (In Russian)
- Sirotnina, O. B. (ed.) (2009). *Razgovornaya rech' v sisteme funktsional'nykh stiley sovremennogo russkogo literaturnogo yazyka. Grammatika: [Colloquial speech in the system of functional styles of the modern Russian literary language: Grammar]*, Moscow, USSR. (In Russian)
- Shpakovskiy, Yu. F. (2007). *Otsenka trudnosti vospriyatiya i optimizatsiya slozhnosti uchebnogo teksta (na materiale tekstov po khimii) [Assessing the difficulty of perception and optimizing the complexity of educational text (based on texts in chemistry)]*, Abstract of Ph.D dissertation, Minsk, Belarus. (In Russian)
- Alderson, J. Ch. (2000). *Assessing reading*, Cambridge University Press, London, UK. (In English)
- Banks, D. (2005). *On the historical origins of nominalized process in scientific text*, *English for Specific Purposes* 24/3, 347–367. DOI: 10.1016/j.esp.2004.08.002 (In English)
- Banks, D. (2008). *The Development of Scientific Writing: Linguistic Features and Historical Context*, Equinox, London, UK. (In English)
- Biber, D. and Gray, B. (2016). *Grammatical complexity in academic English: Linguistic change in writing*, Cambridge University Press, Cambridge, UK. (In English)
- Biber, D., Gray, B. (2013). *Nominalising the verb phrase in academic scientific writing*, in Bas Aaarts et al. (Eds.), *The Verb Phrase in English*, Cambridge University Press, Cambridge, UK. (In English)
- Biber, D., Johansson, S., Leech, G., Conrad, S. and Finegan, E. (2021). *Grammar of spoken and written English*, John Benjamins, Amsterdam, Netherlands. <https://doi.org/10.1075/z232> (In English)
- Biber, D., Johansson, S., Leech, G., Conrad, S. and Finegan, E. (1999). *The Longman grammar of spoken and written English*, Longman, London, England. (In English)
- Biber, D. (1989). *A typology of English texts*, *Linguistics*, 27, 3–43. (In English)
- Crossley, S. A. et al. (2007). *Toward a new readability: A mixed model approach in McNamara, D. S. Trafton, G. (eds.) Proceedings of the 29th annual conference of the Cognitive Science Society*; Austin, TX: Cognitive Science Society, 197–202. (In English)
- DuBay, W. H. (2007). *The Classic Readability Studies*. Costa Mesa. 240 p.

URL: <https://files.eric.ed.gov/fulltext/ED506404.pdf> (Accessed 21 July 2024) (In English)

Eggins, S. (2004). *An introduction to systemic functional linguistics*, London, UK. (In English)

Fang, Z., Schleppegrell, M. J. and Cox, B. E. (2006). Understanding the Language Demands of Schooling: Nouns in Academic Registers, *Journal of Literacy Research*, 38 (3), 247–273.

[https://doi.org/10.1207/s15548430jlr3803\\_1](https://doi.org/10.1207/s15548430jlr3803_1) (In English)

Flesch, R. (1948). A new readability yardstick, *Journal of Applied Psychology*, 32, 221–233. (In English)

Gatiyatullina, G., Solnyshkina, M., Solovyev, V., Danilov, A., Martynova, E. and Yarmakeev, I. (2020). Computing Russian Morphological distribution patterns using RusAC Online Server, *13th International Conference on Developments in eSystems Engineering (DeSE)*, Liverpool, UK, 393–398. DOI: 10.1109/DeSE51703.2020.9450753 (In English)

Halliday, M. A. K. and Martin, J. R. (eds.) (2003). *Writing Science: Literacy and Discursive Power*, Falmer Press, London, UK. (In English)

Just, M. A., Carpenter, P. A. (1987). *The psychology of reading and language comprehension*, Allyn and Bacon, Boston, London. (In English)

Kincaid, J. P., Fishburne, R. P., Rogers, R. L. and Chissom, B. S. (1975). Derivation of new readability formulas (Automated Readability Index, Fog Count, and Fesch Reading Ease Formula) for Navy enlisted personnel, *Research Branch Report*, 8–75. (In English)

Koda, K. (2005). *Insights into second language reading: A cross-linguistic approach*, Cambridge University Press, Cambridge, UK. (In English)

Martin, J. R. (1991). *Nominalization in science and humanities: Distilling knowledge and scaffolding, Functional and systemic linguistics: Approaches and Uses*, De Gruyter Mouton, Berlin, New York, 307–338. <https://doi.org/10.1515/9783110883527.307> (In English)

McLaughlin, G. H. (1969). SMOG Grading – A new readability formula, *Journal of Reading*, 12(8), 639–646. (In English)

Solnyshkina, M., Ivanov, V., Solovyev, V. (2018). Readability formula for Russian texts: A

modified version in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 132–145. DOI: 10.1007/978-3-030-04497-8\_11 (In English)

Solnyshkina, M. I., Kisel'nikov, A. S. (2015). Text complexity: Study phases in Russian linguistics, *Vestnik Tomskogo Gosudarstvennogo Universiteta, Filologiya*, 38 (6), 86–99. DOI: 10.17223/19986645/38/7 (In English)

To, V., Fan, S., Thomas, D. P. (2013). Lexical density and readability: A case study of English textbooks. *The International Journal of Language, Society and Culture*, 37(7), 61–71. (In English)

Vahrusheva, A., Solovyev, V., Solnyshkina, M., Gafiyatova, E., Akhtyamova, S. (2023). Revisiting Assessment of Text Complexity: Lexical and Syntactic Parameters Fluctuations, in Karpov, A., Samudravijaya, K., Deepak, K. T., Hegde, R. M., Agrawal, S. S., Prasanna, S. R. M. (eds) *Speech and Computer. SPECOM 2023. Lecture Notes in Computer Science*, Vol 14338, Springer, Cham, Switzerland. [https://doi.org/10.1007/978-3-031-48309-7\\_35](https://doi.org/10.1007/978-3-031-48309-7_35) (In English)

**Все авторы прочитали и одобрили окончательный вариант рукописи.**

**All authors have read and approved the final manuscript.**

**Конфликты интересов: у авторов нет конфликтов интересов для декларации.**

**Conflicts of interests: the authors have no conflicts of interest to declare.**

**Галия Маратовна Гатиятуллина**, старший преподаватель кафедры теории и практики преподавания иностранных языков, младший научный сотрудник, НИЛ «Текстовая аналитика», Институт филологии и межкультурной коммуникации, Казанский (Приволжский) федеральный университет, Казань, Россия.

**Galiya M. Gatiyatullina**, Senior Lecturer, the Department of Theory and Practice of Teaching Foreign Languages, Research Assistant, Text Analytics Research Laboratory, Institute of Philology and Intercultural Communication, Kazan Federal University, Kazan, Russia.

**Эльзара Васильевна Гафиятова**, доктор филологических наук, доцент, заведующая кафедрой теории и практики преподавания иностранных языков, заместитель директора по международной деятельности и развитию Института филологии и межкультурной коммуникации, Казанский (Приволжский) федеральный университет, Казань, Россия.

**Elzara V. Gafiyatova**, Doctor of Philology, Associate Professor, Head of the Department of Theory and Practice of Teaching Foreign Languages, Vice Director of the Institute of Philology and Intercultural Communication, Kazan Federal University, Kazan, Russia.

**Андрей Владимирович Данилов**, кандидат педагогических наук, доцент кафедры билингвального и цифрового образования, старший научный сотрудник НИЛ «Текстовая аналитика», Институт филологии и межкультурной коммуникации, Казанский (Приволжский) федеральный университет, Казань, Россия.

**Andrei V. Danilov**, PhD in Education, Associate Professor, Department of Bilingual and Digital Education, Senior Researcher at the Text Analytics Research Laboratory, Institute of Philology and Intercultural Communication, Kazan Federal University, Kazan, Russia.